Asymptotic Log-Harnack Inequality and Applications for Stochastic Systems of Infinite Memory*

Jianhai Bao[b),c)], **Feng-Yu Wang**[a),c)], **Chenggui Yuan**[c)]

[a)]Center of Applied Mathematics, Tianjin University, Tianjin 300072, China

[b)]School of Mathematics and Statistics, Central South University, Changsha 410083, China

[c)]Department of Mathematics, Swansea University, Singleton Park, SA2 8PP, UK

jianhai.bao@swansea.ac.uk, wangfy@tju.edu.cn, C.Yuan@swansea.ac.uk

Abstract

The asymptotic log-Harnack inequality is established for several kinds of models on stochastic differential systems with infinite memory: non-degenerate SDEs, neutral SDEs, semi-linear SPDEs, and stochastic Hamiltonian systems. As applications, the following properties are derived for the associated segment Markov semigroups: asymptotic heat kernel estimate, uniqueness of the invariant probability measure, asymptotic gradient estimate (hence, asymptotically strong Feller property), as well as asymptotic irreducibility.

AMS Subject Classification: 60H10, 47G20
Keywords: Asymptotic log-Harnack inequality; asymptotic gradient estimate; asymptotic heat kernel; asymptotic irreducibility

1 Introduction

The dimension-free Harnack inequality was initiated in [15] for elliptic diffusion semigroups on Riemannian manifolds. In case such kind of inequality is unavailable, the log-Harnack inequality was introduced alternatively in [17]. Both inequalities have been investigated extensively and applied to (singular, degenerate) SDEs/SPDEs via coupling by change of measures developed in e.g. [1, 16]; see [19] and references within for more details. In particular, these inequalities imply gradient estimates (hence, the strong Feller property),

*Supported in part by NNSFC (11771326, 11431014,11831014) and a Co-Fund grant.

the uniqueness of invariant probability measures, heat kernel estimates, and irreducibility of the associated Markov semigroups.

However, when the stochastic system is highly degenerate so that these properties are unavailable, the above type Harnack inequalities no longer hold. In this scenario, it is natural to investigate weaker versions of these properties by exploiting Harnack inequalities in the weak version. For instance, the strong Feller property is invalid for degenerate stochastic $2D$ Navier-Stokes equations, whereas the weaker "asymptotically strong Feller" property has been proved in [10] and [26] by making use of asymptotic couplings and "modified log-Harnack inequality", respectively. Since the log-Harnack inequality in the weak version is concerned with long time behavior, below we shall call it "asymptotic log-Harnack inequality".

In this paper, we aim to investigate asymptotic log-Harnack inequality and its applications for SDEs with infinite memory, i.e., the coefficients of SDEs involved depend on the whole history of the system. In this setup, the strong Feller property is invalid (see e.g. [5, 9]), so we are in the weak situation without log-Harnack inequalities. When the memory is finite and the noise is path-independent, the dimension-free Harnack inequality, log-Harnack inequality and gradient estimates have been investigated in [3, 4, 7, 13, 20], to name a few.

Before considering specific models, in Section 2 we present some applications of the asymptotic log-Harnack inequality in a general framework, which are new except the asymptotically strong Feller property derived in [26]. In Sections 3-6, we establish asymptotic log-Harnack inequality for the following stochastic differential systems with infinite memory, respectively, including non-degenerate SDEs, neutral SDEs, semi-linear SPDEs, and stochastic Hamiltonian systems. In the Appendix section, we address the existence and uniqueness of solutions to SDEs with infinite memory under the locally weak monotone condition and the weak coercive condition.

2 Applications of asymptotic log-Harnack inequality

Before we recall the definition on asymptotically strong Feller introduced in [10] for a Markov semigroup P_t, we start with some notation and notions. Let (E,ρ) be a metric space, $\mathscr{B}_b(E)$ the class of bounded measurable functions on E, and $\mathscr{B}_b^+(E)$ the set of positive functions in $\mathscr{B}_b(E)$. A continuous function $d: E\times E\to\mathbb{R}_+ := [0,\infty)$ is called a pseudo-metric if $d(x,x)=0$ and $d(x,y)\le d(x,z)+d(z,y)$ hold for $x,y,z\in E$. For a pseudo-metric d, the transportation cost (which also is called L^1-Wasserstein distance when d is a distance) is defined by

$$W_1^d(\mu_1,\mu_2)=\inf_{\pi\in\mathscr{C}(\mu_1,\mu_2)}\int_{E\times E} d(x,y)\pi(\mathrm{d}x,\mathrm{d}y),\quad \mu_1,\mu_2\in\mathscr{P}(E),$$

where $\mathscr{P}(E)$ stands for the class of probability measures on E, and $\mathscr{C}(\mu_1,\mu_2)$ consists of all couplings of μ_1 and μ_2, that is, $\pi\in\mathscr{C}(\mu_1,\mu_2)$ means $\pi\in\mathscr{P}(E\times E)$ with $\pi(\cdot\times E)=\mu_1$ and $\pi(E\times\cdot)=\mu_2$. An increasing sequence of pseudo-metrics $(d_n)_{n\ge1}$ (i.e., $d_i(\cdot,\cdot)\ge d_j(\cdot,\cdot), i\ge j$) is said to be a totally separating system if $\lim_{n\to\infty} d_n(x,y)=1$ for all $x\neq y$.

Definition 2.1. The Markov semigroup P_t is called asymptotically strong Feller at a point $x \in E$, if there exist a totally separating system of pseudo-metrics $(d_k)_{k\geq 1}$ and a sequence $t_k \uparrow \infty$ such that

$$\inf_{U\in\mathcal{U}_x} \limsup_{k\to\infty} \sup_{y\in U} W_1^{d_k}(P_{t_k}(x,\cdot), P_{t_k}(y,\cdot)) = 0, \tag{2.1}$$

where $\mathcal{U}_x$ denotes the collection of all open sets containing x, and $P_t(x,A) := P_t 1_A(x)$ for $x \in E$ and a measurable set $A \subset E$. P_t is called asymptotically strong Feller if it is asymptotically strong Feller at any $x \in E$.

For a function $f : E \to \mathbb{R}$, define

$$|\nabla f|(x) = \limsup_{y\to x} \frac{|f(x) - f(y)|}{\rho(x,y)}, \quad x \in E.$$

Let $\|\cdot\|_\infty$ be the uniform norm. So, $\|\nabla f\|_\infty = \sup_{x\in E} |\nabla f|(x)$. Set $\mathrm{Lip}(E) := \{f : E \to \mathbb{R}, \|\nabla f\|_\infty < \infty\}$, the family of all Lipschitzian functions on E. Next, we introduce the asymptotic log-Harnack inequality.

Definition 2.2. The following inequality is called an asymptotic log-Harnack inequality of P_t:

$$P_t \log f(x) \leq \log P_t f(y) + \Phi(x,y) + \Psi_t(x,y)\|\nabla \log f\|_\infty, \quad t > 0 \tag{2.2}$$

for any $f \in \mathscr{B}_b^+(E)$ with $\|\nabla \log f\|_\infty < \infty$, where $\Phi, \Psi_t : E \times E \to (0,\infty)$ are measurable with $\Psi_t \downarrow 0$ as $t \uparrow \infty$.

Below, we present some asymptotic properties implied by (2.2). For a measurable set $A \subset E$ and $x \in E$, let $\rho(x,A) = \inf_{y\in A} \rho(x,y)$, i.e., the distance between x and A. Moreover, for any $\varepsilon > 0$, let $A_\varepsilon = \{y \in E : \rho(y,A) < \varepsilon\}$ and A_ε^c be the complement of A_ε.

Theorem 2.1. *Let P_t satisfy (2.2) for some symmetric functions $\Phi, \Psi_t : E \times E \to \mathbb{R}_+$ with $\Psi_t \downarrow 0$ as $t \uparrow \infty$. Then:*

(1) **(Gradient estimate)** *If, for any $x \in E$,*

$$\Lambda(x) := \limsup_{y\to x} \frac{\Phi(x,y)}{\rho(x,y)^2} < \infty, \quad \textit{and} \quad \Gamma_t(x) := \limsup_{y\to x} \frac{\Psi_t(x,y)}{\rho(x,y)} < \infty, \tag{2.3}$$

then, for any $t > 0$ and $f \in \mathrm{Lip}_b(E) := \mathrm{Lip}(E) \cap \mathscr{B}_b(E)$,

$$|\nabla P_t f| \leq \sqrt{2\Lambda}\sqrt{P_t f^2 - (P_t f)^2} + \|\nabla f\|_\infty \Gamma_t. \tag{2.4}$$

In particular, when $\Gamma_t \downarrow 0$ as $t \uparrow \infty$, P_t is asymptotically strong Feller.

(2) **(Asymptotic heat kernel estimate)** *If P_t has an invariant probability measure μ, then, for any $f \in \mathscr{B}_b^+(E)$ with $\|\nabla f\|_\infty < \infty$,*

$$\limsup_{t\to\infty} P_t f(x) \le \log\left(\frac{\mu(\mathrm{e}^f)}{\int_E \mathrm{e}^{-\Phi(x,y)}\mu(\mathrm{d}y)}\right), \qquad x \in E. \tag{2.5}$$

Consequently, for any closed set $A \subset E$ with $\mu(A) = 0$,

$$\lim_{t\to\infty} P_t 1_A(x) = 0, \qquad x \in E. \tag{2.6}$$

(3) **(Uniqueness of invariant probability)** *P_t has at most one invariant probability measure.*

(4) **(Asymptotic irreducibility)** *Let $x \in E$ and $A \subset E$ be a measurable set such that*

$$\delta(x, A) := \liminf_{t\to\infty} P_t(x, A) > 0.$$

Then,

$$\liminf_{t\to\infty} P_t(y, A_\varepsilon) > 0, \quad y \in E, \varepsilon > 0. \tag{2.7}$$

Moreover, for any $\varepsilon_0 \in (0, \delta(x, A))$, there exists a constant $t_0 > 0$ such that

$$P_t(y, A_\varepsilon) > 0 \ \text{ provided } \ t \ge t_0, \quad \Psi_t(x, y) < \varepsilon\varepsilon_0. \tag{2.8}$$

According to the proof of [19, Theorem 1.4.1(4)], if (2.5) holds without limit but for a fixed $t > 0$, then P_t has a density $p_t(x, y)$ with respect to μ satisfying the entropy estimate

$$\int_E p_t(x, y) \log p_t(x, y)\mu(\mathrm{d}y) \le -\log \int_E \mathrm{e}^{-\Phi(x,y)}\mu(\mathrm{d}y).$$

So, (2.5) can be regarded as the asymptotic heat kernel estimate of P_t.

Proof of Theorem 2.1. (**1**) In terms of [26], if $\Gamma_t \downarrow 0$ as $t \uparrow \infty$, (2.4) implies the asymptotically strong Feller property. So, it suffices to prove the gradient estimate (2.4).

For any $x \in E, t > 0$ and $f \in \mathrm{Lip}_b(E)$, we take $x_n \to x$ such that $\varepsilon_n := \rho(x_n, x) \downarrow 0$ and (in case the limit below is negative, write $-f$ instead of f)

$$|\nabla P_t f|(x) = \limsup_{n\to\infty} \frac{P_t f(x_n) - P_t f(x)}{\varepsilon_n}. \tag{2.9}$$

For any constant $c > 0$, (2.2) implies

$$P_t \log(1+c\,\varepsilon_n f)(x_n) \le \log P_t(1+c\,\varepsilon_n f)(x) + \Phi(x_n, x) + \Psi_t(x_n, x)\|\nabla \log(1+c\varepsilon_n f)\|_\infty. \tag{2.10}$$

By Taylor's expansion, for ε_n sufficiently small we have

$$P_t \log(1 + c\,\varepsilon_n f)(x_n) = c\,\varepsilon_n P_t f(x_n) - \frac{c^2\varepsilon_n^2}{2} P_t f^2(x_n) + o(\varepsilon_n^2), \tag{2.11}$$

and

$$\log P_t(1 + c\,\varepsilon_n f)(x) = c\,\varepsilon_n P_t f(x) - \frac{c^2\varepsilon_n^2}{2} (P_t f)^2(x) + o(\varepsilon_n^2). \tag{2.12}$$

Substituting (2.11) and (2.12) into (2.10) yields

$$\begin{aligned} c\,\varepsilon_n(P_t f(x_n) - P_t f(x)) &\le \frac{c^2\varepsilon_n^2}{2}\Big\{P_t f^2(x_n) - (P_t f)^2(x)\Big\} \\ &\quad + \Phi(x_n, x) + \Psi_t(x_n, x)\|\nabla \log(1 + c\,\varepsilon_n f)\|_\infty + o(\varepsilon_n^2) \\ &\le \frac{c^2\varepsilon_n^2}{2}\Big\{P_t f^2(x_n) - (P_t f)^2(x)\Big\} + \Phi(x_n, x) + c\,\varepsilon_n \Psi_t(x_n, x)\|\nabla f\|_\infty + o(\varepsilon_n^2). \end{aligned}$$

Combining this with (2.9), we obtain

$$\begin{aligned} |\nabla P_t f|(x) &\le \limsup_{n\to\infty} \bigg(\frac{c}{2}\Big\{P_t f^2(x_n) - (P_t f)^2(x)\Big\} + \frac{\Phi(x_n, x)}{c\,\varepsilon_n^2} + \frac{\Psi_t(x_n, x)}{\varepsilon_n}\|\nabla f\|_\infty\bigg) \\ &\le \frac{c}{2}\Big\{\limsup_{n\to\infty} P_t f^2(x_n) - (P_t f)^2(x)\Big\} + \frac{\Lambda(x)}{c} + \|\nabla f\|_\infty \Gamma_t(x). \end{aligned}$$

This, in particular, implies $P_t \mathrm{Lip}_b(E) \subset C_b(E)$, so that $\limsup_{n\to\infty} P_t f^2(x_n) = P_t f^2(x)$. Consequently,

$$|\nabla P_t f|(x) \le \frac{c}{2}\Big\{P_t f^2(x) - (P_t f)^2(x)\Big\} + \frac{\Lambda(x)}{c} + \|\nabla f\|_\infty \Gamma_t(x), \quad c > 0.$$

Minimizing the upper bound with respect to $c > 0$, we therefore obtain (2.4).

(**2**) Applying (2.2) to $f = \mathrm{e}^g$ for $g \in \mathscr{B}_b(E)$ with $\|\nabla g\|_\infty < \infty$, we infer that

$$P_t g(x) \le \log(P_t \mathrm{e}^{g(y)}) + \Phi(x, y) + \Psi_t(x, y)\|\nabla g\|_\infty, \qquad x, y \in E.$$

Equivalently,

$$\exp\Big(P_t g(x) - \Phi(x, y) - \Psi_t(x, y)\|\nabla g\|_\infty\Big) \le P_t \mathrm{e}^{g(y)}, \qquad x, y \in E.$$

Integrating with respect to $\mu(\mathrm{d}y)$ on both sides and exploiting the P_t-invariance of μ, we thus derive

$$\mathrm{e}^{P_t g(x)} \int_E \exp\Big(-\Phi(x, y) - \Psi_t(x, y)\|\nabla g\|_\infty\Big)\mu(\mathrm{d}y) \le \mu(\mathrm{e}^g), \qquad x \in E.$$

Hence,

$$P_t g(x) \le \log\left(\frac{\mu(\mathrm{e}^g)}{\int_E \exp\Big(-\Phi(x, y) - \Psi_t(x, y)\|\nabla g\|_\infty\Big)\mu(\mathrm{d}y)}\right), \qquad x \in E.$$

Whence, (2.5) follows by taking $t \to \infty$.

Next, for a closed set $A \subset E$ with $\mu(A) = 0$, let

$$g_k = (1 - k\rho(\cdot, A))^+, \quad k \geq 1.$$

Then $g_k \downarrow 1_A$ as $k \uparrow \infty$. Since $g_k|_A = 1$ and $g_k \geq 0$, we have

$$m \limsup_{t\to\infty} P_t 1_A(x) \leq \limsup_{t\to\infty} P_t(m g_k(x)), \qquad x \in E, m \geq 1.$$

This, together with (2.5), leads to

$$\limsup_{t\to\infty} P_t 1_A(x) \leq \frac{1}{m} \log\left(\frac{\mu(\mathrm{e}^{m g_k})}{\int_E \mathrm{e}^{-\Phi(x,y)}\mu(\mathrm{d}y)}\right), \qquad x \in E, m \geq 1. \tag{2.13}$$

Due to $\mu(A) = 0$, one has

$$\begin{aligned}\lim_{k\to\infty} \mu(\mathrm{e}^{m g_k}) = \mu(\mathrm{e}^{m 1_A}) &= \int_A \mathrm{e}^{m 1_A(x)}\mu(\mathrm{d}x) + \int_{A^c} \mathrm{e}^{m 1_A(x)}\mu(\mathrm{d}x)\\ &= \mathrm{e}^m \mu(A) + \mu(A^c) = 1\end{aligned}$$

so that, by taking $k \to \infty$ in (2.13), we arrive at

$$\limsup_{t\to\infty} P_t 1_A(x) \leq \frac{1}{m} \log\left(\frac{1}{\int_E \mathrm{e}^{-\Phi(x,y)}\mu(\mathrm{d}y)}\right), \qquad x \in E.$$

Therefore, (2.6) holds true by approaching $m \to \infty$.

(**3**) Since the class of invariant probability measures of P_t is convex, and any two different extreme measures in the class are mutually singular (see e.g. [6, Proposition 3.2.5]), it suffices to show that any two invariant probability measures $\mu, \widetilde{\mu}$ are equivalent. For any measurable set $A \subseteq E$ with $\mu(A) = 0$, we aim to prove $\widetilde{\mu}(A) = 0$. Let $\widetilde{A} \subset A$ be a closed set. By the P_t-invariance of $\widetilde{\mu}$, (2.6) and Fatou's lemma, we obtain

$$\widetilde{\mu}(\widetilde{A}) = \limsup_{t\to\infty} \widetilde{\mu}(P_t 1_{\widetilde{A}}) \leq \widetilde{\mu}\Big(\limsup_{t\to\infty} P_t 1_A\Big) = 0.$$

So, one has

$$\widetilde{\mu}(A) = \sup_{\widetilde{A}\subset A, \widetilde{A} \text{ closed}} \widetilde{\mu}(\widetilde{A}) = 0.$$

As a consequence, we conclude that $\widetilde{\mu}$ is absolutely continuous with respect to μ. Similarly, we can infer that μ is absolutely continuous with respect to $\widetilde{\mu}$.

(**4**) Let $f(z) = \varepsilon^{-1}(\varepsilon - \rho(z, A))^+$. Then we have $f|_A = 1$, $f|_{A_\varepsilon^c} = 0$ and $\|\nabla f\|_\infty = \varepsilon^{-1}$. So, for any $n \geq 1$, (2.2) implies

$$\begin{aligned}n P_t(x, A) &\leq P_t \log(\mathrm{e}^{nf})(x) \leq \log P_t(\mathrm{e}^{nf})(y) + \Phi(x, y) + n\Psi_t(x, y)\|\nabla f\|_\infty\\ &\leq \log\big(1 + \mathrm{e}^n P_t(y, A_\varepsilon)\big) + \Phi(x, y) + n\,\varepsilon^{-1}\Psi_t(x, y).\end{aligned}$$

So we have

$$P_t(x,A) \le \frac{1}{n}\log\left(1+\mathrm{e}^n P_t(y,A_\varepsilon)\right) + \frac{1}{n}\Phi(x,y) + \varepsilon^{-1}\Psi_t(x,y). \tag{2.14}$$

If $\liminf_{t\to\infty} P_t(y,A_\varepsilon) = 0$, then, in (2.14), taking $t\to\infty$ followed by letting $n\to\infty$ and using $\Psi_t \to 0$ as $t\to\infty$ yields

$$\delta(x,A) = \liminf_{t\to\infty} P_t(x,A) \le 0,$$

which contradicts to $\delta(x,A) > 0$. Henceforth, (2.7) holds.

Next, take $t_0 > 0$ such that $P_t(x,A) \ge \varepsilon_0$ holds for all $t \ge t_0$. So, if $t \ge t_0$ such that $P_t(y,A_\varepsilon) = 0$, then, due to (2.14) by taking $n\to\infty$, we have

$$\varepsilon_0 \le P_t(x,A) \le \varepsilon^{-1}\Psi_t(x,y).$$

Hence, (2.8) holds. □

3 Non-degenerate SDEs of infinite memory

Let $(\mathbb{R}^d, \langle\cdot,\cdot\rangle, |\cdot|)$ be the d-dimensional Euclidean space. $\mathscr{C} = C((-\infty,0];\mathbb{R}^d)$ denotes the family of all continuous functions $f:(-\infty,0]\to\mathbb{R}^d$. For a fixed constant $r>0$, set

$$\mathscr{C}_r := \Big\{\xi\in\mathscr{C} : \|\xi\|_r := \sup_{-\infty<\theta\le 0}(\mathrm{e}^{r\theta}|\xi(\theta)|) < \infty\Big\}, \tag{3.1}$$

which is a Polish (i.e., complete, separable, metrizable) space with the norm $\|\cdot\|_r$. Since $r>0$ and $\theta\le 0$, the norm $\|\cdot\|_r$ means that the influence of history is exponentially weak with respect to the time parameter, which is a natural feature in the real world.

Let $\mathcal{M}_0 = \mathcal{M}_0((-\infty,0])$ be the set of all probability measures on $(-\infty,0]$. For $\kappa>0$, set

$$\mathcal{M}_\kappa := \Big\{\mu\in\mathcal{M}_0 : \mu^{(\kappa)} := \int_{-\infty}^0 \mathrm{e}^{-\kappa\theta}\mu(\mathrm{d}\theta) < \infty\Big\}.$$

$\mathbb{R}^d\otimes\mathbb{R}^d$ stands for the set of all $n\times n$-matrices with real entries, which is equipped with the Hilbert-Schmidt norm $\|\cdot\|_{HS}$. Denote $\mathscr{B}_b(\mathscr{C}_r)$ by the family of all bounded measurable functions $\phi:\mathscr{C}_r\to\mathbb{R}$ with the uniform norm $\|\phi\|_\infty := \sup_{\xi\in\mathscr{C}_r}|\phi(\xi)|$. For $A\in\mathbb{R}^d\otimes\mathbb{R}^d$, let A^{-1} be its inverse (if it exists) and $\|A\|$ its operator norm.

We consider the following SDE with infinite memory:

$$\mathrm{d}X(t) = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W(t), \quad t>0, \quad X_0 = \xi\in\mathscr{C}_r, \tag{3.2}$$

where, for each fixed $t\ge 0$, $X_t(\cdot)\in\mathscr{C}_r$ is defined by

$$X_t(\theta) := X(t+\theta), \quad \theta\in(-\infty,0],$$

which is called the segment process of $X(t)$, $b:\mathscr{C}_r\to\mathbb{R}^d$, $\sigma:\mathscr{C}_r\to\mathbb{R}^d\otimes\mathbb{R}^d$, and $(W(t))_{t\ge0}$ is a d-dimensional Brownian motion on some complete filtered probability space $(\Omega,\mathscr{F},(\mathscr{F}_t)_{t\ge0},\mathbb{P})$.

To ensure existence and uniqueness of solutions to (3.2) and to establish the asymptotic log-Harnack inequality, we impose the following assumptions on the coefficients b and σ:

(H1) $b \in C(\mathscr{C}_r)$ is bounded on bounded subsets of $\mathscr{C}_r$, and there exists $K_1 > 0$ such that

$$2\langle \xi(0) - \eta(0), b(\xi) - b(\eta) \rangle \leq K_1 \|\xi - \eta\|_r^2, \quad \xi, \eta \in \mathscr{C}_r; \tag{3.3}$$

(H2) There exists $K_2 > 0$ such that

$$\|\sigma(\xi) - \sigma(\eta)\|_{\mathrm{HS}}^2 \leq K_2 \|\xi - \eta\|_r^2, \quad \xi, \eta \in \mathscr{C}_r;$$

(H3) $\|\sigma\|_\infty := \sup_{\xi \in \mathscr{C}_r} \|\sigma(\xi)\| < \infty$, and σ is invertible with $\|\sigma^{-1}\|_\infty := \sup_{\xi \in \mathscr{C}_r} \|\sigma^{-1}(\xi)\| < \infty$.

These two assumptions guarantee that (3.2) admits a unique solution $(X^\xi(t))_{t \geq 0}$ with the initial value $X_0 = \xi \in \mathscr{C}_r$; For see Theorem A.1 below in detail. Moreover, the segment process (or functional solution) $(X_t^\xi)_{t \geq 0}$ enjoys the Markov property; see e.g. [25, Theorem 4.2]. So,

$$P_t f(\xi) = \mathbb{E} f(X_t^\xi), \quad t \geq 0,\, \xi \in \mathscr{C}_r,\, f \in \mathscr{B}_b(\mathscr{C}_r)$$

gives rise to a Markov semigroup P_t. Since the memory is infinite, P_t is not strong Feller; see, for instance, [5, 9].

Assumption (**H3**) is the usual ellipticity condition and will be used to construct couplings by change of measures for asymptotic log-Harnack inequalities.

Theorem 3.1. *Assume* (**H1**)-(**H3**). *For any $r_0 \in (0, r)$, there exists a constant $c > 0$ such that*

$$P_t \log f(\eta) \leq \log P_t f(\xi) + c\, \|\xi - \eta\|_r^2 + c\, \mathrm{e}^{-r_0 t} \|\nabla \log f\|_\infty \|\xi - \eta\|_r \tag{3.4}$$

holds for $\xi, \eta \in \mathscr{C}_r$ and $f \in \mathscr{B}_b^+(\mathscr{C}_r)$ with $\|\nabla \log f\|_\infty < \infty$. Consequently, all assertions in Theorem 2.1 hold for $E = \mathscr{C}_r$, $\rho(\xi, \eta) = \|\xi - \eta\|_r$, $\Lambda = c$, $\Gamma_t = c\, \mathrm{e}^{-r_0 t}$, and $\Phi(\xi, \eta) = c\, \|\xi - \eta\|_r^2$.

To prove (3.4), we construct coupling by change of measures (see for example [19]). Since the memory is infinite, we cannot make the coupling successful at a fixed time, but can make two marginal processes close to each other exponentially fast when $t \to \infty$. The following construction of coupling is due to [18], where SDEs without memory are concerned.

We simply denote $X_t = X_t^\xi$ and $X(t) = X^\xi(t)$, the functional solution and the solution to (3.2) with the initial value $\xi \in \mathscr{C}_r$, respectively. For any $\lambda > r$, where $r > 0$ is given in (3.1), consider the following SDE:

$$\mathrm{d}Y(t) = \{b(Y_t) + \lambda \sigma(Y_t) \sigma^{-1}(X_t)(X(t) - Y(t))\}\mathrm{d}t + \sigma(Y_t)\mathrm{d}W(t), \quad t > 0, \quad Y_0 = \eta \in \mathscr{C}_r. \tag{3.5}$$

With (**H1**)-(**H3**) in hand, we infer that (**D1**) and (**D2**) in Appendix A below hold for

$$\widetilde{b}(\zeta) := b(\zeta) + \lambda \sigma(\zeta) \sigma^{-1}(\widetilde{\zeta})(\widetilde{\zeta}(0) - \zeta(0)), \quad \zeta \in \mathscr{C}_r$$

with fixed $\widetilde{\zeta} \in \mathscr{C}_r$. Thus, under (**H1**)-(**H3**), Theorem A.1 shows that (3.5) has a unique strong solution $(Y(t))_{t \geq 0}$. Let Y_t be the segment process. To examine that Y_t has the semigroup P_t under a probability measure $\mathbb{Q}$, let

$$h(t) = \lambda\, \sigma^{-1}(X_t)(X(t) - Y(t)), \qquad \widetilde{W}(t) = W(t) + \int_0^t h(s)\mathrm{d}s,$$

and define

$$R(t) = \exp\bigg(- \int_0^t \langle h(s), \mathrm{d}W(s)\rangle - \frac{1}{2}\int_0^t |h(s)|^2 \mathrm{d}s \bigg), \quad t \geq 0. \tag{3.6}$$

We have the following result.

Lemma 3.2. *Assume* (**H1**)-(**H3**). *Then,*

$$\sup_{t\in[0,T]} \mathbb{E}\Big(R(t)\log R(t)\Big) < \infty, \quad T > 0. \tag{3.7}$$

Consequently, there exists a unique probability measure $\mathbb{Q}$ *on* $(\Omega, \mathscr{F}_\infty)$ *such that*

$$\frac{\mathrm{d}\mathbb{Q}|_{\mathscr{F}_t}}{\mathrm{d}\mathbb{P}|_{\mathscr{F}_t}} = R(t), \quad t \geq 0. \tag{3.8}$$

Moreover, $\widetilde{W}(t)$ *is a* d*-dimensional Brownian motion under* $\mathbb{Q}$.

Proof. If (3.7) holds, then $(R(t))_{t\geq 0}$ is a locally uniformly integrable martingale, and, by Girsanov's theorem, for any $T > 0$, $(\widetilde{W}(t))_{t\in[0,T]}$ is a d-dimensional Brownian motion under the probability $\mathbb{Q}_T := R(T)\mathbb{P}$. By the martingale property of $R(t)$, the family $(\mathbb{Q}_T)_{T>0}$ is harmonic, so that by Kolmogorov's harmonic theorem, there exists a unique probability measure $\mathbb{Q}$ on $(\Omega, \mathscr{F})$ such that (3.8) holds. Therefore, $(\widetilde{W}(t))_{t\geq 0}$ is a d-dimensional Brownian motion under $\mathbb{Q}$. So, it remains to prove (3.7).

For any $k > \|\xi\|_r + \|\eta\|_r$, define the stopping time

$$\tau_k = \inf\{t \geq 0 : \|X_t\|_r + \|Y_t\|_r \geq k\}.$$

Due to the non-explosion of (3.2) and (3.5) (see (A.1) below for more details), $\tau_k \uparrow \infty$ as $k \uparrow \infty$. Then, $(\widetilde{W}(t))_{t\in[0,T\wedge\tau_k]}$ is a Brownian motion under the weighted probability measure $\mathrm{d}\mathbb{Q}_{T,k} = R(T\wedge\tau_k)\mathrm{d}\mathbb{P}$. By (**H3**), there exists a constant $c_1 > 0$ such that

$$\begin{aligned}\mathbb{E}\Big(R(t\wedge\tau_k)\log R(t\wedge\tau_k)\Big) &= \mathbb{E}_{\mathbb{Q}_{T,k}} \log R(t\wedge\tau_k) = \frac{1}{2}\mathbb{E}_{\mathbb{Q}_{T,k}} \int_0^{t\wedge\tau_k} |h(s)|^2 \mathrm{d}s \\ &\leq c_1 \int_0^T \mathbb{E}_{\mathbb{Q}_{T,k}} \|X_{t\wedge\tau_k} - Y_{t\wedge\tau_k}\|_r^2 \mathrm{d}t, \qquad t \in [0,T].\end{aligned} \tag{3.9}$$

Rewrite respectively (3.2) and (3.5) as

$$\begin{aligned}&\mathrm{d}X(t) = \big\{b(X_t) - \lambda(X(t) - Y(t))\big\}\mathrm{d}t + \sigma(X_t)\mathrm{d}\widetilde{W}(t), \quad t \leq T\wedge\tau_k,\ X_0 = \xi\\ &\mathrm{d}Y(t) = b(Y_t)\mathrm{d}t + \sigma(Y_t)\mathrm{d}\widetilde{W}(t), \quad t \leq T\wedge\tau_k,\ Y_0 = \eta.\end{aligned} \tag{3.10}$$

By Itô's formula and assumptions (**H1**) and (**H2**), under the probability $\mathbb{Q}$ we have

$$\mathrm{d}|X(t) - Y(t)|^2 \leq c_2\|X_t - Y_t\|_r^2 \mathrm{d}t + 2\langle X(t) - Y(t), (\sigma(X_t) - \sigma(Y_t))\mathrm{d}\widetilde{W}(t)\rangle, \quad t \leq T\wedge\tau_k$$

for some constant $c_2 > 0$. Therefore, applying the BDG inequality and using (**H2**) once more, we can find out a constant $C(T, \xi, \eta) > 0$ such that

$$\mathbb{E}\|X_{t\wedge\tau_k} - Y_{t\wedge\tau_k}\|_r^2 \le C(T, \xi, \eta), \quad t \in [0, T].$$

Plugging this into (3.9) leads to

$$\sup_{k\ge 0, t\in[0,T]} \mathbb{E}\Big(R(t\wedge\tau_k)\log R(t\wedge\tau_k)\Big) < \infty, \quad T > 0.$$

This implies (3.7) by Fatou's lemma. □

Next, to deduce asymptotic log-Harnack inequality from the asymptotic coupling (X_t, Y_t), we show that $\|X_t - Y_t\|_r$ decays exponentially fast as $t \to \infty$ in the L^p-norm sense.

Lemma 3.3. *Assume* (**H1**)-(**H3**). *Then, for any $p > 0$ and $r_0 \in (0, r)$, there exist $\lambda, c > 0$ such that the above asymptotic coupling (X_t, Y_t) satisfies*

$$\mathbb{E}_{\mathbb{Q}}\|X_t - Y_t\|_r^p \le c\,\mathrm{e}^{-p\,r_0 t}\|\xi - \eta\|_r^p, \quad t \ge 0. \tag{3.11}$$

Proof. By Jensen's inequality, it suffices to prove for large $p > 0$, for instance, $p > 4$ as we will take below.

Let $Z(t) = X(t) - Y(t), t \in \mathbb{R}$. According to Lemma 3.2, (3.10) holds for all $t \ge 0$, where $\widetilde{W}(t)$ is a d-dimensional Brownian motion under the probability measure $\mathbb{Q}$. By applying Itô's formula and using (**H1**) and (**H2**), there exists $K > 0$ such that for all $\lambda > r$,

$$\mathrm{d}|Z(t)|^2 \le \{-2\lambda|Z(t)|^2 + K\|Z_t\|_r^2\}\mathrm{d}t + 2\langle Z(t), (\sigma(X_t) - \sigma(Y_t))\mathrm{d}\widetilde{W}(t)\rangle, \quad t \ge 0. \tag{3.12}$$

Set

$$M^{(\lambda)}(t) := 2\int_0^t \mathrm{e}^{2\lambda s}\langle Z(s), (\sigma(X_s) - \sigma(Y_s))\mathrm{d}\widetilde{W}(s)\rangle, \quad t \ge 0.$$

Thus, we deduce from (3.12) and the Itô formula that

$$\mathrm{e}^{2\lambda t}|Z(t)|^2 \le |Z(0)|^2 + K\int_0^t \mathrm{e}^{2\lambda s}\|Z_s\|_r^2\mathrm{d}s + M^{(\lambda)}(t), \quad t \ge 0. \tag{3.13}$$

So, letting $\kappa = 2(\lambda - r) > 0$, we obtain

$$\mathrm{e}^{2rt}|Z(t)|^2 \le \mathrm{e}^{-\kappa t}|Z(0)|^2 + K\int_0^t \mathrm{e}^{-\kappa(t-s)}\mathrm{e}^{2rs}\|Z_s\|_r^2\mathrm{d}s + \mathrm{e}^{-\kappa t}M^{(\lambda)}(t), \quad t \ge 0. \tag{3.14}$$

Combining this with the fact that

$$\begin{aligned}\|Z_t\|_r^2 &= \sup_{-\infty<\theta\le 0}(\mathrm{e}^{2r\theta}|Z(t+\theta)|^2) = \mathrm{e}^{-2rt}\sup_{-\infty<s\le t}(\mathrm{e}^{2rs}|Z(s)|^2)\\ &\le \mathrm{e}^{-2rt}\|Z_0\|_r^2 + \mathrm{e}^{-2rt}\sup_{0\le s\le t}(\mathrm{e}^{2rs}|Z(s)|^2),\end{aligned} \tag{3.15}$$

we arrive at

$$
\begin{aligned}
\mathrm{e}^{2rt}\|Z_t\|_r^2 &\le \|Z_0\|_r^2 + \sup_{0\le s\le t}(\mathrm{e}^{2rs}|Z(s)|^2) \\
&\le 2\|Z_0\|_r^2 + K\int_0^t \mathrm{e}^{-\kappa(t-s)}\mathrm{e}^{2rs}\|Z_s\|_r^2\mathrm{d}s + \sup_{s\in[0,t]}\Big(\mathrm{e}^{-\kappa s}M^{(\lambda)}(s)\Big),
\end{aligned}
\tag{3.16}
$$

where in the second procedure we have utilized $t\mapsto \mathrm{e}^{2rt}\|Z_t\|_r^2$ is nondecreasing. By Hölder's inequality, one finds that

$$
\begin{aligned}
\Big(\int_0^{t\wedge\tau_k}\mathrm{e}^{-\kappa(t\wedge\tau_k-s)}\mathrm{e}^{2rs}\|Z_s\|_r^2\mathrm{d}s\Big)^{p/2} &\le \Big(\int_0^\infty \mathrm{e}^{-\frac{p\kappa s}{p-2}}\mathrm{d}s\Big)^{\frac{p-2}{2}}\int_0^{t\wedge\tau_k}\mathrm{e}^{prs}\|Z_s\|_r^p\mathrm{d}s \\
&\le \Big(\frac{p-2}{p\kappa}\Big)^{\frac{p-2}{2}}\int_0^{t\wedge\tau_k}\mathrm{e}^{prs}\|Z_s\|_r^p\mathrm{d}s.
\end{aligned}
\tag{3.17}
$$

On the other hand, taking advantage of [8, Lemma 2.2], we may find out a constant $c_0(p,\lambda)>0$ with $\lim_{\lambda\to\infty}c_0(p,\lambda)=0$ such that

$$
\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}\Big(\sup_{0\le s\le t\wedge\tau_k}\Big(\mathrm{e}^{-\kappa s}M^{(\lambda)}(s)\Big)^{p/2}\Big) \\
&\le c_0(p,\lambda)\mathbb{E}_{\mathbb{Q}}\int_0^{t\wedge\tau_k}\mathrm{e}^{prs}|(\sigma(X_s)-\sigma(Y_s))^*Z(s)|^{p/2}\mathrm{d}s \\
&\le K_2^{p/4}c_0(p,\lambda)\mathbb{E}_{\mathbb{Q}}\int_0^{t\wedge\tau_k}\mathrm{e}^{prs}\|Z_s\|_r^p\mathrm{d}s \\
&\le K_2^{p/4}c_0(p,\lambda)\int_0^t\mathbb{E}_{\mathbb{Q}}(\mathrm{e}^{pr(s\wedge\tau_k)}\|Z_{s\wedge\tau_k}\|_r^p)\mathrm{d}s.
\end{aligned}
\tag{3.18}
$$

Taking (3.17) and (3.18) into consideration, we deduce from (3.16) that, for some $c(p)$, $c(p,\lambda)\in(0,\infty)$ with $c(p,\lambda)\downarrow 0$ as $\lambda\uparrow\infty$,

$$
\mathbb{E}_{\mathbb{Q}}(\mathrm{e}^{p\,r(t\wedge\tau_k)}\|Z_{t\wedge\tau_k}\|_r^p)\le c(p)\|Z_0\|_r^p + c(p,\lambda)\int_0^t\mathbb{E}_{\mathbb{Q}}(\mathrm{e}^{p\,r(s\wedge\tau_k)}\|Z_{s\wedge\tau_k}\|_r^p)\mathrm{d}s,\quad t\ge 0.
$$

By Gronwall's lemma, it follows that

$$
\mathbb{E}_{\mathbb{Q}}(\mathrm{e}^{p\,r(t\wedge\tau_k)}\|Z_{t\wedge\tau_k}\|_r^p)\le c(p)\mathrm{e}^{c(p,\lambda)t}\|Z_0\|_r^p,\quad t\ge 0.
$$

Letting $k\to\infty$, we obtain from Fatou's lemma that

$$
\mathbb{E}_{\mathbb{Q}}\|Z_t\|_r^p\le c(p)\mathrm{e}^{-(p\,r-c(p,\lambda))t}\|Z_0\|_r^p,\quad t\ge 0,
$$

which yields the desired assertion due to $c(p,\lambda)\to 0$ as $\lambda\to\infty$. □

Proof of Theorem 3.1. By Lemma 3.2 and the weak uniqueness of solutions to (3.2), we have

$$
P_tf(\eta)=\mathbb{E}_{\mathbb{Q}}f(Y_t),\quad t\ge 0,\ f\in\mathscr{B}_b(\mathscr{C}_r).
$$

So, for any $f \in \mathscr{B}_b^+(\mathscr{C}_r)$ with $\|\nabla \log f\|_\infty < \infty$, by the definition of $\|\nabla \log f\|_\infty$ and Lemma 3.3, there exists a constant $C > 0$ such that

$$\begin{aligned} P_t \log f(\eta) &= \mathbb{E}_{\mathbb{Q}} \log f(Y_t) = \mathbb{E}_{\mathbb{Q}} \log f(X_t) + \mathbb{E}_{\mathbb{Q}}(\log f(Y_t) - \log f(X_t)) \\ &\le \mathbb{E}(R(t) \log f(X_t)) + \|\nabla \log f\|_\infty \mathbb{E}_{\mathbb{Q}} \|X_t - Y_t\|_r \\ &\le \mathbb{E}(R(t) \log R(t)) + \log P_t f(\xi) + C\mathrm{e}^{-r_0 t} \|\nabla \log f\|_\infty \|\xi - \eta\|_r, \end{aligned} \tag{3.19}$$

where in the last display we have used the Young inequality; see e.g. [2, Lemma 2.4].

Next, it follows from (3.6), (3.11), (**H2**) and (**H3**) that for some constants $C_1, C_2 > 0$,

$$\begin{aligned} &\mathbb{E}(R(t) \log R(t)) = \mathbb{E}_{\mathbb{Q}} \log R(t) = \frac{\lambda^2}{2} \mathbb{E}_{\mathbb{Q}} \int_0^t |\sigma^{-1}(X_t)(X(t) - Y(t))|^2 \mathrm{d}t \\ &\le \frac{C_1 \lambda^2}{2} \int_0^t \mathbb{E}_{\mathbb{Q}} |X(s) - Y(s)|^2 \mathrm{d}s \le \frac{C_1 \lambda^2}{2} \int_0^t \mathbb{E}_{\mathbb{Q}} \|X_s - Y_s\|_r^2 \mathrm{d}s \le C_2 \lambda^2 \|\xi - \eta\|_r^2. \end{aligned}$$

Plugging this back into (3.19) yields (3.4). □

4 Neutral SDEs of infinite memory

Consider the following neutral type SDEs with infinite memory:

$$\mathrm{d}\{X(t) - G(X_t)\} = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W(t), \quad t > 0, \quad X_0 = \xi \in \mathscr{C}_r, \tag{4.1}$$

where b, σ and W are stipulated as in (3.2), and $G : \mathscr{C}_r \to \mathbb{R}^d$, which is, in general, named as the neutral term of (4.1). This kind of equation has been utilized to model some evolution phenomena arising in, e.g., physics, biology and engineering, to name a few; see e.g. [11]. Besides (**H2**) and (**H3**) above, we further assume that

(**A1**) There exists $\delta \in (0, 1)$ such that $|G(\xi) - G(\eta)| \le \delta \|\xi - \eta\|_r$ for any $\xi, \eta \in \mathscr{C}_r$;

(**A2**) $b \in C(\mathscr{C}_r)$ is bounded on bounded subsets of $\mathscr{C}_r$ and there exists an $L > 0$ such that

$$2\langle \xi(0) - \eta(0) - (G(\xi) - G(\eta)), b(\xi) - b(\eta) \rangle \le L \|\xi - \eta\|_r^2, \quad \xi, \eta \in \mathscr{C}_r.$$

Under assumptions (**A1**), (**A2**) and (**H2**), (4.1) has a unique strong solution $(X^\xi(t))_{t \ge 0}$ with the initial value $\xi \in \mathscr{C}_r$ by following exactly the argument of Theorem A.1 below. Let $(X_t^\xi)_{t \ge 0}$ be the corresponding segment process.

Theorem 4.1. *Assume* (**H2**)-(**H3**) and (**A1**)-(**A2**). *Then all assertions in Theorem 3.1 hold true.*

Proof. As in the proof of Theorem 3.1, we construct the following asymptotic coupling by change of measures. Write $(X(t), X_t) = (X^\xi(t), X_t^\xi)$ for notation brevity and consider the coupled neutral SDE with $Y_0 = \eta$:

$$\mathrm{d}\{Y(t) - G(Y_t)\} = \{b(Y_t) + \lambda\sigma(Y_t)\sigma^{-1}(X_t)(X(t) - Y(t) - (G(X_t) - G(Y_t))\}\mathrm{d}t + \sigma(Y_t)\mathrm{d}W(t).$$

For any $t \geq 0$ and $\lambda > r$, let

$$h(t) := \lambda\, \sigma^{-1}(X_t)\{X(t) - Y(t) - (G(X_t) - G(Y_t))\}, \quad \widetilde{W}(t) := W(t) + \int_0^t h(s)\mathrm{d}s.$$

Define

$$R(t) = \exp\Big(- \int_0^t \langle h(s), \mathrm{d}W(s)\rangle - \frac{1}{2}\int_0^t |h(s)|^2 \mathrm{d}s\Big), \quad t \geq 0. \tag{4.2}$$

By a close inspection of argument for Theorem 3.1, it suffices to prove (3.7) and Lemma 3.3 for the present asymptotic coupling $(X(t), Y(t))$. Below, we merely present a brief proof for the later since the former one can be done as that of Lemma 3.2. Let the probability measure $\mathbb{Q}$ be given by (3.8) with $R(t)$ defined in (4.2). Then $\widetilde{W}(t)$ is a d-dimensional Brownian motion under $\mathbb{Q}$. Again let $Z(t) = X(t) - Y(t), t \in \mathbb{R}$, and Z_t be the associated segment process. By following the argument to derive (3.14), (**H2**) and (**A2**) imply

$$\begin{aligned} \mathrm{e}^{2rt}|Z(t) - (G(X_t) - G(Y_t))|^2 &\leq \mathrm{e}^{-\kappa t}|Z(0) - (G(X_0) - G(Y_0))|^2 \\ &\quad + K\int_0^t \mathrm{e}^{-\kappa(t-s)}\mathrm{e}^{2rs}\|Z_s\|_r^2 \mathrm{d}s + \mathrm{e}^{-\kappa t}M^{(\lambda)}(t) \end{aligned} \tag{4.3}$$

for $\kappa := 2(\lambda - r) > 0$, some constant $K > 0$ and

$$M^{(\lambda)}(t) := 2\int_0^t \mathrm{e}^{2\lambda s}\langle Z(s) - (G(X_s) - G(Y_s)), (\sigma(X_s) - \sigma(Y_s))\mathrm{d}\widetilde{W}(s)\rangle.$$

Next, for any $\varepsilon > 0$, it follows from (**A1**) that

$$\begin{aligned} |Z(t)|^2 &\leq (1+\varepsilon)|Z(t) - (G(X_t) - G(Y_t))|^2 + (1 + 1/\varepsilon)|G(X_t) - G(Y_t)|^2 \\ &\leq (1+\varepsilon)|Z(t) - (G(X_t) - G(Y_t))|^2 + (1 + 1/\varepsilon)\delta^2\|Z_t\|_r^2. \end{aligned}$$

This, together with (3.16), yields

$$\mathrm{e}^{2rt}\|Z_t\|_r^2 \leq \|Z_0\|_r^2 + (1+\varepsilon)\sup_{0\leq s\leq t}(\mathrm{e}^{2rs}|Z(s) - (G(X_s) - G(Y_s))|^2) + (1+1/\varepsilon)\delta^2\mathrm{e}^{2rt}\|Z_t\|_r^2.$$

Taking $\varepsilon = \frac{\delta}{1-\delta}$, we derive

$$\mathrm{e}^{2rt}\|Z_t\|_r^2 \leq \frac{1}{1-\delta}\Big\{\|Z_0\|_r^2 + \frac{1}{1-\delta}\sup_{0\leq s\leq t}(\mathrm{e}^{2rs}|Z(s) - (G(X_s) - G(Y_s))|^2)\Big\}.$$

Combining this with (4.3), and noting that (**A1**) implies

$$|\xi(0) - \eta(0) - (G(\xi) - G(\eta))|^2 \leq 4\|\xi - \eta\|_r^2, \quad \xi, \eta \in \mathscr{C}_r,$$

we arrive at

$$\begin{aligned} \mathrm{e}^{prt}\|Z_t\|_r^p \leq c\Big\{&\|Z_0\|_r^p + \Big(\int_0^{t\wedge\tau_k} \mathrm{e}^{-\kappa(t\wedge\tau_k - s)}\mathrm{e}^{2rs}\|Z_s\|_r^2\mathrm{d}s\Big)^{p/2} \\ &+ \sup_{0\leq s\leq t\wedge\tau_k}\Big(\mathrm{e}^{-\kappa s}M^{(\lambda)}(s)\Big)^{p/2}\Big\} \end{aligned} \tag{4.4}$$

for some constant $c > 0$. With the aid of (**A1**) and (**H3**), we observe that (3.17) and (3.18) still hold for $p > 4$. Combining this with $\kappa \to \infty$ as $\lambda \to \infty$, we deduce (3.11) from (4.4). □

Remark 4.1. Indeed, the existence and uniqueness of solutions to (4.1) under the locally weak monotone condition and the weak coercive condition can be obtained by following the argument of Theorem A and constructing the following Euler-Maruyama scheme

$$\mathrm{d}\{X^n(t) - G(X_t^n)\} = b(\widehat{X}_t^n)\mathrm{d}t + \sigma(\widehat{X}_t^n)\mathrm{d}W(t), \quad t > 0, \quad X_0^n = X_0 = \xi,$$

where, for $t \geq 0$, $\widehat{X}_t^n(\theta) := X^n((t+\theta) \wedge t_n),\ t_n := [nt]/n,\ \theta \in (-\infty, 0]$.

5 Semi-linear SPDEs of infinite memory

Let $(\mathbb{H}, \langle \cdot, \cdot \rangle, |\cdot|)$ be a real separable Hilbert space. $\mathscr{C} = C((-\infty, 0]; \mathbb{H})$ denotes the family of all continuous mappings $f : (-\infty, 0] \to \mathbb{H}$, and $\mathscr{C}_r$ is defined as in (3.1). Let $\mathscr{L}(\mathbb{H})$ and $\mathscr{L}_{HS}(\mathbb{H})$ be the spaces of all bounded linear operators and Hilbert-Schmidt operators on $\mathbb{H}$, respectively. Denote $\|\cdot\|$ and $\|\cdot\|_{HS}$ by the operator norm and the Hilbert-Schmidt norm, respectively.

Consider the following semi-linear SPDE on $\mathbb{H}$ with infinite memory:

$$\mathrm{d}X(t) = \{AX(t) + b(X_t)\}\mathrm{d}t + \sigma(X_t)\mathrm{d}W(t), \quad t > 0, \quad X_0 = \xi, \tag{5.1}$$

where $(A, \mathscr{D}(A))$ is a densely defined closed operator on $\mathbb{H}$ generating a C_0-semigroup e^{tA}, $b : \mathscr{C}_r \to \mathbb{H}$, $\sigma : \mathscr{C}_r \to \mathscr{L}(\mathbb{H})$, and $(W(t))_{t\geq 0}$ is a cylindrical Wiener process on $\mathbb{H}$ for a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with the natural filtration $(\mathscr{F}_t)_{t\geq 0}$.

We assume that

(B1) $(-A, \mathscr{D}(A))$ is self-adjoint with discrete spectrum $0 < \lambda_1 \leq \lambda_2 \leq \cdots$ counting multiplicities such that $\sum_{i\geq 1} \lambda_i^{-\alpha} < \infty$ for some $\alpha \in (0, 1)$;

(B2) There exists an $L_0 > 0$ such that

$$|b(\xi) - b(\eta)| + \|\sigma(\xi) - \sigma(\eta)\|_{HS} \leq L_0 \|\xi - \eta\|_r, \quad \xi, \eta \in \mathscr{C}_r;$$

(B3) $\|\sigma\|_\infty := \sup_{\xi\in\mathscr{C}_r} \|\sigma(\xi)\| < \infty$, and $\sigma(\xi)$ is invertible with $\|\sigma^{-1}\|_\infty := \sup_{\xi\in\mathscr{C}_r} \|\sigma^{-1}(\xi)\| < \infty$.

According to **(B2)** and **(B3)**, σ need not, but the difference $\sigma(\xi) - \sigma(\eta)$ does, take values in the space of Hilbert-Schmidt operators. Recall that a continuous adapted process $(X_t^\xi)_{t\geq 0}$ on $\mathscr{C}_r$ is called a mild solution to (5.1) with the initial value $\xi \in \mathscr{C}_r$, if $X_0^\xi = \xi$ and

$$X^\xi(t) = \mathrm{e}^{tA}\xi(0) + \int_0^t \mathrm{e}^{(t-s)A} b(X_s^\xi)\mathrm{d}s + \int_0^t \mathrm{e}^{(t-s)A}\sigma(X_s^\xi)\mathrm{d}W(s), \quad t \geq 0.$$

In terms of the following result, assumptions **(B1)-(B3)** imply the existence and uniqueness of mild solutions to (5.1) as well as asymptotic log-Harnack inequality of the associated Markov semigroup.

Theorem 5.1. *Assume* **(B1)-(B3)**. *Then* (5.1) *has a unique mild solution* $(X_t^\xi)_{t\geq 0}$, *and the associated Markov semigroup* P_t *satisfies all assertions in Theorem 3.1.*

Proof. **(a)** The existence and uniqueness of mild solutions follows from the Banach fixed point theorem by a more or less standard argument under the assumptions **(B1)-(B3)**. Fix $T > 0$ and let

$$\mathscr{D}_T = \Big\{(u(t))_{t\in(-\infty,T]} \text{ is a continuous adapted process on } \mathbb{H} \text{ with } u_0 = \xi \\ \text{and } \mathbb{E}\Big(\sup_{t\in(-\infty,T]}(\mathrm{e}^{rt}|u(t)|^4)\Big) < \infty\Big\}.$$

Then $\mathscr{D}_T$ is a complete metric space with

$$\rho(u,v) := \|u-v\|_{\mathscr{D}_T} := \Big(\mathbb{E}\Big(\sup_{t\in[0,T]}(\mathrm{e}^{rt}|u(t)-v(t)|^4)\Big)\Big)^{\frac{1}{4}}.$$

Observe that the metric ρ is equivalent to the metric below

$$\rho_0(u,v) := \|u-v\|_{\mathscr{D}_T^0} := \Big(\mathbb{E}\Big(\sup_{t\in[0,T]}|u(t)-v(t)|^4\Big)\Big)^{\frac{1}{4}}.$$

By **(B1)-(B3)**, it is easy to see that

$$\Gamma(u)(t) := \mathrm{e}^{tA}\xi(0) + \int_0^t \mathrm{e}^{(t-s)A}b(u_s)\mathrm{d}s + \int_0^t \mathrm{e}^{(t-s)A}\sigma(u_s)\mathrm{d}W(s), \quad t\geq 0, \quad u\in\mathscr{D}_T \tag{5.2}$$

gives rise to a map from $\mathscr{D}_T$ to $\mathscr{D}_T$. Then, by virtue of the fixed point theorem, it remains to find a constant $T_0 > 0$ independent of ξ such that, for any $T \leq T_0$, the map Γ is contractive in $\mathscr{D}_T$ sine the existence and uniqueness of mild solution on the intervals $[T_0, 2T_0], [2T_0, 3T_0], \cdots$ can be done inductively. Below we provide a brief proof for this.

For any $u, v\in\mathscr{D}_T$, by (5.2) we have

$$\mathrm{d}\{\Gamma(u)(t)-\Gamma(v)(t)\} = \big\{A\big(\Gamma(u)(t)-\Gamma(v)(t)\big) + b(u_t) - b(v_t)\big\}\mathrm{d}t + \{\sigma(u_t)-\sigma(v_t)\}\mathrm{d}W(t).$$

According to **(B1)-(B3)**, we may apply Itô's formula to $|\Gamma(u)(t)-\Gamma(v)(t)|^2$ to derive that there exists $c_1 > 0$ such that

$$\begin{aligned}\mathrm{d}|\Gamma(u)(t)-\Gamma(v)(t)|^2 &= 2\big\langle\Gamma(u)(t)-\Gamma(v)(t), A\big(\Gamma(u)(t)-\Gamma(v)(t)\big) + b(u_t)-b(v_t)\big\rangle\mathrm{d}t \\ &\quad + \|\sigma(u_t)-\sigma(v_t)\|_{HS}^2\mathrm{d}t + \mathrm{d}M(t) \\ &\leq \frac{1}{2\sqrt{3}T}|\Gamma(u)(t)-\Gamma(v)(t)|^2\mathrm{d}t + c_1(1+T)\|u_t-v_t\|_r^2\mathrm{d}t + \mathrm{d}M(t),\end{aligned}$$

where we have used the negative definite property of A due to **(B1)** in the last step and set

$$M(t) := 2\int_0^t \big\langle\Gamma(u)(s)-\Gamma(v)(s), \{\sigma(u_s)-\sigma(v_s)\}\mathrm{d}W(s)\big\rangle$$

is a martingale. By the BDG inequality, there exists a constant $c_2 > 0$ such that

$$\begin{aligned}\|\Gamma(u)-\Gamma(v)\|^4_{\mathscr{D}^0_T} &= \mathbb{E}\Big(\sup_{t\in[0,T]}|\Gamma(u)(t)-\Gamma(v)(t)|^4\Big)\\ &\le \frac{1}{4}\|\Gamma(u)-\Gamma(v)\|^4_{\mathscr{D}^0_T} + 3c_1^2\,(1+T)^2T^2\|u-v\|^4_{\mathscr{D}_T} + 3\,\mathbb{E}\Big(\sup_{t\in[0,T]} M(t)^2\Big)\\ &\le \frac{1}{4}\|\Gamma(u)-\Gamma(v)\|^4_{\mathscr{D}^0_T} + 3c_1^2(1+T)^2T^2\|u-v\|^4_{\mathscr{D}_T} + c_2\,\mathbb{E}\langle M(T)\rangle.\end{aligned} \tag{5.3}$$

Note that the definition of $M(t)$ and the assumption **(B2)** imply

$$\begin{aligned}\mathbb{E}\langle M(T)\rangle &\le 4L_0\int_0^T \mathbb{E}(\Gamma(u)(t)-\Gamma(v)(t)|^2\|u_t-v_t\|_r^2)\mathrm{d}t\\ &\le 4L_0T\|\Gamma(u)-\Gamma(v)\|^2_{\mathscr{D}^0_T}\|u-v\|^2_{\mathscr{D}_T}\\ &\le \frac{1}{4c_2}\|\Gamma(u)-\Gamma(v)\|^4_{\mathscr{D}^0_T} + 16c_2L_0^2T^2\|u-v\|^4_{\mathscr{D}_T}.\end{aligned}$$

Putting this into (5.3) gives that

$$\|\Gamma(u)-\Gamma(v)\|^4_{\mathscr{D}^0_T} \le 2(3c_1^2(1+T^2)+16c_2^2L_0^2)T^2\|u-v\|^4_{\mathscr{D}_T}$$

so that

$$\|\Gamma(u)-\Gamma(v)\|^4_{\mathscr{D}_T} \le 2(3c_1^2(1+T^2)+16c_2^2L_0^2)T^2\mathrm{e}^{4rT}\|u-v\|^4_{\mathscr{D}_T}.$$

Therefore, by taking $T_0>0$ such that $2(3c_1^2(1+T_0^2)+16c_2^2L_0^2)T_0^2\mathrm{e}^{4rT_0}<1$, we conclude that Γ is contractive in $\mathscr{D}_T$ for any $T\le T_0$.

(b) It remains to verify the asymptotic log-Harnack inequality (3.4). To this end, we construct an asymptotic coupling by change of measures as follows. Let $(X(t),X_t) = (X^\xi(t),X_t^\xi)$, and for any $\lambda>0$, consider the following SPDE with $Y_0=\eta$:

$$\mathrm{d}Y(t) = \{AY(t)+b(Y_t)+\lambda\sigma(Y_t)\sigma^{-1}(X_t)(X(t)-Y(t))\}\mathrm{d}t+\sigma(Y_t)\mathrm{d}W(t),\quad t>0. \tag{5.4}$$

As shown in **(a)**, assumptions (**B1**)-(**B3**) imply that (5.4) has a unique local mild solution $(Y(t))_{t\ge0}$. Moreover, since the drift $\widetilde{b}(\zeta) := b(\zeta)+\lambda\sigma(\zeta)\sigma^{-1}(\widetilde{\zeta})(\widetilde{\zeta}(0)-\zeta(0))$ for any $\zeta\in\mathscr{C}_r$ and fixed $\widetilde{\zeta}\in\mathscr{C}_r$ is of linear growth due to (**B2**) and (**B3**), we indeed deduce that the unique local mild solution is the global one. Let $(Y_t)_{t\ge0}$ be the associated segment process. For any $t\ge0$ and $\lambda>r$, set

$$h(t) := \lambda\,\sigma^{-1}(X_t)(X(t)-Y(t)),\quad \widetilde{W}(t) := W(t)+\int_0^t h(s)\mathrm{d}s.$$

Define

$$R(t) = \exp\Big(-\int_0^t\langle h(s),\mathrm{d}W(s)\rangle - \frac{1}{2}\int_0^t|h(s)|^2\mathrm{d}s\Big).$$

As explained in the proof of Theorem 4.1, we only need to verify (3.7) and (3.11) for the present framework. By a standard finite-dimensional approximation argument (see for instance [19, Theorem 4.1.3]), these can be easily deduced from assumptions (**B1**)-(**B3**). We therefore skip the details to save space. □

6 Stochastic Hamiltonian systems of infinite memory

In this section, we establish the asymptotic log-Harnack inequality for a class of degenerate SDEs of infinite memory. More precisely, we consider the following stochastic Hamiltonian system of infinite memory on $\mathbb{R}^{2d} := \mathbb{R}^d \times \mathbb{R}^d$

$$\begin{cases} \mathrm{d}X(t) = \lambda Y(t)\mathrm{d}t \\ \mathrm{d}Y(t) = b(X_t, Y_t)\mathrm{d}t + \sigma(X_t, Y_t)\mathrm{d}W(t) \end{cases} \tag{6.1}$$

with the initial value $(X_0, Y_0) = (\xi, \eta) \in \mathscr{C}_r \times \mathscr{C}_r$, where $\lambda > 0$, $b : \mathscr{C}_r \times \mathscr{C}_r \to \mathbb{R}^d$, $\sigma : \mathscr{C}_r \times \mathscr{C}_r \to \mathbb{R}^d \otimes \mathbb{R}^d$, and $(W(t))_{t\geq 0}$ is a d-dimensional Brownian motion defined on the probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$. When the memory is finite or empty, this model has been intensively investigated, see for instance [4, 12, 21, 22, 23, 27] for results on derivative formulas, Harnack inequalities, hypercontractivity, ergodicity, well-posedness, and so forth.

To investigate the present setup with infinite memory, we make the following assumptions.

(**C1**) There exist constants $\beta, L_1 > 0$ such that for any $(\xi, \eta), (\overline{\xi}, \overline{\eta}) \in \mathscr{C}_r \times \mathscr{C}_r$,

$$\langle \beta(\xi(0) - \overline{\xi}(0)) + (\eta(0) - \overline{\eta}(0)), b(\xi, \eta) - b(\overline{\xi}, \overline{\eta}) \rangle \leq L_1(\|\xi - \overline{\xi}\|_r^2 + \|\eta - \overline{\eta}\|_r^2);$$

(**C2**) There exists an $L_2 > 0$ such that for any $(\xi, \eta), (\overline{\xi}, \overline{\eta}) \in \mathscr{C}_r \times \mathscr{C}_r$,

$$\|\sigma(\xi, \eta) - \sigma(\overline{\xi}, \overline{\eta})\|_{HS}^2 \leq L_2(\|\xi - \overline{\xi}\|_r^2 + \|\eta - \overline{\eta}\|_r^2);$$

(**C3**) $\|\sigma\|_\infty := \sup_{\xi,\eta\in\mathscr{C}_r} \|\sigma(\xi, \eta)\| < \infty$, and σ is invertible with $\sup_{\xi,\eta\in\mathscr{C}_r} \|\sigma^{-1}(\xi, \eta)\| < \infty$.

Under assumptions (**C1**) and (**C2**), (6.1) admits from Theorem A.1 in the Appendix A a unique strong solution $(X^\xi(t), Y^\eta(t))_{t\geq 0}$ with the corresponding segment process $(X_t^\xi, Y_t^\eta)_{t\geq 0}$, which is a homogeneous Markov process. Let P_t be the semigroup generated by (X_t^ξ, Y_t^η), i.e., $P_t f(\xi, \eta) = \mathbb{E}f(X_t^\xi, Y_t^\eta)$, $f \in \mathscr{B}_b(\mathscr{C}_r \times \mathscr{C}_r)$.

For $p > 2$ and $\frac{1}{p} < \alpha < \frac{1}{2}$, let

$$\Lambda_{p,\alpha} = \left(\frac{p^{1+p}}{2(p-1)^{p-1}}\right)^{p/2} \left(\frac{\Gamma(1-2\alpha)}{2^{1-2\alpha}}\right)^{p/2} \left(1 - \frac{1}{p}\right)^{p\alpha-1} \Gamma\Big(\frac{p\alpha - 1}{p-1}\Big)^{p-1},$$

where $\Gamma(\cdot)$ is the Gamma function, it is easy to show that the function $(p, \alpha) \mapsto \Lambda_{p,\alpha}$ achieves its infimum at some point (p_0, α_0), i.e.

$$\Lambda_{p_0,\alpha_0} = \inf_{p>2, \frac{1}{p}<\alpha<\frac{1}{2}} \Lambda_{p,\alpha}, \quad p_0 > 2, \quad \frac{1}{p_0} < \alpha_0 < \frac{1}{2}.$$

Moreover, we denote

$$\mu_{p_0} = 2^{3p_0-1}\Big\{(L_1 + L_2/2)^{p_0}(1 - 1/p_0)^{p_0-1} + \Lambda_{p_0,\alpha_0} L_2^{p_0/2}\Big\}. \tag{6.2}$$

Theorem 6.1. Assume **(C1)-(C3)**. If

$$\lambda > r + \frac{1+\beta+2\beta^2}{2\beta}\Big(\frac{\mu_{p_0}}{2p_0 r}\Big)^{\frac{2}{p_0-2}}, \tag{6.3}$$

then there exist $r_0 \in (0, r)$ and a constant $c > 0$ such that

$$\begin{aligned} P_t \log f(\xi,\eta) &\le \log P_t f(\xi',\eta') + c\,(\|\xi-\xi'\|_r^2 + \|\eta-\eta'\|_r^2) \\ &\quad + c\,\mathrm{e}^{-r_0 t}\|\nabla \log f\|_\infty (\|\xi-\xi'\|_r + \|\eta-\eta'\|_r) \end{aligned} \tag{6.4}$$

holds for $(\xi,\eta),(\xi',\eta') \in \mathscr{C}_r \times \mathscr{C}_r$ and $f \in \mathscr{B}_b^+(\mathscr{C}_r\times\mathscr{C}_r)$ with $\|\nabla \log f\|_\infty < \infty$. Consequently, all assertions in Theorem 2.1 hold true.

Proof. Again, we adopt the asymptotic coupling by change of measures. Let $(X(t), Y(t))$ solve (6.1) for $(X_0, Y_0) = (\xi, \eta)$. For $\lambda > 0$ in (6.1) and $\beta > 0$ in (**C1**), consider the following stochastic Hamiltonian system

$$\begin{cases} \mathrm{d}\overline{X}(t) = \lambda\,\overline{Y}(t)\mathrm{d}t \\ \mathrm{d}\overline{Y}(t) = \Big\{b(\overline{X}_t,\overline{Y}_t) + \sigma(\overline{X}_t,\overline{Y}_t)\sigma^{-1}(X_t,Y_t)\Big(\lambda(X(t)-\overline{X}(t)) \\ \qquad\qquad + 2\lambda\beta(Y(t)-\overline{Y}(t))\Big)\Big\}\mathrm{d}t + \sigma(\overline{X}_t,\overline{Y}_t)\mathrm{d}W(t) \end{cases} \tag{6.5}$$

with the initial value $(\overline{X}_0,\overline{Y}_0) = (\overline{\xi},\overline{\eta}) \in \mathscr{C}_r\times\mathscr{C}_r$. Under (**C1**)-(**C3**), according to Theorem A, (6.5) has a unique strong solution $(\overline{X}(t),\overline{Y}(t))_{t\ge 0}$ with the associated segment process $(\overline{X}_t,\overline{Y}_t)_{t\ge0}$. For any $t \ge 0$, let

$$h(t) = \sigma^{-1}(X_t,Y_t)\Big(\lambda(X(t)-\overline{X}(t)) + 2\lambda\,\beta\,(Y(t)-\overline{Y}(t))\Big),\quad \widetilde{W}(t) = W(t) + \int_0^t h(s)\mathrm{d}s.$$

Define

$$R(t) = \exp\bigg(-\int_0^t \langle h(s), \mathrm{d}W(s)\rangle - \frac{1}{2}\int_0^t |h(s)|^2\mathrm{d}s\bigg).$$

As shown in the proof of Theorem 3.1, it suffices to prove (3.7) and Lemma 3.3 for the present coupling $((X_t,Y_t),(\overline{X}_t,\overline{Y}_t))$. For simplicity, we only prove the latter one. It is easy to see that for any $x, y \in \mathbb{R}^d$,

$$\frac{1}{4}(|x|^2+|y|^2) \le V(x,y) := (1/2+\beta^2)|x|^2 + |y|^2/2 + \beta\langle x,y\rangle \le c_\beta(|x|^2+|y|^2), \tag{6.6}$$

where $c_\beta := (1+\beta+2\beta^2)/2$. Set $Z(t) := (X(t)-\overline{X}(t), Y(t)-\overline{Y}(t)),\ t\in\mathbb{R}$. Since (6.1) and (6.5) reduce to

$$\begin{cases} \mathrm{d}X(t) = \lambda\,Y(t)\mathrm{d}t \\ \mathrm{d}Y(t) = \{b(X_t,Y_t) - \lambda(X(t)-\overline{X}(t)) - 2\lambda\beta(Y(t)-\overline{Y}(t))\}\mathrm{d}t + \sigma(X_t,Y_t)\mathrm{d}\widetilde{W}(t) \end{cases}$$

and

$$\begin{cases} \mathrm{d}\overline{X}(t) = \lambda \overline{Y}(t)\mathrm{d}t \\ \mathrm{d}\overline{Y}(t) = b(\overline{X}_t, \overline{Y}_t)\mathrm{d}t + \sigma(\overline{X}_t, \overline{Y}_t)\mathrm{d}\widetilde{W}(t), \end{cases}$$

by Itô's formula we obtain

$$\begin{aligned} \mathrm{d}V(Z(t)) = \Big\{ &\big\langle (1+2\beta^2)(X(t)-\overline{X}(t)) + \beta(Y(t)-\overline{Y}(t)), \lambda(Y(t)-\overline{Y}(t))\big\rangle \\ &+ \big\langle \beta(X(t)-\overline{X}(t)) + Y(t)-\overline{Y}(t), -\lambda(X(t)-\overline{X}(t)) - 2\lambda\beta(Y(t)-\overline{Y}(t))\big\rangle \\ &+ \big\langle \beta(X(t)-\overline{X}(t)) + Y(t)-\overline{Y}(t), b(X_t,Y_t) - b(\overline{X}_t,\overline{Y}_t)\big\rangle \\ &+ \frac{1}{2}\|\sigma(X_t,Y_t)-\sigma(\overline{X}_t,\overline{Y}_t)\|_{HS}^2\Big\}\mathrm{d}t + \mathrm{d}M(t) \\ =: {}& I(t)\mathrm{d}t + \mathrm{d}M(t), \quad t \geq 0, \end{aligned} \tag{6.7}$$

where

$$\mathrm{d}M(t) := \big\langle Y(t)-\overline{Y}(t) + \beta(X(t)-\overline{X}(t)), (\sigma(X_t,Y_t)-\sigma(\overline{X}_t,\overline{Y}_t))\mathrm{d}\widetilde{W}(t)\big\rangle,$$

and by (**C1**) and (**C2**),

$$I(t) \leq -\lambda\beta|Z(t)|^2 + (L_1 + L_2/2)\|Z_t\|_r^2, \quad t \geq 0.$$

Whence, it follows from (6.7) that

$$\mathrm{d}V(Z(t)) \leq \big\{ -\lambda\beta|Z(t)|^2 + (L_1+L_2/2)\|Z_t\|_r^2\big\}\mathrm{d}t + \mathrm{d}M(t), \quad t \geq 0. \tag{6.8}$$

Letting $\lambda' = \frac{\lambda\beta}{2c_\beta}$ such that $2c_\beta\lambda' - \lambda\beta = 0$, and combining this with (6.6), we obtain

$$\begin{aligned} \mathrm{d}(\mathrm{e}^{2\lambda' t}V(Z(t))) &= \mathrm{e}^{2\lambda' t}\{2\lambda' V(Z(t))\mathrm{d}t + \mathrm{d}V(Z(t))\} \\ &\leq \mathrm{e}^{2\lambda' t}\{(2c_\beta\lambda' - \lambda\beta)|Z(t)|^2 + (L_1+L_2/2)\|Z_t\|_r^2\}\mathrm{d}t + \mathrm{e}^{2\lambda' t}\mathrm{d}M(t) \\ &= (L_1+L_2/2)\mathrm{e}^{2\lambda' t}\|Z_t\|_r^2\mathrm{d}t + \mathrm{e}^{2\lambda' t}\mathrm{d}M(t), \quad t \geq 0. \end{aligned}$$

Setting $\kappa = 2(\lambda' - r)$ and using (6.6) again, we derive that

$$\begin{aligned} \mathrm{e}^{2rt}|Z(t)|^2 \leq{}& 4\mathrm{e}^{-\kappa t}V(Z(0)) + 4(L_1+L_2/2)\int_0^t \mathrm{e}^{-\kappa(t-s)}\mathrm{e}^{2rs}\|Z_s\|_r^2\mathrm{d}s \\ &+ 4\mathrm{e}^{-\kappa t}\int_0^t \mathrm{e}^{2\lambda' s}\mathrm{d}M(s). \end{aligned} \tag{6.9}$$

For any $k > \|\xi\|_r + \|\eta\|_r + \|\overline{\xi}\|_r + \|\overline{\eta}\|_r$, define the stopping time

$$\tau_k = \inf\{t \geq 0 : \|X_t\|_r + \|Y_t\|_r + \|\overline{X}_t\|_r + \|\overline{Y}_t\|_r \geq t\}.$$

By the Hölder inequality, one has

$$\Big(\int_0^{t\wedge\tau_k} \mathrm{e}^{-\kappa(t\wedge\tau_k - s)}\mathrm{e}^{2rs}\|Z_s\|_r^2\mathrm{d}s\Big)^{p_0} \leq \frac{(1-1/p_0)^{p_0-1}}{\kappa^{p_0-1}}\int_0^{t\wedge\tau_k} \mathrm{e}^{2p_0rs}\|Z_s\|_r^{2p_0}\mathrm{d}s. \tag{6.10}$$

Moreover, employing [8, Lemma 2.2] leads to

$$
\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}\bigg(\sup_{0\le s\le t\wedge\tau_k}\Big(\mathrm{e}^{-\kappa s}\int_0^s \mathrm{e}^{2\lambda' u}\mathrm{d}M(u)\Big)^{p_0}\bigg)\\
&\le \frac{\Lambda_{p_0,\alpha_0}L_0^{p_0/2}}{\kappa^{p_0/2-1}}\int_0^t \mathbb{E}_{\mathbb{Q}}(\mathrm{e}^{2p_0 r(s\wedge\tau_k)}\|Z_{s\wedge\tau_k}\|_r^{2p_0})\mathrm{d}s,
\end{aligned}
\tag{6.11}
$$

where the explicit expression of Λ_{p_0,α_0} was provided in the last line of the argument of [8, Lemma 2.2]. Thus, taking (6.9), (6.10), and (6.11) into account and employing Fatou's lemma yields

$$
\mathbb{E}(\mathrm{e}^{2p_0 rt}\|Z_t\|_r^{2p_0}) \le c_{p_0,\varepsilon}\|Z_0\|_r^{2p_0} + (1+\varepsilon)\frac{\mu_{p_0}}{\kappa^{p_0/2-1}}\int_0^t \mathbb{E}_{\mathbb{Q}}(\mathrm{e}^{2p_0 rs}\|Z_s\|_r^{2p_0})\mathrm{d}s, \quad \varepsilon>0,
\tag{6.12}
$$

for some constant $c_{p_0,\varepsilon}>0$, where μ_{p_0} was introduced in (6.2). Consequently, the desired assertion follows by taking $\varepsilon>0$ sufficiently small, applying Gronwall's inequality and utilizing (6.3).

□

A Appendix

To make the content self-contained, in this section, we address existence and uniqueness of solutions to (3.2) under the locally weak monotonicity and the weak coercivity. Assume that

(**D1**) $b\in C(\mathscr{C}_r)$ and $\sigma\in C(\mathscr{C}_r)$ are bounded on bounded subsets of $\mathscr{C}_r$, and, for each $k\ge 1$, there is an $L_k>0$ such that for all $\xi,\eta\in\mathscr{C}_r$ with $\|\xi\|_r\vee\|\eta\|_r\le k$,

$$
2\langle\xi(0)-\eta(0), b(\xi)-b(\eta)\rangle + \|\sigma(\xi)-\sigma(\eta)\|_{\mathrm{HS}}^2 \le L_k\|\xi-\eta\|_r^2.
$$

(**D2**) There exists an $L>0$ such that $2\langle\xi(0),b(\xi)\rangle^+ + \|\sigma(\xi)\|_{\mathrm{HS}}^2 \le L(1+\|\xi\|_r^2)$, $\xi\in\mathscr{C}_r$.

Theorem A.1. Let (**D1**) and (**D2**) hold. Then, (3.2) has a unique solution $(X(t))_{t\ge0}$ such that for some $C>0$,

$$
\mathbb{E}\|X_t\|_r^2 \le C\mathrm{e}^{Ct}(1+\|\xi\|_r^2), \quad t\ge0, \quad \xi\in\mathscr{C}_r.
\tag{A.1}
$$

Proof. Below we follow the idea of [14, Theorem 2.3]. Set $N_0:=\{n\in\mathbb{N}: n\ge\frac{r}{\log 2}\}$ and $[s]:=\sup\{k\in\mathbb{Z}: k\le s\}$, the integer par of $s>0$. For any $n\in N_0$, consider an SDE

$$
\mathrm{d}X^n(t) = b(\widehat{X}_t^n)\mathrm{d}t + \sigma(\widehat{X}_t^n)\mathrm{d}W(t), \quad t>0, \quad X_0^n = X_0 = \xi,
\tag{A.2}
$$

where, $\widehat{X}_t^n(\theta) := X^n((t+\theta)\wedge t_n), \theta\in(-\infty,0]$ and $t_n:=[nt]/n$. Define the stopping time

$$
\tau_R^n = \inf\Big\{t\ge0:\ |X^n(t)|\ge R\Big\} = \inf\Big\{t\ge0:\ \|X_t^n\|_r\ge R\Big\},\ R>\|\xi\|_r,\ n\in N_0.
\tag{A.3}
$$

Thanks to $n \in N_0$, we have $\mathrm{e}^{r/n} \le 2$ so that

$$\|\widehat{X}_t^n\|_r \le \|X_t^n\|_r \vee |X^n(t_n)| \le \mathrm{e}^{r(t-t_n)}\|X_t^n\|_r \le 2\|X_t^n\|_r. \tag{A.4}$$

Since b is bounded on bounded subsets of $\mathscr{C}_r$, we get

$$|b(X_t^n)| \le C(R) := \sup_{\|\zeta\|_r \le R} |b(\zeta)| < \infty, \quad R \in (\|\xi\|_r, \infty),\ t \in [0, \tau_R^n]. \tag{A.5}$$

Let $Z^{n,m}(t) = X^n(t) - X^m(t)$ and $p_t^n = X_t^n - \widehat{X}_t^n$. By the notion of τ_R^n, (A.4) implies that

$$\|p_t^n\|_r \le 3R, \quad t \le \tau_R^n. \tag{A.6}$$

By Itô's formula and using (**D1**), (A.4) and (A.5), there are $C, K > 0$ such that

$$\mathrm{d}(\mathrm{e}^{2rt}|Z^{n,m}(t)|^2) \le K\Big\{\sup_{0\le s\le t}(\mathrm{e}^{2rs}|Z^{n,m}(s)|^2) + \mathrm{e}^{2rt}(\|p_t^n\|_r + \|p_t^m\|_r)\Big\}\mathrm{d}t + \mathrm{d}M^{n,m}(t)$$

for any $t \in [0, \tau_R^n \wedge \tau_R^m]$, where $\mathrm{d}M^{n,m}(t) := 2\,\mathrm{e}^{2rt}\big\langle Z^{n,m}(t), (\sigma(\widehat{X}_t^n) - \sigma(\widehat{X}_t^m))\mathrm{d}W(t)\big\rangle$. By the stochastic Grownwall inequality [14, Lemma 5.4], for any $T > 0$, $p \in (0,1)$ and $q > \frac{1+p}{1-p}$, there exists a constant $c_1 > 0$ such that

$$\begin{aligned}\mathbb{E}\Big(\sup_{0\le t\le T\wedge\tau_R^n\wedge\tau_R^m}(\mathrm{e}^{2rt}|Z^{n,m}(t)|^2)\Big)^p &\le c_1\Big(\int_0^T \mathbb{E}(\|p_t^n\|_r^q \mathbf{1}_{\{t\le\tau_R^n\}})\mathrm{d}t\Big)^{p/q}\\ &\quad + c_1\Big(\int_0^T \mathbb{E}(\|p_t^m\|_r^q \mathbf{1}_{\{t\le\tau_R^m\}})\mathrm{d}t\Big)^{p/q}.\end{aligned} \tag{A.7}$$

A straightforward calculation leads to

$$\|p_t^n\|_r \le \int_{t_n}^t |b(\widehat{X}_s^n)|\mathrm{d}s + \sup_{t_n\le s\le t}\Big|\int_{t_n}^s \sigma(\widehat{X}_s^n)\mathrm{d}W(s)\Big|. \tag{A.8}$$

From (A.5) and by the local boundedness of σ and BDG's inequality, for some $M_R > 0$,

$$\begin{aligned}&\lim_{n\to\infty}\mathbb{E}\Big(\int_{t_n}^{t\wedge\tau_R^n}|b(\widehat{X}_s^n)|\mathrm{d}s\Big)^q + \mathbb{E}\Big(\sup_{t_n\le s\le t\wedge\tau_R^n}\Big|\int_{t_n}^s \sigma(\widehat{X}_s^n)\mathrm{d}W(s)\Big|^q\Big)\\ &\le \lim_{n\to\infty}\Big(\frac{C(R)}{n^q} + \frac{M_R}{n^{q/2}}\Big) = 0, \quad t \ge 0.\end{aligned} \tag{A.9}$$

Combining this with (A.8), we make a conclusion that

$$\sup_{t\in[0,T]}\lim_{n\to\infty}\mathbb{E}(\|p_t^n\|_r^q\mathbf{1}_{\{t\le\tau_R^n\}}) = 0, \tag{A.10}$$

which, together with (A.7) for $p = \frac{1}{2}$, implies that

$$\lim_{n,m\to\infty}\mathbb{E}\Big\{\sup_{0\le t\le T\wedge\tau_R^n\wedge\tau_R^m}\|X_t^n - X_t^m\|_r\Big\} = 0. \tag{A.11}$$

So, to ensure that $X^n_\cdot$ converges in probability to a solution of (3.2), it remains to prove

$$\lim_{R\to\infty}\limsup_{n\to\infty}\mathbb{P}(\tau_R^n\le T)=0. \tag{A.12}$$

Indeed, (A.11) and (A.12) yield that

$$\lim_{n,m\to\infty}\mathbb{P}\Big\{\sup_{0\le t\le T}\|X_t^n-X_t^m\|_r\ge\varepsilon\Big\}=0,\quad \varepsilon>0,$$

and thus, due to the completeness of $(\mathscr{C}_r,\|\cdot\|_r)$, there exists a continuous adapted process $(X_t)_{t\in[0,T]}$ on $\mathscr{C}_r$ such that

$$\sup_{0\le t\le T}\|X_t^n-X_t\|_r\to 0\qquad\text{in probability as } n\to\infty.$$

Subsequently, by carrying out a standard argument, we can show that $(X(t))_{t\in[0,T]}$ is the unique functional solution to (3.2) under assumptions (**D1**) and (**D2**). We now proceed to verify (A.12). By Itô's formula, besides (**D2**), there is a constant $c_2>0$ such that

$$\mathrm{d}(\mathrm{e}^{2rt}|X^n(t)|^2)\le c_2\mathrm{e}^{2rt}\Big\{1+|X^n(t)|^2+4\|X_t^n\|_r^2+\|p^n(t)\|_r\cdot|b(\widehat{X}_t^n)|\Big\}\mathrm{d}t+\mathrm{d}M^n(t), \tag{A.13}$$

where $\mathrm{d}M^n(t):=2\,\mathrm{e}^{2rt}\langle X^n(t),\sigma(\widehat{X}_t^n)\mathrm{d}W(t)\rangle$. By combining (A.5) and using (**D2**), BDG's inequality and Gronwall's inequality, there exists a constant $c_3>0$ such that

$$\begin{aligned}\Gamma^{n,R}(t):&=\mathbb{E}\Big(\sup_{0\le s\le t\wedge\tau_R^n}(\mathrm{e}^{2rs}|X^n(s)|^2)\Big)\\ &\le c_3\,\mathrm{e}^{c_3t}\Big\{\|\xi\|_r^2+t+\int_0^t\mathrm{e}^{2rs}\mathbb{E}\Big(|p^n(s)|\mathbf{1}_{\{s\le\tau_R^n\}}\Big)\mathrm{d}s\Big\},\quad t\ge 0\end{aligned} \tag{A.14}$$

holds for some constant $c_3>0$. Next, (A.10), (A.14) and Chebyshev's inequality gives

$$\begin{aligned}&\lim_{R\to\infty}\lim_{n\to\infty}\mathbb{P}(\tau_R^n\le T)=\lim_{R\to\infty}\lim_{n\to\infty}\mathbb{P}\Big(\tau_R^n\le T,\sup_{0\le t\le\tau_R^n\wedge T}|X^n(t)|\ge\frac{R}{4}\Big)\\ &\le\lim_{R\to\infty}\lim_{n\to\infty}\mathbb{P}\Big(\sup_{0\le t\le\tau_R^n\wedge T}|X^n(t)|\ge\frac{R}{4}\Big)\le\lim_{R\to\infty}\lim_{n\to\infty}\frac{16\Gamma^{n,R}(T)}{R^2}=0,\end{aligned}$$

where we used the fact that

$$\Big\{\tau_R^n\le T,\sup_{0\le t\le\tau_R^n\wedge T}|X^n(t)|<\frac{R}{4}\Big\}=\emptyset,$$

by the definition of τ_R^n. So, (A.12) holds.

In the end, by making use of (A.10) and (A.14) and employing Fatou's lemma for $n\to\infty$, we obtain

$$\mathbb{E}\Big(\sup_{0\le s\le t\wedge\tau_R}(\mathrm{e}^{2rs}\|X_s\|_r^2)\Big)\le c_4\big(1+\|\xi\|_r^2\big)\mathrm{e}^{c_4t},$$

where τ^R is defined as in (A.3) for X replacing X^n, which goes to ∞ as $R\to\infty$. Therefore, by approaching $R\uparrow\infty$, we achieve (A.1). □

References

[1] Arnaudon, M., Thalmaier A., Wang, F.-Y., Harnack inequality and heat kernel estimate on manifolds with curvature unbounded below, *Bull. Sci. Math.* **130** (2006), 223–233.

[2] Arnaudon, M., Thalmaier, A., Wang, F.-Y., Gradient estimates and Harnack inequalities on noncompact Riemannian manifolds, *Stoch. Proc. Appl.*, **119** (2009), 3653–3670.

[3] Bao, J., Wang, F.-Y., Yuan, C., Bismut formulae and applications for functional SPDEs, *Bull. Sci. Math.*, **137** (2013), 509–522.

[4] Bao, J., Wang, F.-Y., Yuan, C., Derivative formula and Harnack inequality for degenerate functionals SDEs, *Stoch. Dyn.*, **13** (2013), no. 1, 1250013, 22 pp.

[5] Butkovsky, O., Scheutzow, M., Invariant measures for stochastic functional differential equations, *Electron. J. Probab.*, **22** (2017), Paper No. 98, 23 pp.

[6] Da Prato, G., Zabczyk, J., Ergodicity for Infinite-Dimensional Systems, in: London Mathematical Society, Lecture Note Series, vol. 229, Cambridge University Press, Cambridge, 1996.

[7] Es-Sarhir, A., von Renesse, Max-K., Scheutzow, M., Harnack inequality for functional SDEs with bounded memory, *Electron. Commun. Probab.*, **14** (2009), 560–565.

[8] Es-Sarhir, A., Scheutzow, M., van Gaans, O., Invariant measures for stochastic functional differential equations with superlinear drift term, *Differential Integral Equations*, **23** (2010), 189–200.

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

[10] Hairer, M., Mattingly, J. C., Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing, *Ann. Math.*, **164** (2006), 993–1032.

[11] Mao, X., *Stochastic Differential Equations and Applications*, Horwood, England, Second Edition, 2008.

[12] Mattingly, J. C., Stuart, A. M., Higham, D. J., Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise, *Stochastic Process. Appl.*, **101** (2002), 185–232.

[13] Shao, J., Wang, F.-Y., Yuan, C., Harnack inequalities for stochastic (functional) differential equations with non-Lipschitzian coefficients, *Electron. J. Probab.*, **17** (2012), 18 pp.

[14] von Renesse, M.-K., Scheutzow, M., Existence and uniqueness of solutions of stochastic functional differential equations, *Random Oper. Stoch. Equ.*, **18** (2010), 267–284.

[15] Wang, F.-Y., Logarithmic Sobolev inequalities on noncompact Riemannian manifolds, *Probab. Theory Related Fields*, **109** (1997), 417–424.

[16] Wang, F.-Y., Harnack inequality and applications for stochastic generalized porous media equations, *Ann. Probab.*, **35** (2007), 1333–1350.

[17] Wang, F.-Y., Harnack inequalities on manifolds with boundary and applications, *J. Math. Pures Appl.*, **94** (2010), 304–321.

[18] Wang, F.-Y., *Harnack inequality for SDE with multiplicative noise and extension to Neumann semigroup on nonconvex manifolds,* Ann. Probab. 39(2011), 1449–1467.

[19] Wang, F.-Y., *Harnack inequalities and Applications for Stochastic Partial Differential Equations*, Springer, 2013, Berlin.

[20] Wang, F.-Y., Yuan, C., Harnack inequalities for functional SDEs with multiplicative noise and applications, *Stochastic Process. Appl.*, **121** (2011), 2692–2710.

[21] Wang, F.-Y., Hypercontractivity and applications for stochastic Hamiltonian systems, *J. Funct. Anal.*, **272** (2017), 5360–5383.

[22] Wang, F.-Y., Zhang, X., Degenerate SDE with Hölder-Dini Drift and Non-Lipschitz Noise Coefficient, *SIAM J. Math. Anal.*, **48** (2016), 2189–2226.

[23] Wang, F.-Y., Zhang, X., Degenerate SDEs in Hilbert spaces with rough drifts, *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, **18** (2015), 1550026, 25 pp.

[24] Wang, F.-Y., Zhang, X., Derivative formula and applications for degenerate diffusion semigroups, *J. Math. Pures Appl.*, **99** (2013), 726–740.

[25] Wu, F., Yin, G., Mei, H., Stochastic functional differential equations with infinite delay: existence and uniqueness of solutions, solution maps, Markov properties, and ergodicity, *J. Differential Equations*, **262** (2017), 1226–1252.

[26] Xu, L., A modified log-Harnack inequality and asymptotically strong Feller property, *J. Evol. Equ.*, **11** (2011), 925–942.

[27] Zhang, X., Stochastic flows and Bismut formulas for stochastic Hamiltonian systems, *Stoch. Proc. Appl.*, **120** (2010), 1929–1949.