Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 149 (2022) 1–38

stochastic processes and their applications

www.elsevier.com/locate/spa

Stochastic functional differential equations with infinite delay under non-Lipschitz coefficients: Existence and uniqueness, Markov property, ergodicity, and asymptotic log-Harnack inequality

Ya Wang[a,1], Fuke Wu[a,*,1], George Yin[b,2], Chao Zhu[c,3]

[a] *School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan, Hubei, 430074, PR China*

[b] *Department of Mathematics, University of Connecticut, Storrs, CT 06269-1009, USA*

[c] *Department of Mathematical Sciences, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA*

Received 2 January 2021; received in revised form 17 February 2022; accepted 11 March 2022
Available online 21 March 2022

Abstract

This paper focuses on a class of stochastic functional differential equations with infinite delay and non-Lipschitz coefficients. Under one-sided super-linear growth and non-Lipschitz conditions, this paper establishes the existence and uniqueness of strong solutions and strong Markov properties of the segment processes. Under additional assumption on non-degeneracy of the diffusion coefficient, exponential ergodicity for the segment process is derived by using asymptotic coupling method. In addition, the asymptotic log-Harnack inequality is established for the associated Markovian semigroup by using coupling and change of measures, which implies the asymptotically strong Feller property. Finally, an example is given to demonstrate these results.

MSC: 34K05; 34K30; 34K50; 60J60

Keywords: Stochastic functional differential equation; Infinite delay; Non-Lipschitz coefficient; Ergodicity; Asymptotic log-Harnack inequality; Asymptotic strong Feller property

* Corresponding author.
E-mail addresses: y_wang@hust.edu.cn (Y. Wang), wufuke@hust.edu.cn (F. Wu), gyin@uconn.edu (G. Yin), zhu@uwm.edu (C. Zhu).

[1] The research of the authors was supported in part by the National Natural Science Foundation of China (Grant No. 61873320).

[2] The research of this author was supported in part by the National Science Foundation, United States under grant DMS-2114649.

[3] The research of this author was supported in part by the Simons foundation collaboration, United States grant 523736 and a DIG award from the University of Wisconsin-Milwaukee.

https://doi.org/10.1016/j.spa.2022.03.008
0304-4149/

1. Introduction and motivation

Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ be a complete filtered probability space. For a given $r > 0$, define

$$\mathcal{C}_r = \left\{\phi \in C((-\infty, 0]; \mathbf{R}^d) : \sup_{-\infty<\theta\leq 0} e^{r\theta}|\phi(\theta)| < \infty\right\}$$

with norm $\|\phi\|_r = \sup_{-\infty<\theta\leq 0} e^{r\theta}|\phi(\theta)|$, where $C((-\infty, 0]; \mathbf{R}^d)$ denotes the family of continuous functions from $(-\infty, 0]$ to $\mathbf{R}^d$. Then $(\mathcal{C}_r, \|\cdot\|_r)$ is a Polish space (see [16] for more details on this space and its properties). In this paper, choosing $\mathcal{C}_r$ as the phase space, we consider the following stochastic functional differential equations (SFDEs) with infinite delay

$$dX(t) = b(X_t)dt + \sigma(X_t)dW(t), \;\; X_0 = \xi \in \mathcal{C}_r, \tag{1.1}$$

where $b : \mathcal{C}_r \mapsto \mathbf{R}^d$ and $\sigma : \mathcal{C}_r \mapsto \mathbf{R}^{d\times m}$ are continuous functionals, $W(t)$ is an m-dimensional Wiener process on $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathcal{F}, \mathbb{P})$ and $X_t(\tau) : (-\infty, 0] \ni \tau \mapsto X(t+\tau) \in \mathbf{R}^d$ denotes the segment process. To emphasize the dependence of the solution $X(t)$ on the initial data $X_0 = \xi \in \mathcal{C}_r$, we also write the solution $X(t)$ and the corresponding segment process X_t as $X(t, \xi)$ and $X_t(\xi)$, respectively.

For a stochastic differential equation (SDE) to have a unique global solution with a given initial value, commonly used assumptions are the linear growth and the local Lipschitz conditions. Under these conditions, various asymptotic behaviors of SDEs were also well studied; see for example, [17,19,28]. However, for many important stochastic models, the local Lipschitz condition is a rather restrictive assumption. For example, the diffusion coefficients in the Feller branching diffusion and the Cox–Ingersoll–Ross model are only Hölder continuous. Consequently, stochastic models with non-Lipschitz coefficients have received growing attention lately; see, for example, [9,13,35,36] and the references therein. In [13], Fang and Zhang studied a class of SDEs with non-Lipschitz coefficients and examined the existence and uniqueness of solutions, the dependence with respect to the initial value, and the large deviation principle. While SFDEs provide powerful mathematical tools in modeling and analyzing complex memory-dependent dynamical systems, the studies on SFDEs with non-Lipschitz coefficients are relatively scarce. In this paper, our main aim is to take up these issues for SFDEs with infinite delay. It is well known that solutions of stochastic functional or delay differential equations are non-Markov because of the dependence on the past history. Under non-Lipschitz condition, this paper examines the existence and uniqueness of the global solution $X(t)$ for SFDE (1.1), and the strong Markovian property, ergodicity, asymptotic log-Harnack inequality, and asymptotic strong Feller property for the segment process X_t.

The existence and uniqueness of invariant measures for SFDEs have been investigated in the literature under different settings. For example, by using the tightness criterion of probability measures on a continuous function space (e.g., [18, Theorem 4.10]) and Krylov–Bogoliubov's theorem (e.g., [10, Theorem 3.1.1]), Es-Sarhir et al. investigated the existence of an invariant measure for SFDEs with finite delay under the super-linear drift coefficient in [12]; but the paper did not establish the uniqueness of an invariant measure. By using an asymptotic coupling approach, Hairer et al. [15] obtained uniqueness of the invariant measure for SFDEs with finite delay. Recently, Butkovsky [6], Butkovsky and Scheutzow [8], and Butkovsky et al. [7] further developed Hairer's approach and provided sufficient conditions for existence and uniqueness of invariant measures for SFDEs with finite delay. A crucial assumption in these papers is that the diffusion coefficient is non-degenerate and its right inverse is uniformly bounded. Although such an assumption is removed in [4,5], the coefficients still need to satisfy certain dissipative conditions.

It is worth pointing out that the aforementioned papers only consider SFDEs with finite delay and Lipschitz-type coefficients. Under the dissipative condition, by using the remote start method, Wu et al. [33] obtained the ergodicity for SFDEs with infinite delay and Lipschitz-type coefficients. By using an asymptotic coupling approach, Bao et al. [3] investigated neutral type SFDEs with infinite delay and Lipschitz-type coefficients. Recently, Kulik and Scheutzow [20] established weak ergodic rates for SFDEs with finite delay and Hölder continuous coefficients. Motivated by the aforementioned developments, this paper aims to examine exponential ergodicity in the Wasserstein distance for SFDEs with infinite delay and non-Lipschitz continuous coefficients. In contrast to SFDEs with only finite delay, it is much more difficult to establish a support-type assertion (Lemma 3.4) for SFDEs with infinite delay. This, in turn, leads to much difficulty and subtlety in the construction of a contracting distance-like function satisfying the conditions of the weak Harris Theorem (Theorem 4.8 in [15]). Not only does the support-type assertion for SFDEs with infinite delay depend on the initial condition, but also on a reference number ε. Consequently we have to present an explicit dependence relationship among the parameters involved in the distance-like function and the time variable t.

Moreover, this paper examines the asymptotic log-Harnack inequality for SFDEs with infinite delay and non-Lipschitz continuous coefficients. The dimension-free Harnack inequality was first introduced by Wang [29] for diffusion semigroups on Riemannian manifolds. The weaker version of Harnack inequality (log-Harnack inequality) was established in [25,30] for elliptic diffusion processes. Further developments in the study of these inequalities can be found in [11,27,31,32]. It is worth noting that these two Harnack-type inequalities imply some regularity properties of the associated Markov semigroups such as strong Feller property. In some cases where the stochastic system has no strong Feller property or the above Harnack-type inequalities are unavailable, the modified/asymptotic log-Harnack inequality was introduced in [34], which implies the asymptotic strong Feller property. Recently, by using the asymptotic coupling method, the asymptotic log-Harnack inequality is established by Bao, Wang, and Yuan [2] for several stochastic differential systems with infinite delay, including SFDEs with infinite delay under Lipschitz-type coefficients. This paper aims to further this line of research and derive an asymptotic log-Harnack inequality for SFDEs with infinite delay *and non-Lipschitz* coefficients. To overcome the difficulties from the non-Lipschitz conditions, more delicate computations and stronger condition (see Assumption 4.1) are needed.

The rest of the paper is organized as follows. Section 2 establishes the existence and uniqueness of a global strong solution to (1.1) under weak non-Lipschitz conditions, and proves that the corresponding segment process is a strong Markovian process. Exponential ergodicity is investigated under the non-Lipschitz conditions in Section 3. Section 4 establishes the asymptotic log-Harnack inequality, which leads to the asymptotic strong Feller property for the segment process. Finally, an example is given in Section 5 to demonstrate our results.

To proceed, we introduce some notation and definitions that will be used in later sections. Denote by $\mathbf{R}^d$ the d-dimensional Euclidean space and $|\cdot|$ the Euclidean norm. If $a, b \in \mathbf{R}^d$, $\langle a, b\rangle$ denotes the standard inner product on $\mathbf{R}^d$. If A is a vector or a matrix, its transpose is denoted by $A^\top$. For a matrix A, denote its trace norm by $\|A\| = \sqrt{\mathrm{trace}(A^\top A)}$. $C^\infty(\mathbf{R}^d)$ denotes the family of infinitely differentiable functions $f : \mathbf{R}^d \to \mathbf{R}$. The indicator function of the set G is denoted by $\mathbf{1}_G$. Denote by $\mathcal{M}_0$ the set of probability measures on $(-\infty, 0]$. For any $k > 0$, let us further define $\mathcal{M}_k$, the subset of $\mathcal{M}_0$, by

$$\mathcal{M}_k := \left\{ \mu \in \mathcal{M}_0 : \mu^{(k)} := \int_{-\infty}^0 e^{-k\theta} \mu(d\theta) < \infty \right\}.$$

Let $(E, \mathcal{B}(E), d)$ be a Polish space and denote by $\mathcal{P}(E)$ the family of probability measures on $(E, \mathcal{B}(E), d)$. For $\mu, \nu \in \mathcal{P}(E)$, $\mathcal{C}(\mu, \nu)$ denotes the collection of all couplings of μ and ν, that is, probability measures on $E \times E$ with marginal distributions μ and ν. For the metric d on E, the associated L^1-Wasserstein distance between two probability measures $\mu, \nu \in \mathcal{P}(E)$ is defined as follows:

$$\mathbb{W}_d(\mu, \nu) = \inf_{\Pi \in \mathcal{C}(\mu,\nu)} \int_{E\times E} d(\xi, \eta)\Pi(d\xi, d\eta).$$

For probability measures $\mu, \nu \in \mathcal{P}(E)$ satisfying $\mu \ll \nu$, the Kullback–Leibler divergence of μ from ν is defined by

$$D_{KL}(\mu \parallel \nu) := \int_E \log \frac{d\mu}{d\nu} d\mu = \int_E \frac{d\mu}{d\nu} \log \frac{d\mu}{d\nu} d\nu.$$

2. Existence and uniqueness of solution and Markov property

This section is devoted to the existence and uniqueness of a solution to SFDEs (1.1) and Markov properties of the segment processes under non-Lipschitz conditions. Our approach can be described as follows. First, we establish the existence of a weak solution to (2.2) (which can be regarded as an approximation to (1.1)) and the pathwise uniqueness result. Therefore, by the Yamada–Watanabe principle [35], Eq. (2.2) has a unique global strong solution. This implies that (1.1) has a unique maximal local strong solution. Then we show that the maximal local strong solution to (1.1) is non-explosive under Assumption 2.2. This leads to the desired assertion that (1.1) has a unique global strong solution. Then we show the strong Markov and Feller properties of the segment process X_t to (1.1).

Existence and uniqueness of solution

To characterize the non-Lipschitz coefficients of (1.1), we introduce the following class of functions:

$$\mathcal{U} = \left\{ u \mid_{(0,\infty)\to[1,\infty)} : \int_{0+} \frac{ds}{su(s)} = \infty, \;\; s \mapsto su(s) \text{ is increasing and concave} \right\}.$$

One can verify $u(s) = \log(\mathrm{e} \vee s^{-1})$ and $u(s) = \log\big((1 + s^{-1}) \vee e\big) \in \mathcal{U}$. Noting that $\lim_{s\to 0} su(s) = 0$ for $u \in \mathcal{U}$, we set $0u(0) = \lim_{s\to 0} su(s) = 0$ without loss of generality. To ensure the existence and uniqueness of the solution, we make the following assumptions.

Assumption 2.1. b is continuous and bounded on bounded subset of $\mathcal{C}_r$. There exist a positive constant δ and a function $u \in \mathcal{U}$ such that for all $k > 0$ and $\phi, \psi \in \mathcal{C}_r$ with $\|\phi\|_r \vee \|\psi\|_r \le k$ and $\|\phi - \psi\|_r \le \delta$,

$$2\langle \phi(0) - \psi(0), b(\phi) - b(\psi)\rangle_+ + \|\sigma(\phi) - \sigma(\psi)\|^2 \le L_k \|\phi - \psi\|_r^2 u(\|\phi - \psi\|_r^2), \tag{2.1}$$

where L_k is a positive constant depending on k and $a_+ := \max\{0, a\}$ for any $a \in \mathbf{R}$.

Note that the non-Lipschitz condition (2.1) in Assumption 2.1 is only required to hold in a small neighborhood of the diagonal line $\phi = \psi$ in $\mathcal{C}_r \times \mathcal{C}_r$ with $\|\phi\|_r \vee \|\psi\|_r \le k$ for all $k > 0$. This is in stark contrast to the standard local Lipschitz condition, which significantly relaxes the conditions used in the literature such as [2,33].

Assumption 2.2. σ is bounded on bounded subset of $\mathcal{C}_r$ and there exists a non-decreasing function $\zeta(\cdot) : [0, \infty) \mapsto (0, \infty)$ such that $\int_0^\infty 1/\zeta(x)dx = \infty$ and for all $\phi \in \mathcal{C}_r$,

$$2\langle \phi(0), b(\phi)\rangle + \|\sigma(\phi)\|^2 \le \zeta(\|\phi\|_r^2).$$

Theorem 2.1. *Under Assumption 2.1, (1.1) has a unique maximal local strong solution for any initial data $X_0 = \xi \in \mathcal{C}_r$. Under Assumption 2.2, any maximal local strong solution to (1.1) is non-explosive in any finite time almost surely.*

Proof. We divide the proof into two steps. The first step shows that under Assumption 2.1, Eq. (1.1) has a unique maximal local strong solution. The second step proves that any maximal local solution to Eq. (1.1) is non-explosive in any finite time a.s. under Assumption 2.2.

Step 1: Maximal local strong solution to (1.1). For any $m \ge 1$, we can find $h_m \in C^\infty(\mathbf{R})$ with compact support contained in $\mathbf{S}_{m+1}$ such that $h_m|_{\mathbf{S}_m} = 1$, where $\mathbf{S}_m := \{x \in \mathbf{R} : |x| \le m\}$. Let

$$b_m(\phi) = b(\phi)h_m(\|\phi\|_r), \quad \sigma_m(\phi) = \sigma(\phi)h_m(\|\phi\|_r).$$

Since b and σ are bounded on bounded subset of $\mathcal{C}_r$, b_m and σ_m are uniformly bounded on $\mathcal{C}_r$ and satisfy Assumption 2.1 for any $m \ge 1$.

Fix $m \ge 1$ arbitrarily. We first consider the following equation

$$dX(t) = b_m(X_t)dt + \sigma_m(X_t)dW(t). \tag{2.2}$$

Noting that b_m and σ_m are uniformly bounded on $\mathcal{C}_r$, there exist two sequences of uniformly bounded local Lipschitz continuous (i.e., Lipschitz continuous on bounded subset of $\mathcal{C}_r$) functions $\{b_m^n\}_{n\ge1}$ and $\{\sigma_m^n\}_{n\ge1}$ such that Assumption 2.1 holds for b_m^n, σ_m^n uniformly in n (that is, the constants do not depend on n). Moreover, b_m^n and σ_m^n converge to b_m and σ_m as $n \to \infty$ uniformly on each compact subset of $\mathcal{C}_r$. Therefore for each $n \in \mathbb{N}$, the following equation

$$dX^n(t) = b_m^n(X_t^n)dt + \sigma_m^n(X_t^n)dW(t), \quad X_0^n = \xi$$

has a unique global strong solution. Note that b_m^n, σ_m^n are uniformly bounded, that is, there exists a constant C_m independent of n such that

$$|b_m^n(\phi)| \vee \|\sigma_m^n(\phi)\| \le C_m, \quad \forall \phi \in \mathcal{C}_r.$$

Hence it is easy to see that for any $0 \le s, t \le T < \infty$, we have

$$\begin{aligned}\sup_{n\ge1}\mathbb{E}|X^n(t) - X^n(s)|^4 &\le 8\sup_{n\ge1}\mathbb{E}\left|\int_s^t b_m^n(X_v^n)dv\right|^4 + 8\sup_{n\ge1}\mathbb{E}\left|\int_s^t \sigma_m^n(X_v^n)dW(v)\right|^4 \\ &\le 8C_m^4|t-s|^4 + 8C_P C_m|t-s|^2 \\ &\le 8C_m^4(|t-s|^2 + C_P)|t-s|^2,\end{aligned} \tag{2.3}$$

where C_p is the coefficient of the Burkholder–Davis–Gundy inequality. Since $T > 0$ is arbitrary, according to [18, Problem 2.4.11], (2.3) implies that $\{\mathbb{P}^n\}_{n\ge1}$, the family of probability law of X^n, is tight. Hence there exists a probability measure $\mathbb{P}^\infty$ on $C([0,\infty); \mathbf{R}^d)$ such that $\mathbb{P}^n$ (up to a sub-sequence) converges weakly to $\mathbb{P}^\infty$. Let $\mathcal{G}_t = \sigma\{\omega(s) : 0 \le s \le t\}$, $t \ge 0$, $\omega \in C([0,\infty); \mathbf{R}^d)$. Then the coordinate process

$$Z(t)(\omega) := \omega(t), \quad t > 0, \omega \in C([0,\infty); \mathbf{R}^d)$$

is $\mathcal{G}_t$-adapted. Note that

$$X^n(t) - X^n(0) - \int_0^t b_m^n(X_s^n)ds$$

is a martingale relative to $(\mathbb{P}, \mathcal{F}_t)$ with cross-variation given by

$$\sum_{k=1}^m \int_0^t \{(\sigma_m^n)_{ik}(\sigma_m^n)_{jk}\}(X_s^n)ds, \quad 1 \le i, j \le d.$$

Since $\mathbb{P}^n$ is the distribution of X^n on $C([0, \infty); \mathbf{R}^d)$,

$$M^n(t) := Z(t) - Z(0) - \int_0^t b_m^n(Z_s)ds$$

is a martingale relative to $(\mathbb{P}^n, \mathcal{G}_t)$ with cross-variation

$$\langle M_i^n, M_j^n\rangle(t) = \sum_{k=1}^m \int_0^t \{(\sigma_m^n)_{ik}(\sigma_m^n)_{jk}\}(Z_s)ds, \quad 1 \le i, j \le d,$$

where Z_s denotes the corresponding segment process (by choosing $Z_0 = X_0^n = \xi$). Since b_m^n is uniformly bounded, and as $n \to \infty$, b_m^n converges to b_m uniformly on compact subsets of $\mathcal{C}_r$, $M^n(t)$ converges to

$$M(t) := Z(t) - Z(0) - \int_0^t b_m(Z_s)ds$$

on $C([0, \infty); \mathbf{R}^d)$ and the convergence is uniform on compact subsets of $\mathcal{C}_r$. Then for any given $s < t$ and $A \in \mathcal{G}_s$, by [18, Problem 2.4.12] and the martingale property of $M^n(t)$ relative to $(\mathbb{P}^n, \mathcal{G}_t)$

$$\mathbb{E}^\infty 1_A M(t) = \lim_{n\to\infty} \mathbb{E}^n 1_A M^n(t) = \lim_{n\to\infty} \mathbb{E}^n 1_A M^n(s) = \mathbb{E}^\infty 1_A M(s),$$

where $\mathbb{E}^n$ denotes the expectation operator with respect to the measure $\mathbb{P}^n$ and likewise, $\mathbb{E}^\infty$ denotes the expectation with respect to $\mathbb{P}^\infty$. This implies that $M(t)$ is a $\mathbb{P}^\infty$-martingale. In addition, noting that σ_m^n is uniformly bounded, and as $n \to \infty$, σ_m^n converges to σ_m uniformly on each compact subset of $\mathcal{C}_r$, by a similar argument as before, we obtain

$$\langle M_i, M_j\rangle(t) = \sum_{k=1}^m \int_0^t \{(\sigma_m)_{ik}(\sigma_m)_{jk}\}(Z_s)ds, \quad 1 \le i, j \le d.$$

Then it follows from [17, Theorem II.7.1′] that there exists an m-dimensional Brownian motion $\widetilde{W}$ on an extended probability space of $(C([0, \infty); \mathbf{R}^d), \mathcal{B}(C([0, \infty); \mathbf{R}^d)), \{\mathcal{G}_t\}, \mathbb{P}^\infty)$ such that

$$M(t) = \int_0^t \sigma_m(Z_s)d\widetilde{W}(s).$$

As a result, $Z(t) = Z(0) + \int_0^t b_m(Z_s)ds + \int_0^t \sigma_m(Z_s)d\widetilde{W}(s)$ and hence Eq. (2.2) has a weak solution.

Now we prove the pathwise uniqueness for Eq. (2.2). Suppose that X and Y satisfy

$$X(t) = \xi(0) + \int_0^t b_m(X_s)ds + \int_0^t \sigma_m(X_s)dW(s), \quad X_0 = \xi,$$
$$Y(t) = \eta(0) + \int_0^t b_m(Y_s)ds + \int_0^t \sigma_m(Y_s)dW(s), \quad Y_0 = \eta,$$

for all $t \geq 0$. Assume $\|\xi - \eta\|_r < \delta_0 \leq \delta$ and define the stopping time

$$S_{\delta_0} = \inf\{t \geq 0 : |X(t) - Y(t)| > \delta_0\}.$$

For $R > \|\xi\|_r \vee \|\eta\|_r$, define another stopping time

$$T_R = \inf\{t \geq 0 : |X(t)| \vee |Y(t)| > R\}.$$

Applying Itô's formula and using Assumption 2.1, we have

$$\begin{aligned}
&|X(t \wedge T_R \wedge S_{\delta_0}) - Y(t \wedge T_R \wedge S_{\delta_0})|^2 \\
&\quad = |\xi(0) - \eta(0)|^2 + \int_0^{t\wedge T_R \wedge S_{\delta_0}} 2(X(v) - Y(v))^\top(\sigma_m(X_v) - \sigma_m(Y_v))dW(v) \\
&\qquad + \int_0^{t\wedge T_R \wedge S_{\delta_0}} \big(2\langle X(v) - Y(v), b_m(X_v) - b_m(Y_v)\rangle + \|\sigma_m(X_v) - \sigma_m(Y_v)\|^2\big)\, dv \\
&\quad \leq |\xi(0) - \eta(0)|^2 + \int_0^{t\wedge T_R \wedge S_{\delta_0}} L_R \|X_v - Y_v\|_r^2 u(\|X_v - Y_v\|_r^2) dv \\
&\qquad +2\int_0^{t\wedge T_R \wedge S_{\delta_0}} (X(v) - Y(v))^\top(\sigma_m(X_v) - \sigma_m(Y_v))dW(v).
\end{aligned} \tag{2.4}$$

By the Burkholder–Davis–Gundy inequality, Assumption 2.1, and the Young inequality, we have

$$\begin{aligned}
&2\mathbb{E}\Big[\sup_{0\leq s\leq t} \int_0^{s\wedge T_R \wedge S_{\delta_0}} (X(v) - Y(v))^\top(\sigma_m(X_v) - \sigma_m(Y_v))dW(v)\Big] \\
&\quad \leq 2\sqrt{32}\mathbb{E}\left(\int_0^{t\wedge T_R \wedge S_{\delta_0}} |(X(v) - Y(v))^\top(\sigma_m(X_v) - \sigma_m(Y_v))|^2 dv\right)^{\frac{1}{2}} \\
&\quad \leq \frac{1}{2}\mathbb{E}\Big[\sup_{0\leq s\leq t} |X(s\wedge T_R \wedge S_{\delta_0}) - Y(s\wedge T_R \wedge S_{\delta_0})|^2\Big] \\
&\qquad +64 L_R \mathbb{E}\int_0^{t\wedge T_R \wedge S_{\delta_0}} \|X_s - Y_s\|_r^2 u(\|X_s - Y_s\|_r^2) ds.
\end{aligned} \tag{2.5}$$

Noting that

$$\|X_s - Y_s\|_r^2 \leq \|\xi - \eta\|_r^2 \ \vee \sup_{0\leq v\leq s} |X(v) - Y(v)|^2$$

and that the function $su(s)$ is nondecreasing, we get

$$\begin{aligned}
&\|X_s - Y_s\|_r^2 u(\|X_s - Y_s\|_r^2) \\
&\quad \leq \|\xi - \eta\|_r^2 u(\|\xi - \eta\|_r^2) + \sup_{0\leq v\leq s} |X(v) - Y(v)|^2 u(\sup_{0\leq v\leq s} |X(v) - Y(v)|^2).
\end{aligned}$$

Denote $\Delta(t) = |X(t) - Y(t)|^2$. Combining this with (2.4) and (2.5), we arrive at

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{0\leq s\leq t} \Delta(s\wedge T_R \wedge S_{\delta_0})\Big] \\
&\quad \leq 2\|\xi - \eta\|_r^2 + 130 L_R \mathbb{E}\int_0^{t\wedge T_R \wedge S_{\delta_0}} \|\xi - \eta\|_r^2 u(\|\xi - \eta\|_r^2) ds \\
&\qquad + 130 L_R \mathbb{E}\int_0^{t\wedge T_R \wedge S_{\delta_0}} \sup_{0\leq v\leq s} \Delta(v) u(\sup_{0\leq v\leq s} \Delta(v)) ds
\end{aligned}$$

$$
\begin{aligned}
&\le 2\|\xi-\eta\|_r^2 + 130L_R\|\xi-\eta\|_r^2 u(\|\xi-\eta\|_r^2)t \\
&\quad + 130L_R\mathbb{E}\int_0^t \sup_{0\le v\le s} \Delta(v\wedge T_R\wedge S_{\delta_0})u(\sup_{0\le v\le s} \Delta(v\wedge T_R\wedge S_{\delta_0}))ds \\
&\le 2\|\xi-\eta\|_r^2 + 130L_R\|\xi-\eta\|_r^2 u(\|\xi-\eta\|_r^2)t \\
&\quad + 130L_R\int_0^t \mathbb{E}\Big[\sup_{0\le v\le s} \Delta(v\wedge T_R\wedge S_{\delta_0})\Big]u(\mathbb{E}\sup_{0\le v\le s} \Delta(v\wedge T_R\wedge S_{\delta_0}))ds,
\end{aligned}
$$

where we have used the concavity of $su(s)$ and Jensen's inequality to derive the last inequality. Define $\varsigma(t) = \mathbb{E}\sup_{0\le s\le t} \Delta(s\wedge T_R\wedge S_{\delta_0})$. Then we have

$$
0 \le \varsigma(t) \le 2\|\xi-\eta\|_r^2 + 130L_R\|\xi-\eta\|_r^2 u(\|\xi-\eta\|_r^2)t + 130L_R\int_0^t \varsigma(s)u(\varsigma(s))ds =: \iota(t).
$$

Define $G(t) = \int_1^t \frac{1}{su(s)}ds$ for $t > 0$. Since $\int_{0+} \frac{1}{su(s)}ds = \infty$, $\lim_{t\downarrow 0} G(t) = -\infty$. In addition, G is nondecreasing and satisfies $G(t) > -\infty$ for $t > 0$. Then we have

$$
\begin{aligned}
G(\varsigma(t)) &\le G(\iota(t)) = G(\iota(0)) + \int_0^t G'(\iota(s))d\iota(s) \\
&= G(2\|\xi-\eta\|_r^2) + \int_0^t \frac{130L_R\|\xi-\eta\|_r^2 u(\|\xi-\eta\|_r^2) + 130L_R\varsigma(s)u(\varsigma(s))}{\iota(s)u(\iota(s))}ds \\
&\le G(2\|\xi-\eta\|_r^2) + 260tL_R,
\end{aligned}
\tag{2.6}
$$

where we used $su(s)$ being nondecreasing to derive the last inequality. It is readily seen that the right-hand side of (2.6) converges to $-\infty$ as $\|\xi-\eta\|_r \to 0$, so does the left-hand side. Therefore, we obtain

$$
\begin{aligned}
\lim_{\|\xi-\eta\|_r\to 0} \varsigma(t) &= \lim_{\|\xi-\eta\|_r\to 0} \mathbb{E}\Big[\sup_{0\le s\le t} \Delta(s\wedge T_R\wedge S_{\delta_0})\Big] \\
&= \lim_{\|\xi-\eta\|_r\to 0} \mathbb{E}\Big[\sup_{0\le s\le t\wedge T_R\wedge S_{\delta_0}} |X(s)-Y(s)|^2\Big] \\
&= 0.
\end{aligned}
\tag{2.7}
$$

In particular, if $\|\xi-\eta\|_r = 0$, we have $\mathbb{E}[\sup_{0\le s\le t\wedge T_R\wedge S_{\delta_0}} |X(s)-Y(s)|^2] = 0$. This, together with Fatou's lemma and the fact that $\lim_{R\to\infty} T_R = \infty$ a.s., further leads to $\mathbb{E}[\sup_{0\le s\le t\wedge S_{\delta_0}} |X(s)-Y(s)|^2] = 0$. Note that on the event $\{S_{\delta_0} \le t\}$, we have $|X(S_{\delta_0}) - Y(S_{\delta_0})| = \delta_0$. Thus we have $\delta_0\mathbb{P}\{S_{\delta_0} \le t\} \le \mathbb{E}[\sup_{0\le s\le t\wedge S_{\delta_0}} |X(s)-Y(s)|^2] = 0$ and hence $\mathbb{P}\{S_{\delta_0} \le t\} = 0$. Then

$$
\begin{aligned}
\mathbb{E}\Big[\sup_{0\le s\le t} |X(s)-Y(s)|^2\Big] &= \mathbb{E}\Big[\sup_{0\le s\le t} |X(s)-Y(s)|^2\mathbf{1}_{\{t\le S_{\delta_0}\}}\Big] \\
&\quad + \mathbb{E}\Big[\sup_{0\le s\le t} |X(s)-Y(s)|^2\mathbf{1}_{\{t> S_{\delta_0}\}}\Big] \\
&\le \mathbb{E}\Big[\sup_{0\le s\le t\wedge S_{\delta_0}} |X(s)-Y(s)|^2\Big] + 0 = 0;
\end{aligned}
$$

the second summand above equals zero because $\mathbb{E}[\sup_{0\le s\le t} |X(s)-Y(s)|^2] < \infty$ thanks to the uniform boundedness of b_m and σ_m. Therefore, it follows that $\mathbb{P}\{\sup_{0\le s\le t} |X(s)-Y(s)| = 0\} = 1$ and hence the pathwise uniqueness for (2.2) holds. Consequently, according to the Yamada–Watanabe principle, (2.2) has a unique global strong solution.

Finally, by a standard argument (e.g., [33, Lemma 3.1]), (1.1) has a unique maximal local strong solution.

Step 2: Non-explosion. Let $X(t)$ be the maximal local strong solution to (1.1) and τ be its explosion time or life time, that is, $\limsup_{t\to\tau}|X(t)|=\infty$. For any $t\in[0,\tau)$, by Itô's formula, using Assumption 2.2 gives

$$\begin{aligned}|X(t)|^2 &= |X(0)|^2+\int_0^t 2\langle X(s),b(X_s)\rangle+\|\sigma(X_s)\|^2ds+2\int_0^t X^\top(s)\sigma(X_s)dW(s)\\ &\le |\xi(0)|^2+\int_0^t \zeta(\|X_s\|_r^2)ds+N(t),\end{aligned}$$

where $\{N(t)\}_{t\in[0,\tau)}$ is a continuous local martingale with $N(0)=0$. Note that $\|X_s\|_r^2\le\|\xi\|_r^2+\sup_{0\le v\le s}|X(v)|^2$ and $\zeta(\cdot)$ is non-decreasing. We see that

$$\|\xi\|_r^2+|X(t)|^2\le 2\|\xi\|_r+\int_0^t \zeta(\|\xi\|_r^2+\sup_{0\le v\le s}|X(v)|^2)ds+N(t).$$

Then by the stochastic Gronwall lemma (see, e.g., [23, Lemma 5.1]), we have $\tau=\infty$ almost surely. Hence the maximal local strong solution $X(t)$ is actually non-explosive in any finite time under Assumption 2.2.

Combining the results of Steps 1 and 2 completes the proof of this theorem. □

Remark 2.2. We now give a specific construction of sequences $\{b_m^n\}_{n\ge1}$ and $\{\sigma_m^n\}_{n\ge1}$. Take a sequence of non-negative, twice continuously differentiable functions $\{\rho_n\}_{n\ge1}$ such that

$$\mathrm{supp}(\rho_n)\subset\left\{x\in\mathbf{R}^n:|x|\le\frac{1}{n}\right\}\quad\text{and}\quad\int_{\mathbf{R}^n}\rho_n(x)dx=1.$$

Note that $\mathcal{C}_r$ is isomorphic to $C([-1,0];\mathbf{R}^d)$ (see [33]). Then $\mathcal{C}_r$ has the Schauder basis $\{e_i\}_{i=1}^\infty\in\mathcal{C}_r$ since $C([-1,0];\mathbf{R}^d)$ has the corresponding basis (see, e.g., [22]). Let Q_n denote the projection mapping from $\mathcal{C}_r$ to $\{e_1,e_2,\dots,e_n\}$, that is, $Q_n(\sum_{i=1}^\infty x_ie_i)=\sum_{i=1}^n x_ie_i$. $\bar{Q}_n(\phi)$ denotes the coordinate coefficients of the projection of $\phi\in\mathcal{C}_r$ on $\{e_1,e_2,\dots,e_n\}$, that is, $Q_n(\phi)=\bar{Q}_n(\phi)^\top(e_1,e_2,\dots,e_n)^\top$. Define

$$\begin{aligned}b_m^n(\phi)&=\int_{\mathbf{R}^n}\rho_n\big(x-\bar{Q}_n(\phi)\big)b_m\left(\sum_{i=1}^n x_ie_i\right)dx,\\ \sigma_m^n(\phi)&=\int_{\mathbf{R}^n}\rho_n\big(x-\bar{Q}_n(\phi)\big)\sigma_m\left(\sum_{i=1}^n x_ie_i\right)dx.\end{aligned}$$

It is readily verified that $\{b_m^n\}$ and $\{\sigma_m^n\}$ defined above satisfy the desired property.

Remark 2.3. By using a similar approach as in [33, Theorem 4.1], it is easy to verify that the segment process X_t is continuous and $\mathcal{F}_t$-adapted.

Markov property

Proposition 2.4. *Let Assumptions 2.1 and 2.2 hold. Then the segment process $X=(X_t)_{t\ge0}$ to (1.1) is a strong Markov process.*

Proof. By using the standard technique (see, e.g., [8, Proposition 4.1]), it is easy to see from the strong uniqueness or pathwise uniqueness that the segment process X_t to (1.1) is Markov. In addition, since X_t has continuous trajectories, the strong Markov property follows from the Feller property (see, e.g., [24, Theorem 3.3.1]).

For $R > \|\xi\|_r \vee \|\eta\|_r$, define the stopping times

$$\tau_R(\xi) = \inf\{t \geq 0 : \|X_t(\xi)\|_r > R\}, \quad \text{and} \quad \tau_R(\eta) = \inf\{t \geq 0 : \|X_t(\eta)\|_r > R\}.$$

In fact, $\tau_R(\xi) = \inf\{t \geq 0 : |X(t;\xi)| > R\}$ for $R > \|\xi\|_r$. Then the non-explosion implies that for any $\varepsilon > 0$, η with $\|\xi - \eta\|_r < \delta_0$, there exists an $R > 0$ large enough such that

$$\mathbb{P}\{\tau_R(\xi) \wedge \tau_R(\eta) < t\} < \frac{\varepsilon}{4}, \tag{2.8}$$

where the constant R does not depend on η. In addition,

$$\begin{aligned} &\delta_0^2 \mathbb{P}\{\bar{S}_{\delta_0} \leq t \wedge \tau_R(\xi) \wedge \tau_R(\eta)\} \\ &\quad \leq \mathbb{E}|X(t \wedge \tau_R(\xi) \wedge \tau_R(\eta) \wedge \bar{S}_{\delta_0}; \xi) - X(t \wedge \tau_R(\xi) \wedge \tau_R(\eta) \wedge \bar{S}_{\delta_0}; \eta)|^2, \end{aligned}$$

where $\bar{S}_{\delta_0} := \inf\{t \geq 0 : |X(t;\xi) - X(t;\eta)| \geq \delta_0\} = \inf\{t \geq 0 : \|X_t(\xi) - X_t(\eta)\|_r \geq \delta_0\}$. Then, for any $\varepsilon > 0$,

$$\begin{aligned} &\mathbb{P}\{\|X_t(\xi) - X_t(\eta)\|_r > \varepsilon\} \\ &\quad = \mathbb{P}\{\|X_t(\xi) - X_t(\eta)\|_r > \varepsilon, \tau_R(\xi) \wedge \tau_R(\eta) < t\} \\ &\qquad + \mathbb{P}\{\|X_t(\xi) - X_t(\eta)\|_r > \varepsilon, \tau_R(\xi) \wedge \tau_R(\eta) \geq t, \bar{S}_{\delta_0} > t\} \\ &\qquad + \mathbb{P}\{\|X_t(\xi) - X_t(\eta)\|_r > \varepsilon, \tau_R(\xi) \wedge \tau_R(\eta) \geq t, \bar{S}_{\delta_0} \leq t\} \\ &\quad \leq \mathbb{P}\{\tau_R(\xi) \wedge \tau_R(\eta) < t\} + \mathbb{P}\{\|X_{t\wedge\tau_R(\xi)\wedge\tau_R(\eta)\wedge\bar{S}_{\delta_0}}(\xi) - X_{t\wedge\tau_R(\xi)\wedge\tau_R(\eta)\wedge\bar{S}_{\delta_0}}(\eta)\|_r > \varepsilon\} \\ &\qquad + \mathbb{P}\{\bar{S}_{\delta_0} \leq t \wedge \tau_R(\xi) \wedge \tau_R(\eta)\} \\ &\quad \leq \mathbb{P}\{\tau_R(\xi) \wedge \tau_R(\eta) < t\} + \frac{1}{\varepsilon^2}\mathbb{E}\|X_{t\wedge\tau_R(\xi)\wedge\tau_R(\eta)\wedge\bar{S}_{\delta_0}}(\xi) - X_{t\wedge\tau_R(\xi)\wedge\tau_R(\eta)\wedge\bar{S}_{\delta_0}}(\eta)\|_r^2 \\ &\qquad + \frac{1}{\delta_0^2}\mathbb{E}|X(\xi, t \wedge \tau_R(\xi) \wedge \tau_R(\eta) \wedge \bar{S}_{\delta_0}) - X(\eta, t \wedge \tau_R(\xi) \wedge \tau_R(\eta) \wedge \bar{S}_{\delta_0})|^2. \end{aligned} \tag{2.9}$$

Similar to (2.7), we have

$$\lim_{\|\eta-\xi\|_r \to 0} \mathbb{E}\Big[\sup_{0\leq s\leq t\wedge\tau_R(\xi)\wedge\tau_R(\eta)\wedge\bar{S}_{\delta_0}} |X(s;\xi) - X(s;\eta)|^2\Big] = 0 \tag{2.10}$$

Recall that $\|X_t - Y_t\|_r^2 \leq \|\xi - \eta\|_r^2 + \sup_{0\leq s\leq t} |X(s) - Y(s)|^2$. It follows from (2.8), (2.9), and (2.10) that there exists a positive constant $\tilde{\delta}_0 < \delta_0$ such that for $\|\xi - \eta\|_r \leq \tilde{\delta}_0$, we have

$$\mathbb{P}\{\|X_t(\xi) - X_t(\eta)\|_r > \varepsilon\} \leq \varepsilon. \tag{2.11}$$

Since $\varepsilon > 0$ is arbitrary, $\|X_t(\xi) - X_t(\eta)\|_r$ converges to 0 in probability as $\|\xi - \eta\|_r \to 0$. Thus $X_t(\xi)$ converges to $X_t(\eta)$ in distribution as $\xi \to \eta$ in the norm $\|\cdot\|_r$. This implies further that for any bounded continuous function $F : \mathcal{C}_r \to \mathbf{R}$, $\mathbb{E}[F(X_t(\xi))]$ converges to $\mathbb{E}[F(X_t(\eta))]$ as $\xi \to \eta$ in the norm $\|\cdot\|_r$. Therefore, X_t is a Feller process. The proof is completed. □

3. Ergodicity

To prove the ergodicity we need to modify Assumptions 2.1 and 2.2 as follows:

Assumption 3.1. b is continuous and bounded on bounded subset of $\mathcal{C}_r$ and satisfies the one-sided linear growth condition, that is, there exists a constant $L > 0$ such that for any $\phi \in \mathcal{C}_r$,

$$\langle \phi(0), b(\phi) \rangle \le L(1 + \|\phi\|_r^2).$$

And there exist a function $u \in \mathcal{U}$ and positive constants K, δ and $\beta \in (0, 1)$ such that for all $\phi, \psi \in \mathcal{C}_r$ with $\|\phi - \psi\|_r \le \delta$,

$$\begin{aligned} &2\langle \phi(0) - \psi(0), b(\phi) - b(\psi) \rangle_+ + \|\sigma(\phi) - \sigma(\psi)\|^2 \\ &\quad \le K\big[(\|\phi - \psi\|_r^2 u(\|\phi - \psi\|_r^2)) \wedge \|\phi - \psi\|_r^{1+\beta}\big]. \end{aligned} \tag{3.1}$$

From Assumption 3.1, the diffusion coefficient σ satisfies the following linear growth condition: there exists a positive constant $\bar{K}$ such that for any $\phi \in \mathcal{C}_r$

$$\|\sigma(\phi)\| \le \bar{K}(1 + \|\phi\|_r).$$

In addition, to construct asymptotic couplings by change of measures, we need to impose the following condition on the diffusion coefficient σ.

Assumption 3.2. For any $\phi \in \mathcal{C}_r$, the matrix $\sigma(\phi)$ admits a right inverse $\sigma^{-1}(\phi)$ and

$$\|\sigma^{-1}\|_\infty := \sup_{\phi \in \mathcal{C}_r} \|\sigma^{-1}(\phi)\| < \infty.$$

In this paper, we consider the following function on $\mathcal{C}_r \times \mathcal{C}_r$: for $\xi, \eta \in \mathcal{C}_r$

$$d_{N,\gamma}(\xi, \eta) = (N\|\xi - \eta\|_r^\gamma) \wedge 1, \quad N \ge 1, \quad \gamma \in (0, \beta).$$

Clearly, each $d_{N,\gamma}$ is a bounded metric on $\mathcal{C}_r$ and is equivalent to the usual distance $\|\cdot - \cdot\|_r$ in the sense of topology. Therefore, its corresponding L^1-Wasserstein distance is a metric on $\mathcal{P}(\mathcal{C}_r)$, and convergence in this metric is equivalent to weak convergence in $\mathcal{P}(\mathcal{C}_r)$, where $\mathcal{P}(\mathcal{C}_r)$ denotes the family of probability measures on $\mathcal{C}_r$. Denote

$$d_\gamma(\xi, \eta) := d_{1,\gamma}(\xi, \eta) = \|\xi - \eta\|_r^\gamma \wedge 1, \quad \xi, \eta \in \mathcal{C}_r.$$

In addition, it follows from Proposition 2.4 that the segment process $(X_t)_{t \ge 0}$ of Eq. (1.1) is a strong Markov process on $(\mathcal{C}_r, \mathcal{B}(\mathcal{C}_r))$ with transition functions $P_t(\xi, \cdot) := \mathbb{P}(X_t(\xi) \in \cdot)$. The associated Markovian semigroup operators are given by

$$P_t f(\xi) = \mathbb{E} f(X_t(\xi)) = \int_{\mathcal{C}_r} f(\eta) P_t(\xi, d\eta), \qquad t \ge 0, \ f \in \mathcal{B}_b(\mathcal{C}_r), \ \xi \in \mathcal{C}_r.$$

and

$$(P_t \mu)(A) = \int_{\mathcal{C}_r} P_t(\xi, A) \mu(d\xi), \quad \mu \in \mathcal{P}(\mathcal{C}_r), \ A \in \mathcal{B}(\mathcal{C}_r).$$

Theorem 3.1. *Let Assumptions 3.1 and 3.2 hold. Assume also that there exist a continuous functional $V : \mathcal{C}_r \to [0, \infty)$ with $\lim_{\|\xi\|_r \to \infty} V(\xi) = \infty$ and constants $C_V, \theta > 0$ such that*

$$P_t V(\xi) := \int_{\mathcal{C}_r} V(\eta) P_t(\xi, d\eta) \le C_V e^{-\theta t} V(\xi) + C_V \tag{3.2}$$

holds for all $\xi \in \mathcal{C}_r$ *and* $t \geq 0$*. Then* P_t *has a unique invariant probability measure* π*, and for any* $\gamma \in (0, \beta)$*, there exist constants* $C, \rho > 0$ *such that*

$$\mathbb{W}_{d_\gamma^V}(P_t(\xi, \cdot), \pi) \leq Ce^{-\rho t}\sqrt{1 + V(\xi)}, \quad t \geq 0, \tag{3.3}$$

where $d_\gamma^V(\xi, \eta) := \sqrt{d_\gamma(\xi, \eta)(1 + V(\xi) + V(\eta))}$*.*

We first present two crucial lemmas before proving the above theorem.

Lemma 3.2. *Let Assumptions 3.1 and 3.2 hold. Then for any* $h > 0$ *and* $\gamma \in (0, \beta)$*, there exists an* $N(h, \gamma) > 0$ *such that for any* $N \geq N(h, \gamma)$ *and all* $\xi, \eta \in \mathcal{C}_r$ *with* $d_{N,\gamma}(\xi, \eta) < 1$*,*

$$\mathbb{W}_{d_{N,\gamma}}\big(P_h(\xi, \cdot), P_h(\eta, \cdot)\big) \leq \theta_1 d_{N,\gamma}(\xi, \eta) \tag{3.4}$$

holds for some $\theta_1 \in (0, 1)$*. Moreover, there exists a constant* $\theta_h > 0$ *such that for all* $\xi, \eta \in \mathcal{C}_r$

$$\mathbb{W}_{d_{N,\gamma}}\big(P_t(\xi, \cdot), P_t(\eta, \cdot)\big) \leq \theta_h d_{N,\gamma}(\xi, \eta), \quad \forall t \in [0, h]. \tag{3.5}$$

Proof. We adopt the idea in [20] to prove this lemma. The proof is divided into four steps. In Step 1, we construct an asymptotic coupling and then give the deviation bound between the asymptotic coupling processes X, Y in Step 2. An application of the triangle inequality gives the desired result (3.4) in Step 3. Finally the estimation (3.5) is established in Step 4.

Fix $\xi, \eta \in \mathcal{C}_r$ arbitrarily and denote $\upsilon = \|\xi - \eta\|_r$. Assume without loss of generality that $\upsilon > 0$. Consider the following equation

$$dY(t) = b(Y_t)dt + \sigma(Y_t)dW(t) + \upsilon^{\gamma-1}(X(t) - Y(t))1_{\{t\leq\tau\}}dt, \quad Y_0 = \eta, \tag{3.6}$$

where $X(t)$ denotes the solution to (1.1) with $X_0 = \xi$ and $\tau = \inf\{t \geq 0 : |X(t) - Y(t)| \geq 2\upsilon\}$. Clearly, Theorem 2.1 implies that under the assumptions of Lemma 3.2 the system of coupling equations involving (1.1) and (3.6) has a unique strong solution (X, Y).

Step 1. Fix some arbitrary $h > 0$. We first prove that there exist positive constants $\theta_2 \in (0, 1)$, κ_1, κ_2, and $\hat{\upsilon}$ small enough such that for $\|\xi - \eta\|_r < \hat{\upsilon}$,

$$\mathbb{P}\big\{\|X_h - Y_h\|_r \geq \theta_2\|\xi - \eta\|_r\big\} \leq C_1(h)\|\xi - \eta\|_r^{\kappa_1} \exp\big\{-C_2(h)\|\xi - \eta\|_r^{-\kappa_2}\big\}, \tag{3.7}$$

where $C_1(h)$ and $C_2(h)$ denote some constants depending on h. Applying Itô's formula yields that

$$|X(t) - Y(t)|^2 = |\xi(0) - \eta(0)|^2 + \int_0^t A(s)ds + \int_0^t \Sigma(s)dW(s),$$

where

$$A(s) = 2\langle X(s) - Y(s), b(X_s) - b(Y_s)\rangle + \|\sigma(X_s) - \sigma(Y_s)\|^2 - 2\upsilon^{\gamma-1}|X(s) - Y(s)|^2 1_{\{s\leq\tau\}},$$

and

$$\Sigma(s) = 2(X(s) - Y(s))^\top(\sigma(X_s) - \sigma(Y_s)).$$

Without loss of generality, we suppose $\delta \leq 1$ and $2\upsilon \leq \delta$. Observing that $\|X_s - Y_s\|_r \leq 2\upsilon$ for $s \leq \tau$, by Assumption 3.1, we have

$$A(s) \leq -2\upsilon^{\gamma-1}|X(s) - Y(s)|^2 + 4K\upsilon^{1+\beta} \text{ and } |\Sigma(s)| \leq 8\sqrt{K}\upsilon^{\frac{3+\beta}{2}} \quad \text{for all } s \leq \tau. \tag{3.8}$$

Let $\lambda = 2\upsilon^{\gamma-1}$, $A = 4K\upsilon^{1+\beta}$ and $B = 64K\upsilon^{3+\beta}$. Then for any $\delta_0 \in (0, 1/2)$, we have

$$A\lambda^{-1} = 2K\upsilon^{1+\beta}\upsilon^{1-\gamma} = 2K\upsilon^{2+\beta-\gamma}, \quad B^{1/2}\lambda^{-\delta_0} = 2^{3-\delta_0}\sqrt{K}\upsilon^{2+\frac{\beta-1}{2}+\delta_0(1-\gamma)}.$$

Since $\gamma < \beta$, we have $\frac{\beta-1}{2} + \frac{1}{2}(1-\gamma) = \frac{\beta-\gamma}{2} > 0$. Recall $0 < \upsilon < 1$. We can fix $\delta_0 < 1/2$ sufficiently close to 1/2 and then choose $\kappa > 0$ small enough such that

$$A\lambda^{-1} \le 2K\upsilon^{2+\kappa}, \quad B^{1/2}\lambda^{-\delta_0} \le 8\sqrt{K}\upsilon^{2+2\kappa}.$$

Then there exists a $\upsilon_0 \in (0, 1)$ such that for any $\upsilon \in (0, \upsilon_0)$,

$$\upsilon^{-\kappa} \ge (1 + 8\log 2)^{1/2}(\Gamma(\delta_0) + \sup_{x>0} x^{\delta_0}e^{-x})\big(h^{1-2\delta_0} + \upsilon^{\kappa}(\Gamma(\delta_0) + \sup_{x>0} x^{\delta_0}e^{-x})\big),$$

where $\Gamma(\cdot)$ is the Gamma function. Then applying Lemma A.2 in the Appendix gives that for all $\upsilon \in (0, \upsilon_0)$,

$$\mathbb{P}(H) \le c_1\upsilon^{2\kappa\delta_0}\left(h^{\frac{5-2\delta_0}{2}} + 1\right)\exp\left\{-c_2\upsilon^{-2\kappa}(h^{\frac{1-2\delta_0}{2}} + 1)^{-2}\right\}, \tag{3.9}$$

where $c_1, c_2 > 0$ depend only on δ_0, and

$$H := \left\{\sup_{0\le t\le \tau\wedge h}\left\{|X(t) - Y(t)|^2 - e^{-2\upsilon^{\gamma-1}t}|\xi(0) - \eta(0)|^2\right\} \ge (2K + 8\sqrt{K})\upsilon^{2+\kappa}\right\}.$$

Recall that $\kappa > 0$ and the definition of the stopping time τ. Then on the set $\{\tau < \infty\}$, the inequality

$$|X(\tau) - Y(\tau)|^2 - e^{-2\upsilon^{\gamma-1}\tau}|\xi(0) - \eta(0)|^2 \ge (2K + 8\sqrt{K})\upsilon^{2+\kappa}$$

implies

$$4\upsilon^2 - e^{-2\upsilon^{\gamma-1}\tau}|\xi(0) - \eta(0)|^2 \ge (2K + 8\sqrt{K})\upsilon^{2+\kappa}.$$

Furthermore, we can choose a constant υ_1 small enough such that for any $\upsilon \in (0, \upsilon_1)$

$$4\upsilon^2 - e^{-2\upsilon^{\gamma-1}\tau}|\xi(0) - \eta(0)|^2 \ge 4\upsilon^2 - e^{-2\upsilon^{\gamma-1}\tau}\upsilon^2 \ge 3\upsilon^2 \ge (2K + 8\sqrt{K})\upsilon^{2+\kappa}. \tag{3.10}$$

This implies that

$$\{\tau < h\} \subset H, \quad \text{for all } \upsilon \in (0, \upsilon_1). \tag{3.11}$$

Therefore, it follows from (3.9) and (3.11) that for $\upsilon \in (0, \upsilon_0 \wedge \upsilon_1)$,

$$\begin{aligned}
&\mathbb{P}\left\{\sup_{0\le t\le h}\left\{|X(t) - Y(t)|^2 - e^{-2\upsilon^{\gamma-1}t}|\xi(0) - \eta(0)|^2\right\} \ge (2K + 8\sqrt{K})\upsilon^{2+\kappa}\right\}\\
&\quad= \mathbb{P}\left\{\sup_{0\le t\le h}\left\{|X(t) - Y(t)|^2 - e^{-2\upsilon^{\gamma-1}t}|\xi(0) - \eta(0)|^2\right\} \ge (2K + 8\sqrt{K})\upsilon^{2+\kappa}, \tau < h\right\}\\
&\qquad+\mathbb{P}\left\{\sup_{0\le t\le h}\left\{|X(t) - Y(t)|^2 - e^{-2\upsilon^{\gamma-1}t}|\xi(0) - \eta(0)|^2\right\} \ge (2K + 8\sqrt{K})\upsilon^{2+\kappa}, \tau \ge h\right\}\\
&\quad\le \mathbb{P}\big\{\tau < h\big\} + \mathbb{P}(H)\\
&\quad\le 2\mathbb{P}(H)\\
&\quad\le 2c_1\upsilon^{2\kappa\delta_0}\left(h^{\frac{5-2\delta_0}{2}} + 1\right)\exp\left\{-c_2\upsilon^{-2\kappa}(h^{\frac{1-2\delta_0}{2}} + 1)^{-2}\right\}.
\end{aligned} \tag{3.12}$$

On the other hand, the inequality

$$\sup_{0\le t\le h}\left\{|X(t) - Y(t)|^2 - e^{-2\upsilon^{\gamma-1}t}|\xi(0) - \eta(0)|^2\right\} < (2K + 8\sqrt{K})\upsilon^{2+\kappa} \tag{3.13}$$

implies that

$$\sup_{0\le t\le h} e^{2rt}|X(t)-Y(t)|^2 \le \sup_{0\le t\le h} e^{2(r-\upsilon^{\gamma-1})t}\|\xi-\eta\|_r^2 + e^{2rh}(2K+8\sqrt{K})\upsilon^{2+\kappa}.$$

These, together with the definition of norm $\|\cdot\|_r$, yield that

$$\begin{aligned}\|X_h-Y_h\|_r^2 &\le e^{-2rh}\|\xi-\eta\|_r^2 \vee e^{-2rh}\sup_{0\le t\le h} e^{2rt}|X(t)-Y(t)|^2\\ &\le e^{-2rh}\|\xi-\eta\|_r^2 \vee \left[e^{-2rh}\sup_{0\le t\le h} e^{2(r-\upsilon^{\gamma-1})t}\|\xi-\eta\|_r^2 + (2K+8\sqrt{K})\upsilon^{2+\kappa}\right]\\ &= \left[e^{-2rh} \vee \left(e^{-2rh}\sup_{0\le t\le h} e^{2(r-\upsilon^{\gamma-1})t} + (2K+8\sqrt{K})\upsilon^{\kappa}\right)\right]\|\xi-\eta\|_r^2.\end{aligned}$$

Noting that $h>0$, $\gamma\in(0,1)$, and $\kappa>0$, there exist constants $\upsilon_2>0$ and $\theta_2<1$ such that for all $\upsilon\in(0,\upsilon_2)$,

$$r<\upsilon^{\gamma-1} \quad\text{and}\quad e^{-2rh}+(2K+8\sqrt{K})\upsilon^{\kappa}<\theta_2^2<1. \tag{3.14}$$

Therefore, inequality (3.13) implies

$$\|X_h-Y_h\|_r<\theta_2\|\xi-\eta\|_r, \quad \forall\, \upsilon\in(0,\upsilon_2).$$

That is, for any $\upsilon\in(0,\upsilon_2)$,

$$\begin{aligned}&\left\{\sup_{0\le t\le h}\left\{|X(t)-Y(t)|^2-e^{-2\upsilon^{\gamma-1}t}|\xi(0)-\eta(0)|^2\right\}<(2K+8\sqrt{K})\upsilon^{2+\kappa}\right\}\\ &\quad\subset\left\{\|X_h-Y_h\|_r<\theta_2\|\xi-\eta\|_r\right\}.\end{aligned}$$

Take $\upsilon_3=\min\{\upsilon_0,\upsilon_1,\upsilon_2\}$. Then for any $\upsilon\in(0,\upsilon_3)$, we have

$$\begin{aligned}&\mathbb{P}\{\|X_h-Y_h\|_r\ge\theta_2\|\xi-\eta\|_r\}\\ &\quad\le\mathbb{P}\left\{\sup_{0\le t\le h}\left\{|X(t)-Y(t)|^2-e^{-2\upsilon^{\gamma-1}t}|\xi(0)-\eta(0)|^2\right\}\ge(2K+8\sqrt{K})\upsilon^{2+\kappa}\right\}\\ &\quad\le 2c_1\upsilon^{2\kappa\delta_0}\left(h^{\frac{5-2\delta_0}{2}}+1\right)\exp\left\{-c_2\upsilon^{-2\kappa}(h^{\frac{1-2\delta_0}{2}}+1)^{-2}\right\},\\ &\quad= 2c_1\|\xi-\eta\|_r^{2\kappa\delta_0}\left(h^{\frac{5-2\delta_0}{2}}+1\right)\exp\left\{-c_2\|\xi-\eta\|_r^{-2\kappa}(h^{\frac{1-2\delta_0}{2}}+1)^{-2}\right\},\end{aligned} \tag{3.15}$$

which yields (3.7) as desired.

Step 2. Now we estimate the change of the law of the segment process caused by the additional drift term, that is, we show that

$$\|\mathrm{Law}(Y_h)-P_h(\eta,\cdot)\|_{TV}\le C_3(h)\|\xi-\eta\|_r^{\gamma},$$

where $C_3(h)$ is a constant depending on h. Let

$$J(t)=\sigma^{-1}(Y_t)\upsilon^{\gamma-1}(X(t)-Y(t))1_{\{t\le\tau\}}, \qquad \widetilde{W}(t)=W(t)+\int_0^t J(s)ds,$$

and

$$R(t)=\exp\left\{-\int_0^t J(s)dW(s)-\frac{1}{2}\int_0^t|J(s)|^2ds\right\}.$$

By Assumption 3.2, we obtain

$$|J(t)|\le 2\|\sigma^{-1}\|_\infty\upsilon^{\gamma}=2\|\sigma^{-1}\|_\infty\|\xi-\eta\|_r^{\gamma}, \quad \forall t\ge 0. \tag{3.16}$$

Then the Girsanov theorem reveals that $\{\widetilde{W}(t) : 0 \le t \le h\}$ is a Wiener process under the probability measure $\mathbb{Q}^{(h)}$ with $d\mathbb{Q}^{(h)} = R(h)d\mathbb{P}$. Furthermore, we can rewrite (3.6) as:

$$dY(t) = b(Y_t)dt + \sigma(Y_t)d\widetilde{W}(t), \quad Y_0 = \eta.$$

Therefore, on the new probability space $(\Omega, \mathcal{F}, \mathbb{Q}^{(h)})$, Y solves (1.1) with W replaced by $\widetilde{W}$ up to the time h. The weak uniqueness of (1.1) implies the law of Y_h under $\mathbb{Q}^{(h)}$ equals $P_h(\eta, \cdot)$. Combining with (3.16), this means that

$$\begin{aligned}\|\mathrm{Law}(Y_h) - P_h(\eta,\cdot)\|_{\mathrm{TV}} &\le \|\mathbb{P} - \mathbb{Q}^{(h)}\|_{\mathrm{TV}} \\ &\le \sqrt{\frac{1}{2}\int_\Omega \log\left(\frac{d\mathbb{P}}{d\mathbb{Q}^{(h)}}\right) d\mathbb{P}} \\ &\le \sqrt{-\frac{1}{2}\mathbb{E}\log R(h)} \\ &\le \sqrt{h}\|\sigma^{-1}\|_\infty \|\xi - \eta\|_r^\gamma, \end{aligned} \tag{3.17}$$

where we used the Pinsker inequality (see, e.g., [7, Lemma A.1. (A.1)]) in the second inequality.

Step 3. We now estimate the bound of $\mathbb{W}_{d_{N,\gamma}}(P_h(\xi,\cdot), P_h(\eta,\cdot))$. Noting that $d_{N,\gamma} \le 1$, by using the triangle inequality we have

$$\begin{aligned}\mathbb{W}_{d_{N,\gamma}}(P_h(\xi,\cdot), P_h(\eta,\cdot)) &\le W_{d_{N,\gamma}}(P_h(\xi,\cdot), Law(Y_h(\eta))) + W_{d_{N,\gamma}}(Law(Y_h(\eta)), P_h(\eta,\cdot)) \\ &\le \mathbb{E}d_{N,\gamma}(X_h(\xi), Y_h(\eta)) + \|Law(Y_h(\eta)) - P_h(\eta,\cdot)\|_{TV} \\ &\le \mathbb{E}d_{N,\gamma}(X_h(\xi), Y_h(\eta))1_{\{\|X_h(\xi)-Y_h(\eta)\|_r < \theta_2\|\xi-\eta\|_r\}} \\ &\quad + \mathbb{P}\{\|X_h(\xi) - Y_h(\eta)\|_r \ge \theta_2\|\xi-\eta\|_r\} \\ &\quad + \|Law(Y_h(\eta)) - P_h(\eta,\cdot)\|_{TV}. \end{aligned} \tag{3.18}$$

When $d_{N,\gamma}(\xi,\eta) < 1$, we have

$$\mathbb{E}d_{N,\gamma}(X_h(\xi), Y_h(\eta))1_{\{\|X_h(\xi)-Y_h(\eta)\|_r < \theta_2\|\xi-\eta\|_r\}} < N\theta_2^\gamma\|\xi-\eta\|_r^\gamma = \theta_2^\gamma d_{N,\gamma}(\xi,\eta). \tag{3.19}$$

Therefore, substituting (3.15), (3.17), and (3.19) into (3.18) yields

$$\begin{aligned}&\mathbb{W}_{d_{N,\gamma}}\big(P_h(\xi,\cdot), P_h(\eta,\cdot)\big) \\ &\le \theta_2^\gamma d_{N,\gamma}(\xi,\eta) + 2c_1\|\xi-\eta\|_r^{2\kappa\delta_0}\left(h^{\frac{5-2\delta_0}{2}}+1\right)\exp\left\{-c_2\|\xi-\eta\|_r^{-2\kappa}(h^{\frac{1-2\delta_0}{2}}+1)^{-2}\right\} \\ &\quad + \sqrt{h}\|\sigma^{-1}\|_\infty\|\xi-\eta\|_r^\gamma \\ &\le \left(\theta_2^\gamma + \frac{2c_1\left(h^{\frac{5-2\delta_0}{2}}+1\right)\exp\left\{-c_2\|\xi-\eta\|_r^{-2\kappa}(h^{\frac{1-2\delta_0}{2}}+1)^{-2}\right\}}{N\|\xi-\eta\|_r^{\gamma-2\kappa\delta_0}}\right) d_{N,\gamma}(\xi,\eta) \\ &\quad + \frac{\sqrt{h}\|\sigma^{-1}\|_\infty}{N} d_{N,\gamma}(\xi,\eta) \end{aligned} \tag{3.20}$$

provided that $d_{N,\gamma}(\xi,\eta) < 1$. Since $\lim_{\upsilon\to 0} 2c_1\upsilon^{-\gamma+2\kappa\delta_0}e^{-c_2\upsilon^{-2\kappa}(h^{\frac{1-2\delta_0}{2}}+1)^{-2}} = 0$, we have

$$\widehat{C} := \sup_{\upsilon\in(0,1)} 2c_1\upsilon^{-\gamma+2\kappa\delta_0}e^{-c_2\upsilon^{-2\kappa}(h^{\frac{1-2\delta_0}{2}}+1)^{-2}} < \infty.$$

Since $\theta_2 \in (0, 1)$ and $\gamma > 0$, we can further choose an N_1 such that

$$\theta_2^\gamma + \frac{\widehat{C}\left(h^{\frac{5-2\delta_0}{2}} + 1\right) + \sqrt{h}\|\sigma^{-1}\|_\infty}{N_1} < 1$$

is satisfied. Taking $N_2 = \upsilon_3^{-\gamma}$, then

$$d_{N_2,\gamma}(\xi, \eta) < 1 \Leftrightarrow N_2\|\xi - \eta\|_r^\gamma < 1 \Rightarrow \|\xi - \eta\|_r < \upsilon_3.$$

Then it follows from (3.20) that the desired assertion holds for $N(h, \gamma) := \max\{N_1, N_2\}$ and

$$\theta_1 := \theta_2^\gamma + \frac{\widehat{C}\left(h^{\frac{5-2\delta_0}{2}} + 1\right) + \sqrt{h}\|\sigma^{-1}\|_\infty}{N(h, \gamma)} < 1. \tag{3.21}$$

Thus, proved (3.4).

Step 4. We now prove (3.5). Noting that $W_{d_{N,\gamma}}(\cdot, \cdot) \le 1$ which follows from the fact $d_{N,\gamma}(\cdot, \cdot) \le 1$, it suffices to prove (3.5) for $\xi, \eta \in \mathcal{C}_r$ with $d_{N,\gamma}(\xi, \eta) < 1$. Indeed, when $d_{N,\gamma}(\xi, \eta) = 1$, $W_{d_{N,\gamma}}(\cdot, \cdot) \le 1 = d_{N,\gamma}(\xi, \eta)$; that is, (3.5) holds for $\theta_h = 1$. For any $t \in [0, h]$, using inequality (3.12) and the subsequent computations (with h replaced by t in appropriate places), we can find a $\hat{\theta}_2 > 0$ such that for all $\xi, \eta \in \mathcal{C}_r$ with $d_{N,\gamma}(\xi, \eta) < 1$,

$$\begin{aligned}\sup_{0\le t\le h} &\mathbb{P}\left\{\|X_t - Y_t\|_r \ge \hat{\theta}_2\|\xi - \eta\|_r\right\} \\ &\le 2c_1\|\xi - \eta\|_r^{2\kappa\delta_0}\left(h^{\frac{5-2\delta_0}{2}} + 1\right)\exp\left\{-c_2\|\xi - \eta\|_r^{-2\kappa}(h^{\frac{1-2\delta_0}{2}} + 1)^{-2}\right\},\end{aligned}$$

and

$$\sup_{0\le t\le h} \|\mathrm{Law}(Y_t) - P_t(\eta, \cdot)\|_{\mathrm{TV}} \le \sqrt{h}\|\sigma^{-1}\|_\infty\|\xi - \eta\|_r^\gamma.$$

Then (3.5) follows from similar calculations as those in (3.18)–(3.20). This proof is completed. □

Remark 3.3. For any given h and γ, by (3.10) and (3.14), for the above δ_0 and κ, there exists a constant $\widehat{N}(h, \gamma)$ large enough such that we can find a $\theta_2 < 1$ and for any $\upsilon := \|\xi - \eta\|_r < \widehat{N}(h, \gamma)^{-1/\gamma}$,

$$e^{-2rh} + (2K + 8\sqrt{K})\upsilon^\kappa < \theta_2^2 < 1,$$

$$\upsilon^{-\kappa} \ge (1 + 8\log 2)^{1/2}(\Gamma(\delta_0) + \sup_{x>0} x^{\delta_0}e^{-x})\big(h^{1-2\delta_0} + \upsilon^\kappa(\Gamma(\delta_0) + \sup_{x>0} x^{\delta_0}e^{-x})\big),$$

and

$$\theta_2^\gamma + \frac{\widehat{C}\left(h^{\frac{5-2\delta_0}{2}} + 1\right) + \sqrt{h}\|\sigma^{-1}\|_\infty}{\widehat{N}(h, \gamma)} < 1.$$

Since for any given $N \ge \widehat{N}(h, \gamma)$, $d_{N,\gamma}(\xi, \eta) < 1$ implies $\upsilon := \|\xi - \eta\|_r < \widehat{N}(h, \gamma)^{-1/\gamma}$ and

$$\theta_1 := \theta_2^\gamma + \frac{\widehat{C}\left(h^{\frac{5-2\delta_0}{2}} + 1\right) + \sqrt{h}\|\sigma^{-1}\|_\infty}{\widehat{N}(h, \gamma)} < 1,$$

Lemma 3.2 holds for the above $\widehat{N}(h, \gamma)$ and θ_1.

In addition, suppose there exist a constant $\widetilde{N}$ and a function f such that

$$f(\widetilde{N}) \geq \frac{2}{r}, \qquad \frac{1}{\widetilde{N}}\widehat{C}\left(f(\widetilde{N})^{\frac{5-2\delta_0}{2}}+1\right)+\frac{1}{\widetilde{N}}\sqrt{f(\widetilde{N})}\|\sigma^{-1}\|_\infty < 1-\left(\frac{1}{2}\right)^{\gamma},$$

and for $\upsilon < \widetilde{N}^{-1/\gamma}$, $(2K+8\sqrt{K})\upsilon^\kappa < 1/8$ and

$$\upsilon^{-\kappa} \geq (1+8\log 2)^{1/2}\left(\Gamma(\delta_0)+\sup_{x>0} x^{\delta_0}e^{-x}\right)\left[f(\widetilde{N})^{1-2\delta_0}+\upsilon^\kappa\left(\Gamma(\delta_0)+\sup_{x>0} x^{\delta_0}e^{-x}\right)\right].$$

Then in light of (3.14), we can choose $\theta_2 < 1/2$ so that

$$\widetilde{\theta}_1 := \frac{1}{\widetilde{N}}\widehat{C}\left(f(\widetilde{N})^{\frac{5-2\delta_0}{2}}+1\right)+\frac{1}{\widetilde{N}}\sqrt{f(\widetilde{N})}\|\sigma^{-1}\|_\infty+\theta_2^\gamma < 1.$$

Then it follows from the proof of Lemma 3.2, for any $\xi, \eta \in \mathcal{C}_r$ satisfying $d_{\widetilde{N},\gamma}(\xi,\eta) < 1$,

$$\mathbb{W}_{d_{\widetilde{N},\gamma}}\left(P_{\widetilde{h}}(\xi,\cdot), P_{\widetilde{h}}(\eta,\cdot)\right) \leq \widetilde{\theta}_1 d_{\widetilde{N},\gamma}(\xi,\eta),$$

where $\widetilde{h} = f(\widetilde{N})$.

Lemma 3.4. *Under the conditions of Lemma 3.2, for any R, $\varepsilon > 0$, there exists a constant $t_{R,\varepsilon} > 0$ satisfying $e^{-rt_{R,\varepsilon}}R < \varepsilon$ such that for any $t \geq t_{R,\varepsilon}$*

$$\inf_{\xi\in\mathbf{B}_R} \mathbb{P}\left\{\|X_t(\xi)\|_r \leq \varepsilon\right\} > 0,$$

where $\mathbf{B}_R := \{\xi \in \mathcal{C}_r : \|\xi\|_r \leq R\}$.

Proof. Fix $R, \varepsilon > 0$ arbitrarily. Let $\|\xi\|_r < R$. Consider the following equation

$$dY(t) = b(Y_t)dt + \sigma(Y_t)dW(t) - \lambda_0 Y(t)dt, \qquad Y_0 = \xi, \tag{3.22}$$

where λ_0 is some positive constant to be determined later. Then for any $t, \tilde{\kappa} > 0$ and $\delta_0 \in (0, 1/2)$, by [20, Lemma B.1], we obtain

$$\mathbb{P}\left\{\sup_{0\leq s\leq t\wedge\bar{\tau}_n}\left\{|Y(s)|^2 - e^{-2\lambda_0 s}|\xi(0)|^2\right\} \geq \frac{L(1+n^2)}{2\lambda_0}+\frac{2\sqrt{L}(1+n^2)\tilde{\kappa}}{\lambda_0^{\delta_0}}\right\} \leq c_3 e^{-c_4\tilde{\kappa}^2}, \tag{3.23}$$

where $\bar{\tau}_n := \inf\{s \geq 0 : \|Y_s\|_r > n\}$, $n > \|\xi\|_r$, and the constants c_3 and $c_4 > 0$ depend only on t and δ_0. Note that the inequality

$$\sup_{0\leq s\leq t}\left\{|Y(s)|^2 - e^{-2\lambda_0 s}|\xi(0)|^2\right\} < \frac{L(1+n^2)}{2\lambda_0}+\frac{2\sqrt{L}(1+n^2)\tilde{\kappa}}{\lambda_0^{\delta_0}}$$

implies

$$e^{-2rt}\sup_{0\leq s\leq t} e^{2rs}|Y(s)|^2 \leq e^{-2rt}\sup_{0\leq s\leq t} e^{2(r-\lambda_0)s}R^2+\frac{L(1+n^2)}{2\lambda_0}+\frac{2\sqrt{L}(1+n^2)\tilde{\kappa}}{\lambda_0^{\delta_0}},$$

which further means that

$$\|Y_t(\xi)\|_r^2 \leq e^{-2rt}\|\xi\|_r^2 \vee \left(e^{-2rt}\sup_{0\leq s\leq t} e^{2(r-\lambda_0)s}R^2+\frac{L(1+n^2)}{2\lambda_0}+\frac{2\sqrt{L}(1+n^2)\tilde{\kappa}}{\lambda_0^{\delta_0}}\right).$$

Summarizing the above observations, we have

$$\left\{\sup_{0\le s\le t}\left\{|Y(s)|^2-e^{-2\lambda_0 s}|\xi(0)|^2\right\}<\frac{L(1+n_0^2)}{2\lambda_0}+\frac{2\sqrt{L}(1+n_0^2)\tilde{\kappa}}{\lambda_0^{\delta_0}}\right\}$$
$$\subset\left\{\|Y_t(\xi)\|_r^2\le e^{-2rt}\|\xi\|_r^2\right.$$
$$\left.\vee\left(e^{-2rt}\sup_{0\le s\le t}e^{2(r-\lambda_0)s}R^2+\frac{L(1+n^2)}{2\lambda_0}+\frac{2\sqrt{L}(1+n^2)\tilde{\kappa}}{\lambda_0^{\delta_0}}\right)\right\}. \tag{3.24}$$

By Theorem 2.1, Eq. (3.22) has a unique global strong solution. Therefore for any fixed $t\ge t_{R,\varepsilon}$, with $t_{R,\varepsilon}>0$ satisfying $e^{-rt_{R,\varepsilon}}R<\varepsilon$, there exists a constant n_0 large enough such that

$$\mathbb{P}\{\bar{\tau}_{n_0}<t\}\le\frac{1}{4}. \tag{3.25}$$

Since $e^{-2rt_{R,\varepsilon}}R^2<\varepsilon^2$ and the constants c_3 and c_4 depend only on t and δ_0, we can choose $\tilde{\kappa}>0$ such that $c_3e^{-c_4\tilde{\kappa}^2}\le 1/4$. Furthermore, there exists a $\hat{\lambda}_0>r$ such that

$$e^{-2rt_{R,\varepsilon}}\sup_{0\le s\le t_{R,\varepsilon}}e^{2(r-\hat{\lambda}_0)s}R^2+\frac{L(1+n_0^2)}{2\hat{\lambda}_0}+\frac{2\sqrt{L}(1+n_0^2)\tilde{\kappa}}{\hat{\lambda}_0^{\delta_0}}<\varepsilon^2.$$

Note that $e^{-2rt_{R,\varepsilon}}R^2<\varepsilon^2$. Hence, by (3.24) with $\lambda_0=\hat{\lambda}_0$, for the above $t\ge t_{R,\varepsilon}$ fixed, we have

$$\left\{\sup_{0\le s\le t}\left\{|Y(s)|^2-e^{-2\hat{\lambda}_0 s}|\xi(0)|^2\right\}<\frac{L(1+n_0^2)}{2\hat{\lambda}_0}+\frac{2\sqrt{L}(1+n_0^2)\tilde{\kappa}}{\hat{\lambda}_0^{\delta_0}}\right\}$$
$$\subset\left\{\|Y_t(\xi)\|_r^2\le e^{-2rt}\|\xi\|_r^2\vee\left(e^{-2rt}\sup_{0\le s\le t}e^{2(r-\hat{\lambda}_0)s}R^2+\frac{L(1+n^2)}{2\hat{\lambda}_0}+\frac{2\sqrt{L}(1+n^2)\tilde{\kappa}}{\hat{\lambda}_0^{\delta_0}}\right)\right\}$$
$$\subset\left\{\|Y_t(\xi)\|_r\le\varepsilon\right\}.$$

Therefore, we can use (3.23) and (3.25) to obtain

$$\begin{aligned}
&\mathbb{P}\left\{\|Y_t(\xi)\|_r\le\varepsilon\right\}\\
&\ge\mathbb{P}\left\{\sup_{0\le s\le t}\left\{|Y(s)|^2-e^{-2\hat{\lambda}_0 s}|\xi(0)|^2\right\}<\frac{L(1+n_0^2)}{2\hat{\lambda}_0}+\frac{2\sqrt{L}(1+n_0^2)\tilde{\kappa}}{\hat{\lambda}_0^{\delta_0}}\right\}\\
&=1-\mathbb{P}\left\{\sup_{0\le s\le t}\left\{|Y(s)|^2-e^{-2\hat{\lambda}_0 s}|\xi(0)|^2\right\}\ge\frac{L(1+n_0^2)}{2\hat{\lambda}_0}+\frac{2\sqrt{L}(1+n_0^2)\tilde{\kappa}}{\hat{\lambda}_0^{\delta_0}},\bar{\tau}_{n_0}\ge t\right\}\\
&\qquad-\mathbb{P}\left\{\sup_{0\le s\le t}\left\{|Y(s)|^2-e^{-2\hat{\lambda}_0 s}|\xi(0)|^2\right\}\ge\frac{L(1+n_0^2)}{2\hat{\lambda}_0}+\frac{2\sqrt{L}(1+n_0^2)\tilde{\kappa}}{\hat{\lambda}_0^{\delta_0}},\bar{\tau}_{n_0}<t\right\}\\
&\ge1-\mathbb{P}\left\{\sup_{0\le s\le t\wedge\bar{\tau}_{n_0}}\left\{|Y(s)|^2-e^{-2\hat{\lambda}_0 s}|\xi(0)|^2\right\}\ge\frac{L(1+n_0^2)}{2\hat{\lambda}_0}+\frac{2\sqrt{L}(1+n_0^2)\tilde{\kappa}}{\hat{\lambda}_0^{\delta_0}}\right\}\\
&\quad-\mathbb{P}\left\{\bar{\tau}_{n_0}<t\right\}\\
&\ge1-\frac{1}{4}-\frac{1}{4}=\frac{1}{2}.
\end{aligned} \tag{3.26}$$

Recall that b and σ satisfy the one-sided linear growth and linear growth condition, respectively. Therefore a standard argument yields that for the t given above, we have

$$\widehat{C}_1 := \sup_{\xi\in\mathbf{B}_R} \mathbb{E} \sup_{s\in[0,t]} |Y(s)|^2 < \infty. \tag{3.27}$$

Let

$$\bar{R}(s) := \exp\left\{\int_0^s \langle \hat{\lambda}_0\sigma^{-1}(Y_s)Y(s), dW(s)\rangle - \frac{1}{2}\int_0^s \hat{\lambda}_0^2|\sigma^{-1}(Y_s)Y(s)|^2 ds\right\}, \quad s \geq 0.$$

Next we show that

$$\sup_{0\leq s\leq t} \mathbb{E}\bar{R}(s)\log\bar{R}(s) < \infty. \tag{3.28}$$

To this end, we define a sequence of stopping time $\tau_n := \inf\{s \geq 0 : |Y(s)| > n\}$ for $n \in \mathbb{N}$. Then $\bar{R}(\cdot \wedge \tau_n)$ is a nonnegative martingale for each $n \in \mathbb{N}$. Let $\mathbb{Q}_n(A) := \mathbb{E}[\bar{R}(t \wedge \tau_n)1_A]$ for $A \in \mathcal{F}_t$ and $n \in \mathbb{N}$, which is a consistent family of probability measures. In addition, the process

$$\overline{W}^{(n)}(s) := W(s) - \int_0^s \hat{\lambda}_0\sigma^{-1}(Y_v)Y(v)dv, \quad 0 \leq s \leq t \wedge \tau_n$$

is a $\mathbb{Q}_n$ Brownian motion and Y satisfies the SFDEs

$$\begin{cases} dY(s) = b(Y_s)ds + \sigma(Y_s)d\overline{W}^{(n)}(s), & 0 \leq s \leq t \wedge \tau_n, \\ Y_0 = \xi, \end{cases}$$

under $\mathbb{Q}_n$. Using a similar argument as that for (3.27), we can show that $\mathbb{E}_{\mathbb{Q}_n}[\sup_{0\leq s\leq t\wedge\tau_n} |Y(s)|^2] \leq \bar{L} < \infty$, where the positive constant $\bar{L} = \bar{L}(t, \|\xi\|_r)$ is independent of $n \in \mathbb{N}$. Therefore, for any $s \in [0, t]$ and $n \in \mathbb{N}$, we have from Assumption 3.2 that

$$\begin{aligned} \mathbb{E}[\bar{R}(s \wedge \tau_n)\log\bar{R}(s \wedge \tau_n)] = \mathbb{E}_{\mathbb{Q}_n}[\log\bar{R}(s \wedge \tau_n)] &= \frac{1}{2}\mathbb{E}_{\mathbb{Q}_n}\left[\int_0^{s\wedge\tau_n} \hat{\lambda}_0^2|\sigma^{-1}(Y_u)Y(u)|^2 du\right] \\ &\leq \frac{1}{2}\hat{\lambda}_0^2\|\sigma^{-1}\|_\infty \int_0^s \mathbb{E}_{\mathbb{Q}^n}|Y(u \wedge \tau_n)|^2 du \leq C(\xi, t) < \infty. \end{aligned}$$

Therefore, passing to the limit and utilizing Fatou's lemma lead to (3.28), which further implies that $\{\bar{R}(s), s \in [0, t]\}$ is a uniformly integrable martingale and hence $\{\bar{R}(s), s \in [0, t]\}$ is a nonnegative martingale with $\mathbb{E}[\bar{R}(t)] = 1$. Then, by the Girsanov Theorem, $\overline{W}(s) = W(s) - \int_0^s \hat{\lambda}_0\sigma^{-1}(Y_v)Y(v)dv$, $s \in [0, t]$ is a Brownian motion under $\mathbb{Q}$, where $\mathbb{Q}(A) = \mathbb{E}[\bar{R}(t)1_A]$, $A \in \mathcal{F}_t$. Moreover, we can rewrite (3.22) as

$$dY(s) = b(Y_s)ds + \sigma(Y_s)d\overline{W}(s), \quad 0 \leq s \leq t, Y_0 = \xi.$$

In other word, Y solves (1.1) up to time t under $\mathbb{Q}$. In view of the pathwise uniqueness for (1.1), it follows from the Yamada–Watanabe Theorem (see, e.g., [18, Proposition 5.3.20] or [18, Corollary 5.3.23]) that for any $\xi \in \mathcal{C}_r$, there exists a measurable mapping $\Phi_\xi : C([0, t]; \mathbf{R}^d) \to C([0, t]; \mathbf{R}^d)$ such that

$$X|_{[0,t]} = \Phi_\xi(W|_{[0,t]}), \quad Y|_{[0,t]} = \Phi_\xi(\overline{W}|_{[0,t]}),$$

where X and Y denote the solutions to Eqs. (1.1) and (3.22) with initial data ξ, respectively. By [7, Theorem A.2] and (3.27), we have

$$\sup_{\xi\in\mathbf{B}_R} D_{KL}\left(\mathrm{Law}(\overline{W}|_{[0,t]} \parallel \mathrm{Law}(W|_{[0,t]}))\right)$$
$$\leq \frac{\hat\lambda_0^2}{2}\mathbb{E}\int_0^t |\sigma^{-1}(Y_s)Y(s)|^2 ds \leq \frac{\hat\lambda_0^2\widehat{C}_1 t}{2}\|\sigma^{-1}\|_\infty^2 =: \widehat{C}_2. \tag{3.29}$$

Denote

$$D_t = \left\{x \in \mathcal{C}([0,t];\mathbf{R}^d) : \sup_{0\leq s\leq t} e^{r(s-t)}|\Phi_\xi(x)(s)| \leq \varepsilon\right\}.$$

Recall $e^{-2rt_{R,\varepsilon}}R^2 < \varepsilon^2, t \geq t_{R,\varepsilon}$ and the definition of the norm $\|\cdot\|_r$. It follows from (3.26), (3.29), and [7, Lemma A.1] that for any $M > 1$ and $\xi \in \mathbf{B}_R$, we have

$$\begin{aligned}
\mathbb{P}\left\{\|X_t(\xi)\|_r \leq \varepsilon\right\} &= \mathbb{P}\left\{\sup_{0\leq s\leq t} e^{r(s-t)}|X(s)| \leq \varepsilon\right\} \\
&= \mathbb{P}\left\{\sup_{0\leq s\leq t} e^{r(s-t)}|\Phi_\xi(W|_{[0,t]})(s)| \leq \varepsilon\right\} \\
&= \mathrm{Law}(W|_{[0,t]})(D_t) \\
&\geq \frac{\mathrm{Law}(\overline{W}|_{[0,t]})(D_t)}{M} - \frac{D_{KL}\left(\mathrm{Law}(\overline{W}|_{[0,t]} \parallel \mathrm{Law}(W|_{[0,t]}))\right) + \log 2}{M\log M} \\
&\geq \frac{1}{2M} - \frac{\widehat{C}_2 + \log 2}{M\log M}.
\end{aligned}$$

Taking $M = \exp\{4\widehat{C}_2 + 4\log 2\}$, we have

$$\inf_{\xi\in\mathbf{B}_R}\mathbb{P}\left\{\|X_t(\xi)\|_r \leq \varepsilon\right\} \geq \frac{1}{4\exp\{4\widehat{C}_2 + 4\log 2\}} > 0.$$

This is the desired assertion. □

Remark 3.5. It is worth noting that Lemma 3.4 holds only for $t \geq t_{R,\varepsilon}$ depending on the initial data and ε. This is the essential difference between SFDEs with infinite delay and finite delay. Consequently we have to obtain the explicit dependence between N and h in Lemma 3.2 (see Remark 3.3) to prove Theorem 3.1.

Now we present the proof of Theorem 3.1.

Proof of Theorem 3.1. Since $\lim_{\|\xi\|_r\to\infty} V(\xi) = \infty$, there exists a constant $R > 0$ such that $\{\xi \in \mathcal{C}_r : V(\xi) \leq 4C_V\} \subset \mathbf{B}_R$. Let

$$h = \frac{\log(R\vee 1) + 2 + \gamma^{-1}\log(2N)}{r}.$$

By Lemma 3.2 and Remark 3.3, there exist constant $N > 0$ large enough and $\theta_1 \in (0,1)$ such that

$$\mathbb{W}_{d_{N,\gamma}}\left(P_h(\xi,\cdot), P_h(\eta,\cdot)\right) \leq \theta_1 d_{N,\gamma}(\xi,\eta) \tag{3.30}$$

provided $d_{N,\gamma}(\xi,\eta) < 1$, that is, $d_{N,\gamma}$ is contractive for P_h. Let $\varepsilon = (2N)^{-1/\gamma}$. Then we have

$$e^{-rh}R \leq e^{-2}\left(\frac{1}{2N}\right)^{\frac{1}{\gamma}} < \varepsilon.$$

Therefore, for the above R, h and ε, it follows from Lemma 3.4 that

$$\varrho := \inf_{\xi\in\mathbf{B}_R}\mathbb{P}\left\{\|X_h(\xi)\|_r \leq \varepsilon\right\} > 0. \tag{3.31}$$

Fix ξ and $\eta \in \mathbf{B}_R$ arbitrarily, and construct independent $\mathcal{C}_r$-valued random variable ς_1, ς_2 such that

$$\mathrm{Law}(\varsigma_1) = P_h(\xi, \cdot) \quad \text{and} \quad \mathrm{Law}(\varsigma_2) = P_h(\eta, \cdot).$$

Then by (3.31), $d_{N,\gamma} \le 1$ and $\gamma \in (0, 1)$, we get

$$\begin{aligned}
&\mathbb{W}_{d_{N,\gamma}}\big(P_h(\xi, \cdot), P_h(\eta, \cdot)\big)\\
&\quad \le \mathbb{E} d_{N,\gamma}(\varsigma_1, \varsigma_2)\\
&\quad = \mathbb{E} d_{N,\gamma}(\varsigma_1, \varsigma_2) 1_{\{\|\varsigma_1\|_r > \varepsilon \text{ or } \|\varsigma_2\|_r > \varepsilon\}} + \mathbb{E} d_{N,\gamma}(\varsigma_1, \varsigma_2) 1_{\{\|\varsigma_1\|_r \le \varepsilon, \|\varsigma_2\|_r \le \varepsilon\}}\\
&\quad \le \mathbb{P}\{\|\varsigma_1\|_r > \varepsilon \text{ or } \|\varsigma_2\|_r > \varepsilon\} + N(2\varepsilon)^\gamma \mathbb{P}\{\|\varsigma_1\|_r \le \varepsilon, \|\varsigma_2\|_r \le \varepsilon\}\\
&\quad \le 1 - \varrho^2 + 2^{\gamma-1}\varrho^2 = 1 - (1 - 2^{\gamma-1})\varrho^2 < 1.
\end{aligned}$$

This implies that $\mathbf{B}_R$ is $d_{N,\gamma}$-small for $\mathbb{P}_h$, which further implies that $\{\xi \in \mathcal{C}_r : V(\xi) \le 4C_V\}$ is $d_{N,\gamma}$-small for $\mathbb{P}_h$.

Since $d_{N,\gamma}$ is a metric and is equivalent to $\|\cdot - \cdot\|_r$, $\mathcal{P}(\mathcal{C}_r)$ is complete under the metric $\mathbb{W}_{d_{N,\gamma}}$ and P_t is also Feller under $d_{N,\gamma}$. Therefore it follows from [15, Theorem 4.8] that P_t has a unique invariant probability measure π. Moreover, there exists a $t_* > 0$ such that

$$\mathbb{W}_{d^V_{N,\gamma}}\big(P_{t_*}\mu, P_{t_*}\nu\big) \le \frac{1}{2}\mathbb{W}_{d^V_{N,\gamma}}\big(\mu, \nu\big) \tag{3.32}$$

for all $\mu, \nu \in \mathcal{P}(\mathcal{C}_r)$. In addition, In light of (3.2) and (3.5), using the Hölder inequality gives that for $t \in [0, t_*]$, there exists a constant $\widetilde{C} > 0$ such that

$$\begin{aligned}
&\mathbb{W}_{d^V_{N,\gamma}}\big(P_t(\xi, \cdot), P_t(\eta, \cdot)\big)\\
&\quad = \inf_{\Pi \in \mathcal{C}(P_t(\xi,\cdot), P_t(\eta,\cdot))} \int_{\mathcal{C}_r \times \mathcal{C}_r} \big(d_{N,\gamma}(x, y)(1 + V(x) + V(y))\big)^{\frac{1}{2}} \Pi(dx, dy)\\
&\quad \le \inf_{\Pi \in \mathcal{C}(P_t(\xi,\cdot), P_t(\eta,\cdot))} \left(\int_{\mathcal{C}_r \times \mathcal{C}_r} d_{N,\gamma}(x, y) \Pi(dx, dy)\right)^{\frac{1}{2}}\\
&\qquad \times \left(\int_{\mathcal{C}_r \times \mathcal{C}_r} (1 + V(x) + V(y)) \Pi(dx, dy)\right)^{\frac{1}{2}}\\
&\quad \le \big(\mathbb{W}_{d_{N,\gamma}}(P_t(\xi, \cdot), P_t(\eta, \cdot))\big)^{\frac{1}{2}} \big(1 + C_V e^{-\theta t}(V(\xi) + V(\eta)) + 2C_V\big)^{\frac{1}{2}}\\
&\quad \le \theta_{t_*}^{\frac{1}{2}} \big(d_{N,\gamma}(\xi, \eta)\big)^{\frac{1}{2}} \big(1 + C_V e^{-\theta t}(V(\xi) + V(\eta)) + 2C_V\big)^{\frac{1}{2}}\\
&\quad \le \widetilde{C} d^V_{N,\gamma}(\xi, \eta),
\end{aligned} \tag{3.33}$$

where $\mathcal{C}(P_t(\xi, \cdot), P_t(\eta, \cdot))$ denotes the family of couplings of $P_t(\xi, \cdot)$ and $P_t(\eta, \cdot)$. Note that the Wasserstein distance $\mathbb{W}_{d^V_{N,\gamma}}(\cdot, \cdot)$ is convex. Then by using the semigroup property of P_t and Jensen's inequality, it follows from (3.32) and (3.33) that for any $t > 0$ and $\xi \in \mathcal{C}_r$,

$$\begin{aligned}
\mathbb{W}_{d^V_{N,\gamma}}\big(P_t(\xi, \cdot), \pi\big) &= \mathbb{W}_{d^V_{N,\gamma}}\big(P_t(\xi, \cdot), P_t\pi\big)\\
&= \mathbb{W}_{d^V_{N,\gamma}}\big(P_{[t/t_*]t_*} P_{t-[t/t_*]t_*}(\xi, \cdot), P_{[t/t_*]t_*} P_{t-[t/t_*]t_*}\pi\big)\\
&\le \frac{1}{2^{[t/t_*]}} \mathbb{W}_{d^V_{N,\gamma}}\big(P_{t-[t/t_*]t_*}(\xi, \cdot), P_{t-[t/t_*]t_*}\pi\big)\\
&\le \frac{1}{2^{[t/t_*]}} \int_{\mathcal{C}_r} \mathbb{W}_{d^V_{N,\gamma}}\big(P_{t-[t/t_*]t_*}(\xi, \cdot), P_{t-[t/t_*]t_*}(\eta, \cdot)\big)\pi(d\eta)
\end{aligned}$$

$$\begin{aligned}
&\leq \widetilde{C}\frac{1}{2^{[t/t_*]}}\int_{\mathcal{C}_r} d^V_{N,\gamma}(\xi,\eta)\pi(d\eta)\\
&\leq \widetilde{C}\frac{1}{2^{[t/t_*]}}\sqrt{1+V(\xi)+\pi(V)}. \qquad (3.34)
\end{aligned}$$

In addition, it is easy to see from (3.2) that $\pi(V)<\infty$. Therefore, (3.34) implies that there exist constants C and $\rho>0$ such that

$$\mathbb{W}_{d^V_{N,\gamma}}\big(P_t(\xi,\cdot),\pi\big)\leq Ce^{-\rho t}\sqrt{1+V(\xi)}, \quad t\geq 0.$$

Since $d^V_\gamma(\xi,\eta)\leq d^V_{N,\gamma}(\xi,\eta)$, (3.3) follows and the proof is completed. □

Remark 3.6. Similar to (3.34), it is easy to observe from (3.32) that for any $\mu,\nu\in\mathcal{P}(\mathcal{C}_r)$ with $(\mu\times\nu)(d^V_{N,\gamma}(\cdot,\cdot))<\infty$, there exists some $c>0$ such that

$$\mathbb{W}_{d^V_{N,\gamma}}(P_t\mu,P_t\nu)\leq ce^{-\rho t}\mathbb{W}_{d^V_{N,\gamma}}(\mu,\nu), \quad t\geq 0.$$

4. Asymptotic log-Harnack inequality

To establish the asymptotic log-Harnack inequality for Eq. (1.1), we need to impose the following stronger conditions on the coefficients b and σ.

Assumption 4.1. b is continuous and bounded on bounded subset of $\mathcal{C}_r$. In addition, there exist a positive constant K_1, a decreasing continuous function $u\in\mathcal{U}$ and a probability measure $\mu\in\mathcal{M}_{2r}$ such that for any $\xi,\eta\in\mathcal{C}_r$,

$$\begin{aligned}
&2\langle\phi(0)-\psi(0),b(\phi)-b(\psi)\rangle_+ + \|\sigma(\phi)-\sigma(\psi)\|^2\\
&\quad\leq K_1\int_{-\infty}^0|\phi(\theta)-\psi(\theta)|^2u(|\phi(\theta)-\psi(\theta)|^2)\mu(d\theta).
\end{aligned}$$

Moreover, the function $s\mapsto su(s^2)$ is increasing and there exist constants $K_2>0$ and $0<\alpha<1$ such that the function $u\in\mathcal{U}$ satisfies the following inequality

$$su^2(s)\leq K_2\big((su(s))^\alpha+su(s)\big), \quad \forall s>0. \qquad (4.1)$$

It is easy to observe that the function $u(s)=\log(\mathrm{e}^2\vee s^{-1})\in\mathcal{U}$ is decreasing and satisfies (4.1), and $su(s^2)$ is increasing.

Assumption 4.2. The functional σ satisfies $\|\sigma\|_\infty:=\sup_{\phi\in\mathcal{C}_r}\|\sigma(\phi)\|<\infty$ and for any $\phi\in\mathcal{C}_r$, $\sigma(\phi)$ admits a right inverse $\sigma^{-1}(\phi)$ and $\|\sigma^{-1}\|_\infty:=\sup_{\phi\in\mathcal{C}_r}\|\sigma(\phi)\|<\infty$.

For convenience, we first present some notation and definitions to be used in this section. Let (E,d) be a Polish space. For a function $f:E\to\mathbf{R}$ and any $x\in E$, denote

$$|Df(x)|=\limsup_{d(x,y)\to 0}\frac{|f(x)-f(y)|}{d(x,y)}.$$

We further denote $\|Df\|_\infty=\sup_{x\in E}|Df(x)|$. An increasing sequence $(d_n)_{n\geq1}$ of bounded, continuous pseudo-metrics on (E,d) is called *totally separating* if for every $x\neq y$, it holds that $\lim_{n\to\infty}d_n(x,y)=1$.

Definition 4.1 (*[14, Definition 3.8]*). A Markov semigroup $(P_t)_{t\geq 0}$ satisfies the *asymptotic strong Feller property* if it is Feller and there exist a sequence of positive real numbers $(t_n)_{n\geq 1}$ and a totally separating sequence $(d_n)_{n\geq 1}$ of pseudo-metrics such that for every $x \in E$,

$$\inf_{U\in\mathcal{U}_x} \limsup_{n\to\infty} \sup_{y\in U} W_{d_n}\big(P_{t_n}(x,\cdot), P_{t_n}(y,\cdot)\big) = 0,$$

where $\mathcal{U}_x := \{U \subseteq E : x \in U \text{ and } U \text{ is an open set}\}$.

For any $\lambda > r$, consider the following SFDEs with infinite delay

$$\begin{cases} dX(t) = b(X_t)dt + \sigma(X_t)dW(t), \\ dY(t) = \{b(Y_t) + \lambda\sigma(Y_t)\sigma^{-1}(X_t)\Gamma(X(t), Y(t))\}dt + \sigma(Y_t)dW(t), \end{cases} \tag{4.2}$$

with the initial data $X_0 = \xi$ and $Y_0 = \eta$, where

$$\Gamma(X(t), Y(t)) = \begin{pmatrix} (X^1(t) - Y^1(t))u(|X^1(t) - Y^1(t)|^2) \\ (X^2(t) - Y^2(t))u(|X^2(t) - Y^2(t)|^2) \\ \vdots \\ (X^d(t) - Y^d(t))u(|X^d(t) - Y^d(t)|^2) \end{pmatrix}. \tag{4.3}$$

Under Assumption 4.1, it is easy to see that b and σ satisfy Assumptions 2.1 and 2.2. Since the first equation of (4.2) does not depend on Y, it has a unique solution. To show the existence and uniqueness of solution to the second equation of (4.2), it suffices to verify that

$$\tilde{b}(\phi) := \lambda\sigma(\phi)\sigma^{-1}(\xi)\Gamma(\xi(0), \phi(0)), \quad \phi \in \mathcal{C}_r$$

as a drift satisfies Assumptions 2.1 and 2.2 for any fixed $\xi \in \mathcal{C}_r$. This, however, follows directly from Assumptions 4.1 and 4.2; see Proposition A.3. Therefore we can apply Theorem 2.1 again to conclude that the second equation of (4.2) has a unique strong solution. Summarizing the above observation yields that (4.2) has a unique strong solution (X, Y). Let

$$h(t) = \lambda\sigma^{-1}(X_t)\Gamma(X(t), Y(t)), \quad B(t) = W(t) + \int_0^t h(s)ds,$$

and define

$$R(t) = \exp\left\{-\int_0^t \langle h(s), dW(s)\rangle - \frac{1}{2}\int_0^t |h(s)|^2 ds\right\}.$$

Further, define the stopping time

$$\tau_n = \inf\{t \geq 0 : \|X_t\|_r \vee \|Y_t\|_r \geq n\}, \quad n \in \mathbb{N}.$$

Recalling that $0u(0) = 0$ and u is continuous, (4.1) implies that for any fixed $n \geq 1$, $\{R(t \wedge \tau_n)\}_{t\geq 0}$ is a martingale. Thus it follows from Girsanov's theorem that for any fixed $T \geq 0$, $\{B(t \wedge \tau_n)\}_{t\in[0,T]}$ is a d-dimensional Wiener process under the probability measure $d\mathbb{Q}_{T,n} := R(T \wedge \tau_n)d\mathbb{P}$. For $t \leq T \wedge \tau_n$, rewrite (4.2) as

$$\begin{cases} dX(t) = \{b(X_t) - \lambda\Gamma(X(t), Y(t))\}dt + \sigma(X_t)dB(t), \quad X_0 = \xi, \\ dY(t) = b(Y_t)dt + \sigma(Y_t)dB(t), \quad Y_0 = \eta. \end{cases} \tag{4.4}$$

Since $X(t)$ and $Y(t)$ depend on the whole history, it is impossible to construct a successful coupling for X_t and Y_t. Thus, we aim to establish the asymptotic result. In order to establish the asymptotic log-Harnack inequality for SFDEs with infinite delay, we have to verify that $\{B(t)\}_{t\in[0,\infty)}$ is a Wiener process on some probability space. We now present the following result.

Lemma 4.1. *Let Assumptions 4.1 and 4.2 hold. Then for all $\lambda \geq r + K_1\mu^{(2r)}/2 + 1/2$, we have*

$$\sup_{t\geq 0} \mathbb{E}\left[R(t)\log(R(t))\right] < \infty. \tag{4.5}$$

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

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = R(t), \quad \forall t \geq 0. \tag{4.6}$$

Moreover, $\{B(t), \mathcal{F}_t; 0 \leq t < \infty\}$ is a d-dimensional Wiener process on $(\Omega, \mathcal{F}_\infty, \mathbb{Q})$.

Proof. Let $Z(t) = X(t) - Y(t)$. By Assumption 4.2 and (4.1), we have

$$\begin{aligned}
\mathbb{E}[R(t\wedge\tau_n)\log(R(t\wedge\tau_n))] &= \mathbb{E}_{\mathbb{Q}_{t,n}} \log(R(t\wedge\tau_n)) \\
&= \mathbb{E}_{\mathbb{Q}_{t,n}}\left(-\int_0^{t\wedge\tau_n}\langle h(s), dB(s)\rangle + \frac{1}{2}\int_0^{t\wedge\tau_n}|h(s)|^2 ds\right) \\
&\leq \frac{d}{2}\lambda^2\|\sigma^{-1}\|_\infty^2 \mathbb{E}_{\mathbb{Q}_{t,n}}\int_0^{t\wedge\tau_n}|Z(s)|^2u^2(|Z(s)|^2)ds \\
&\leq \frac{dK_2}{2}\lambda^2\|\sigma^{-1}\|_\infty^2 \mathbb{E}_{\mathbb{Q}_{t,n}}\int_0^{t\wedge t_n}\left(|Z(s)|^2u(|Z(s)|^2)\right)^\alpha ds \\
&\quad + \frac{dK_2}{2}\lambda^2\|\sigma^{-1}\|_\infty^2 \mathbb{E}_{\mathbb{Q}_{t,n}}\int_0^{t\wedge\tau_n}|Z(s)|^2u(|Z(s)|^2)ds,
\end{aligned} \tag{4.7}$$

where $\mathbb{E}_{\mathbb{Q}_{t,n}}$ denotes the expectation operator with respect to the probability measure $\mathbb{Q}_{t,n}$. For some positive constant $r_0 < r$, applying the Itô formula and using Assumption 4.1 give

$$\begin{aligned}
e^{2r_0(t\wedge\tau_n)}|Z(t\wedge\tau_n)|^2 &= |\xi(0)-\eta(0)|^2 - 2\lambda\int_0^{t\wedge\tau_n} e^{2r_0 s}\sum_{i=1}^d |Z^i(s)|^2 u(|Z^i(s)|^2)ds \\
&\quad + \int_0^{t\wedge\tau_n} e^{2r_0 s}(2r_0|Z(s)|^2 + 2\langle Z(s), b(X_s) - b(Y_s)\rangle \\
&\quad + \|\sigma(X_s)-\sigma(Y_s)\|^2)ds \\
&\quad + 2\int_0^{t\wedge\tau_n} e^{2r_0 s} Z(s)^\top(\sigma(X_s)-\sigma(Y_s))dB(s) \\
&\leq \|\xi-\eta\|_r^2 + 2(r_0-\lambda)\int_0^{t\wedge\tau_n} e^{2r_0 s}\sum_{i=1}^d |Z^i(s)|^2 u(|Z^i(s)|^2)ds \\
&\quad + K_1\int_0^{t\wedge\tau_n}\int_{-\infty}^0 e^{2r_0 s}|Z(s+\theta)|^2u(|Z(s+\theta)|^2)\mu(d\theta)ds \\
&\quad + 2\int_0^{t\wedge\tau_n} e^{2r_0 s} Z(s)^\top(\sigma(X_s)-\sigma(Y_s))dB(s).
\end{aligned} \tag{4.8}$$

By the Tonelli theorem and a substitution technique, we have

$$\begin{aligned}
&\int_0^{t\wedge\tau_n}\int_{-\infty}^0 e^{2r_0 s}|Z(s+\theta)|^2u(|Z(s+\theta)|^2)\mu(d\theta)ds \\
&\quad = \int_0^{t\wedge\tau_n}\int_{-\infty}^{-s} e^{2r_0 s}|Z(s+\theta)|^2u(|Z(s+\theta)|^2)\mu(d\theta)ds
\end{aligned}$$

$$
\begin{aligned}
&+\int_0^{t\wedge\tau_n}\int_{-s}^0 e^{2r_0s}|Z(s+\theta)|^2u(|Z(s+\theta)|^2)\mu(d\theta)ds\\
\le&\int_0^{t\wedge\tau_n}\int_{-\infty}^{-s} e^{2r_0s}e^{-2r(s+\theta)}\big(e^{2r(s+\theta)}|Z(s+\theta)|^2u(e^{2r(s+\theta)}|Z(s+\theta)|^2)\big)\mu(d\theta)ds\\
&+\int_{-\infty}^0\int_0^{t\wedge\tau_n} e^{2r_0s-2r_0\theta}|Z(s)|^2u(|Z(s)|^2)ds\mu(d\theta)\\
\le&\|\xi-\eta\|_r^2u(\|\xi-\eta\|_r^2)\int_0^{t\wedge\tau_n}\int_{-\infty}^{-s} e^{2r_0s}e^{-2r(s+\theta)}\mu(d\theta)ds\\
&+\mu^{(2r)}\int_0^{t\wedge\tau_n} e^{2r_0s}|Z(s)|^2u(|Z(s)|^2)ds\\
\le&\frac{\mu^{(2r)}}{2r-2r_0}\|\xi-\eta\|_r^2u(\|\xi-\eta\|_r^2)+\mu^{(2r)}\int_0^{t\wedge\tau_n} e^{2r_0s}|Z(s)|^2u(|Z(s)|^2)ds\\
\le&\frac{\mu^{(2r)}}{2r-2r_0}\|\xi-\eta\|_r^2u(\|\xi-\eta\|_r^2)+\mu^{(2r)}\int_0^{t\wedge\tau_n} e^{2r_0s}\sum_{i=1}^d|Z^i(s)|^2u(|Z^i(s)|^2)ds,
\end{aligned}
\tag{4.9}
$$

where the first inequality follows from the fact that $asu(s)\le asu(as)$ for $a\in[0,1]$ and the last inequality follows from the fact that $(s+t)u(s+t)\le su(s)+tu(t)$ for $s,t\ge 0$, which follows from the fact that $u(\cdot)$ is decreasing. Substituting (4.9) into (4.8) yields that

$$
\begin{aligned}
e^{2r_0(t\wedge\tau_n)}|Z(t\wedge\tau_n)|^2=&\Big(1+\frac{K_1}{2r-2r_0}\mu^{(2r)}\Big)\|\xi-\eta\|_r^2u(\|\xi-\eta\|_r^2)\\
&+(2r_0+K_1\mu^{(2r)}-2\lambda)\int_0^{t\wedge\tau_n} e^{2r_0s}\sum_i^d|Z^i(s)|^2u(|Z^i(s)|^2)ds\\
&+2\int_0^{t\wedge\tau_n} e^{2r_0s}Z(s)^\top(\sigma(X_s)-\sigma(Y_s))dB(s).
\end{aligned}
\tag{4.10}
$$

Note that $r_0<r$ and $\lambda\ge r+K_1\mu^{(2r)}/2+1/2$. Thus we have

$$
\begin{aligned}
\int_0^{t\wedge\tau_n} e^{2r_0s}|Z(s)|^2u(|Z(s)|^2)ds\le&\int_0^{t\wedge\tau_n} e^{2r_0s}\sum_{i=1}^d|Z^i(s)|^2u(|Z^i(s)|^2)ds\\
\le&\Big(1+\frac{K_1}{2r-2r_0}\mu^{(2r)}\Big)\|\xi-\eta\|_r^2u(\|\xi-\eta\|_r^2)\\
&+2\int_0^{t\wedge\tau_n} e^{2r_0s}Z(s)^\top(\sigma(X_s)-\sigma(Y_s))dB(s).
\end{aligned}
$$

Taking expectation with respect to $\mathbb{E}_{\mathbb{Q}_{t,n}}$ on both sides of the above inequality, we obtain

$$
\mathbb{E}_{\mathbb{Q}_{t,n}}\int_0^{t\wedge\tau_n} e^{2r_0s}|Z(s)|^2u(|Z(s)|^2)ds\le\Big(1+\frac{K_1}{2r-2r_0}\mu^{(2r)}\Big)\|\xi-\eta\|_r^2u(\|\xi-\eta\|_r^2).
\tag{4.11}
$$

By virtue of Hölder's inequality and (4.11), we have

$$
\begin{aligned}
&\mathbb{E}_{\mathbb{Q}_{t,n}}\int_0^{t\wedge\tau_n}\big(|Z(s)|^2u(|Z(s)|^2)\big)^\alpha ds\\
&=\mathbb{E}_{\mathbb{Q}_{t,n}}\int_0^{t\wedge\tau_n} e^{-2r_0\alpha s}\left(e^{2r_0s}|Z(s)|^2u(|Z(s)|^2)\right)^\alpha ds
\end{aligned}
$$

$$
\begin{aligned}
&\leq \left\{\mathbb{E}_{\mathbb{Q}_{t,n}} \int_0^{t\wedge\tau_n} e^{\frac{-2r_0\alpha s}{1-\alpha}} ds\right\}^{1-\alpha} \left\{\mathbb{E}_{\mathbb{Q}_{t,n}} \int_0^{t\wedge\tau_n} e^{2r_0 s}|Z(s)|^2 u(|Z(s)|^2) ds\right\}^{\alpha} \\
&\leq \left(\frac{1-\alpha}{2r_0\alpha}\right)^{1-\alpha} A^{\alpha},
\end{aligned}
\tag{4.12}
$$

where $A := (1 + \frac{K_1}{2r-2r_0}\mu^{(2r)})\|\xi - \eta\|_r^2 u(\|\xi - \eta\|_r^2)$. Substituting (4.11) and (4.12) into (4.7) gives that

$$
\sup_{t\geq 0, n\geq 1} \mathbb{E}\big[R(t\wedge\tau_n)\log(R(t\wedge\tau_n))\big] \leq \frac{dK_2}{2}\lambda^2\|\sigma^{-1}\|_\infty^2 \Big[\Big(\frac{1-\alpha}{2r_0\alpha}\Big)^{1-\alpha} A^{\alpha} + A\Big]. \tag{4.13}
$$

By Fatou's lemma, it follows from (4.13) that

$$
\sup_{t\geq 0} \mathbb{E}\big[R(t)\log(R(t))\big] \leq \frac{dK_2}{2}\lambda^2\|\sigma^{-1}\|_\infty^2 \Big[\Big(\frac{1-\alpha}{2r_0\alpha}\Big)^{1-\alpha} A^{\alpha} + A\Big] < \infty.
$$

This establishes (4.5).

We now prove (4.6). We first show that $\{R(t), \mathcal{F}_t;\ 0 \leq t < \infty\}$ is a uniformly integrable martingale. The uniform integrability follows from (4.5) directly. Therefore we only need to prove that $R(t)$ is a martingale. By the Dominated Convergence Theorem and the martingale property of $R(t\wedge\tau_n)$, for any $t > s$, we have

$$
\mathbb{E}[R(t)|\mathcal{F}_s] = \mathbb{E}[\lim_{n\to\infty} R(t\wedge\tau_n)|\mathcal{F}_s] = \lim_{n\to\infty} \mathbb{E}[R(t\wedge\tau_n)|\mathcal{F}_s] = \lim_{n\to\infty} R(s\wedge\tau_n) = R(s),
$$

which implies $\{R(t), \mathcal{F}_t;\ 0 \leq t < \infty\}$ is a martingale.

Since $\{R(t)\}_{t\geq 0}$ is a uniformly integrable martingale, it follows from the Submartingale Convergence Theorem (see, e.g., [18, Theorem 3.15]) that the limit $R(\infty) := \lim_{t\to\infty} R(t)$ exists for almost all $\omega \in \Omega$ and $R(\infty)$ is an integrable random variable. Moreover, $\{R(t), \mathcal{F}_t : 0 \leq t \leq \infty\}$ is a martingale (see, e.g., [18, Problem 3.20]). Define a probability measure on $\mathcal{F}_\infty$ as follows

$$
\mathbb{Q}(A) = \mathbb{E}[1_A R(\infty)] \quad \text{for } A \in \mathcal{F}_\infty.
$$

Because $\{R(t), \mathcal{F}_t : 0 \leq t \leq \infty\}$ is a martingale, $\mathbb{Q}(A) = \mathbb{E}[1_A R(t)]$ for $A \in \mathcal{F}_t$, $t \geq 0$. Hence (4.6) holds. Additionally, in light of Girsanov's theorem, for each fixed $T > 0$, $\{B(t), \mathcal{F}_t; 0 \leq t \leq T\}$ is a d-dimensional Wiener process on $(\Omega, \mathcal{F}_T, \mathbb{Q}_T)$, where $\mathbb{Q}_T(A) := \mathbb{E}[1_A R(T)] = \mathbb{Q}(A)$, $\forall A \in \mathcal{F}_T$. As a result, $\{B(t), \mathcal{F}_t; 0 \leq t < \infty\}$ is a d-dimensional Wiener process on $(\Omega, \mathcal{F}_\infty, \mathbb{Q})$. This proof is completed. □

Lemma 4.2. *Let Assumptions 4.1 and 4.2 hold. Then for any $r_0 \in (0, r)$, there exists a constant $C_1 > 0$ such that the asymptotic coupling (X_t, Y_t) satisfies*

$$
\mathbb{E}_{\mathbb{Q}}\|X_t - Y_t\|_r^2 \leq C_1\|\xi - \eta\|_r^2 u(\|\xi - \eta\|_r^2) e^{-2r_0 t}. \tag{4.14}
$$

Proof. Since $\{B(t), \mathcal{F}_t; 0 \leq t < \infty\}$ is a d-dimensional Wiener process, inequality (4.11) is still valid for the probability measure $\mathbb{Q}$ in place of $\mathbb{E}_{\mathbb{Q}_{t,n}}$. In addition, noting that the solution $X(t)$ to (1.1) is non-explosive, by (4.11) and Fatou's lemma, we have

$$
\mathbb{E}_{\mathbb{Q}} \int_0^t e^{2r_0 s}|Z(s)|^2 u(|Z(s)|^2) ds \leq \Big(1 + \frac{K_1}{2r - 2r_0}\mu^{(2r)}\Big)\|\xi - \eta\|_r^2 u(\|\xi - \eta\|_r^2). \tag{4.15}
$$

By the Burkholder–Davis–Gundy inequality, and using Assumption 4.1 and the Cauchy–Schwarz inequality, we have

$$\begin{aligned}
&2\mathbb{E}_{\mathbb{Q}} \sup_{0\le s\le t} \int_0^s e^{2r_0 v} Z(v)^\top(\sigma(X_v)-\sigma(Y_v))dB(v) \\
&\le 8\sqrt{2}\mathbb{E}_{\mathbb{Q}}\left(\int_0^t e^{4r_0 s}|Z(s)|^2\|\sigma(X_s)-\sigma(Y_s)\|^2 ds\right)^{\frac{1}{2}} \\
&\le \frac{1}{2}\mathbb{E}_{\mathbb{Q}} \sup_{0\le s\le t} e^{2r_0 s}|Z(s)|^2 + 64K_1\mathbb{E}_{\mathbb{Q}}\int_0^t\int_{-\infty}^0 e^{2r_0 s}|Z(s+\theta)|^2 u(|Z(s+\theta)|^2)\mu(d\theta)ds \\
&\le \frac{1}{2}\mathbb{E}_{\mathbb{Q}} \sup_{0\le s\le t} e^{2r_0 s}|Z(s)|^2 + \frac{64K_1}{2r-2r_0}\mu^{(2r)}\|\xi-\eta\|_r^2 u(\|\xi-\eta\|_r^2) \\
&\quad +64K_1\mu^{(2r)}\mathbb{E}_{\mathbb{Q}}\int_0^t e^{2r_0 s}|Z(s)|^2 u(|Z(s)|^2)ds.
\end{aligned} \tag{4.16}$$

It follows from (4.10), (4.15), and (4.16) that

$$\mathbb{E}_{\mathbb{Q}} \sup_{0\le s\le t} e^{2r_0 s}|Z(s)|^2 \le \widetilde{C}_1\|\xi-\eta\|_r^2 u(\|\xi-\eta\|_r^2),$$

where $\widetilde{C}_1 := 2+128K_1\mu^{(2r)}+(65+64K_1\mu^{(2r)})\frac{K_1\mu^{(2r)}}{r-r_0}$. Recall the definition of the norm $\|\cdot\|_r$. Noting that $0<r_0<r$, we obtain

$$\begin{aligned}
\|Z_t\|_r^2 &= \sup_{\theta\le 0} e^{2r\theta}|Z(t+\theta)|^2 \\
&\le \sup_{s\le t} e^{2r_0(s-t)}|Z(s)|^2 \le e^{-2r_0 t}\|\xi-\eta\|_r^2 + e^{-2r_0 t}\sup_{0\le s\le t} e^{2r_0 s}|Z(s)|^2.
\end{aligned}$$

Thus, we have

$$\mathbb{E}_{\mathbb{Q}}\|Z_t\|_r^2 \le (\widetilde{C}_1+1)\|\xi-\eta\|_r^2 u(\|\xi-\eta\|_r^2)e^{-2r_0 t}.$$

Hence (4.14) holds for $C_1=\widetilde{C}_1+1$. This proof is completed. □

Theorem 4.3. *Let Assumptions 4.1 and 4.2 hold. Then for any $r_0\in(0,r)$, there exists a constant $C_2>0$ such that for $t>0$, the asymptotic log-Harnack inequality*

$$\begin{aligned}
P_t\log f(\eta) &\le \log P_t f(\xi) + C_2(\|\xi-\eta\|_r^2 u(\|\xi-\eta\|_r^2))^\alpha + C_2\|\xi-\eta\|_r^2 u(\|\xi-\eta\|_r^2) \\
&\quad + C_2 e^{-r_0 t}\|D\log f\|_\infty\sqrt{\|\xi-\eta\|_r^2 u(\|\xi-\eta\|_r^2)}
\end{aligned} \tag{4.17}$$

holds for any $\xi,\eta\in\mathcal{C}_r$ and $f\in\mathcal{B}_b^+(\mathcal{C}_r)$ with $f\ge 1$ and $\|D\log f\|_\infty<\infty$. Consequently, the Markov semigroup P_t is asymptotically strong Feller.

Proof. In light of Lemma 4.1 and the weak uniqueness of solution to (1.1), Y_t also has the Markov semigroup P_t under the probability measure $\mathbb{Q}$, i.e., $P_t f(\eta)=\mathbb{E}_{\mathbb{Q}} f(Y_t)$ for any $t\ge 0$ and $f\in\mathcal{B}_b(\mathcal{C}_r)$. Therefore, for any $f\in\mathcal{B}_b^+$ with $f\ge 1$ and $\|D\log f\|_\infty<\infty$, by the definition of $\|D\log f\|_\infty$ and the Young inequality (see e.g., [1, Lemma 2.4]), we obtain

$$\begin{aligned}
P_t\log f(\eta) &= \mathbb{E}_{\mathbb{Q}}\log f(Y_t) = \mathbb{E}_{\mathbb{Q}}\log f(X_t) + \mathbb{E}_{\mathbb{Q}}(\log f(Y_t)-\log f(X_t)) \\
&\le \mathbb{E}R(t)\log f(X_t) + \|D\log f\|_\infty \mathbb{E}_{\mathbb{Q}}\|X_t-Y_t\|_r \\
&\le \mathbb{E}R(t)\log R(t) + \log P_t f(\xi) + \|D\log f\|_\infty\left(\mathbb{E}_{\mathbb{Q}}\|X_t-Y_t\|_r^2\right)^{\frac{1}{2}}.
\end{aligned} \tag{4.18}$$

Recall that

$$\mathbb{E}\big[R(t)\log R(t)\big] \le \frac{dK_2}{2}\lambda^2\|\sigma^{-1}\|_\infty^2\left(\left(\frac{1-\alpha}{2r_0\alpha}\right)^{1-\alpha}A^\alpha + A\right), \tag{4.19}$$

where $A := (1+\frac{K_1}{2r-2r_0}\mu^{(2r)})\|\xi-\eta\|_r^2 u(\|\xi-\eta\|_r^2)$. Substituting (4.14) and (4.19) into (4.18) yields

$$P_t\log f(\eta) \le \log P_t f(\xi) + \frac{dK_2}{2}\lambda^2\|\sigma^{-1}\|_\infty^2\left(\left(\frac{1-\alpha}{2r_0\alpha}\right)^{1-\alpha}A^\alpha + A\right)$$
$$+ e^{-r_0 t}\|D\log f\|_\infty\sqrt{C_1\|\xi-\eta\|_r^2 u(\|\xi-\eta\|_r^2)}.$$

Therefore (4.17) holds for

$$C_2 = \max\left\{\sqrt{C_1}, \frac{dK_2}{2}\lambda^2\|\sigma^{-1}\|_\infty^2(\frac{1-\alpha}{2r_0})^{1-\alpha}(1+\frac{K_1}{2r-2r_0}\mu^{(2r)}),\right.$$
$$\left.\frac{dK_2}{2}\lambda^2\|\sigma^{-1}\|_\infty^2(1+\frac{K_1}{2r-2r_0}\mu^{(2r)})\right\}.$$

Finally, in view of [34, Theorem 1.4], (4.17) implies that the Markov semigroup P_t is asymptotically strong Feller. This completes this proof. □

Remark 4.4. When the diffusion term depends on the history of the solution, the SFDEs might have a reconstruction property (see, e.g., [26]), which causes the laws of segment processes with different initial data to be mutually singular. This indicates that the strong Feller property and the ergodicity under the total variational distance are invalid. This is the reason why this paper only shows the asymptotic log-Harnack inequality and the exponential ergodicity under Wasserstein distance.

5. Example

In this section we study a concrete example to illustrate the main results of the paper.

Example 5.1. Consider the following 1-dimensional stochastic functional differential equation with infinite delay

$$dX(t) = \left\{-\gamma_1 X(t) + \int_{-\infty}^0 \Phi(X(t+\theta))\mu(d\theta)\right\}dt$$
$$+ \left\{1 + \int_{-\infty}^0 1\wedge|\Phi(X(t+\theta))|\mu(d\theta)\right\}dW(t), \tag{5.1}$$

where $\gamma_1 > 0$, $\mu \in \mathcal{M}_{2r}$ and $W(t)$ is a 1-dimensional Wiener process. $\Phi(x)$ is a continuous function satisfying

$$\Phi(x) = \begin{cases} \Phi_1(x), & x \in (-\infty, -\frac{1}{2}] \\ \Phi_2(x), & x \in [-\frac{1}{2}, \frac{1}{2}] \\ \Phi_3(x), & x \in [\frac{1}{2}, \infty), \end{cases}$$

where Φ_1, Φ_3 are Lipschitz continuous in that

$$|\Phi_1(x) - \Phi_1(y)|^2 \le \beta_1|x-y|^2, \qquad x, y \in \left(-\infty, -\frac{1}{2}\right],\ \beta_1 > 0,$$

$$|\Phi_3(x) - \Phi_3(y)|^2 \le \beta_3 |x-y|^2, \quad x, y \in \Big[\frac{1}{2}, \infty\Big), \ \beta_3 > 0,$$

and Φ_2 is Hölder-Dini continuous and satisfies

$$|\Phi_2(x) - \Phi_2(y)|^2 \le \beta_2 |x-y|^2 \log \frac{e^2}{|x-y|^2}, \quad x, y \in \Big[-\frac{1}{2}, \frac{1}{2}\Big], \ \beta_2 > 0. \tag{5.2}$$

If $2\gamma_1 > 1 + 6(\beta_1 + \beta_2 + \beta_3)\mu^{(2r)}$, then the assertions of Theorems 2.1, 3.1, and 4.3 hold.

Proof. Clearly, the diffusion coefficient satisfies Assumption 4.2. Now we first verify (5.1) satisfies Assumption 4.1 with the function $u(s) = \log(s^{-1}e^2 \vee e^2), s > 0$. Set $0u(0) = \lim_{s\to 0} su(s) = 0$ and it is easy to verify that $u(s) = \log(s^{-1}e^2 \vee e^2) \in \mathcal{U}$ is decreasing and satisfies (4.1), and $su(s^2)$ is increasing . Since $s\log(s^{-1}e^2 \vee e^2)$ is increasing and $\log(s^{-1}e^2 \vee e^2) \ge 1$, we see that for any $x, y \in \mathbf{R}$

$$\begin{aligned} |\Phi(x) - \Phi(y)|^2 &\le 3(\beta_1+\beta_2+\beta_3)|x-y|^2 \log\left(\frac{e^2}{|x-y|^2} \vee e^2\right) \\ &=: A_1 |x-y|^2 \log\left(\frac{e^2}{|x-y|^2} \vee e^2\right). \end{aligned} \tag{5.3}$$

Since $s\log(s^{-1}e^2 \vee e^2) \le 1 + 2s$ for $s \ge 0$, (5.3) implies

$$|\Phi(x)|^2 \le 2(1+\varepsilon)A_1|x|^2 + (1+\varepsilon)A_1 + (1+\frac{1}{\varepsilon})|\Phi(0)|^2 =: 2(1+\varepsilon)A_1|x|^2 + A_2. \tag{5.4}$$

By the Cauchy inequality and the Hölder inequality, for any $\phi, \psi \in \mathcal{C}_r$ we have

$$\begin{aligned} &2\langle \phi(0) - \psi(0), b(\phi) - b(\psi)\rangle_+ + \|\sigma(\phi) - \sigma(\psi)\|^2 \\ &= 2\Big\langle \phi(0) - \psi(0), -\gamma_1(\phi(0) - \psi(0)) + \int_{-\infty}^0 \Phi(\phi(\theta)) - \Phi(\psi(\theta))\mu(d\theta)\Big\rangle_+ \\ &\quad + \left|\int_{-\infty}^0 (1 \wedge |\Phi(\phi(\theta))|) - (1 \wedge |\Phi(\psi(\theta))|)\mu(d\theta)\right|^2 \\ &\le \Big(-2\gamma_1|\phi(0) - \psi(0)|^2 + \gamma_1|\phi(0) - \psi(0)|^2 \\ &\quad + \frac{1}{\gamma_1}\int_{-\infty}^0 |\Phi(\phi(\theta)) - \Phi(\psi(\theta))|^2\mu(d\theta)\Big)_+ \\ &\quad + \int_{-\infty}^0 |\Phi(\phi(\theta)) - \Phi(\psi(\theta))|^2\mu(d\theta) \\ &\le \left(\frac{1}{\gamma_1} + 1\right)\int_{-\infty}^0 |\Phi(\phi(\theta)) - \Phi(\psi(\theta))|^2\mu(d\theta) \\ &\le A_1\left(\frac{1}{\gamma_1} + 1\right)\int_{-\infty}^0 |\phi(\theta) - \psi(\theta)|^2 \log\left(\frac{e^2}{|\phi(\theta) - \psi(\theta)|^2} \vee e^2\right)\mu(d\theta). \end{aligned} \tag{5.5}$$

This implies Assumption 4.1 holds for $u(s) = \log(s^{-1}e^2 \vee e^2), s > 0$ and $K_1 = A_1(1/\gamma_1 + 1)$. Therefore, (5.1) has a unique global strong solution and the asymptotic log-Harnack inequality (4.17) in Theorem 4.3 holds. In addition, since $\lim_{s\to 0} s^{1-\beta}\log(e^2/s^2) = 0$ for any given

$\beta \in (0, 1)$, we have

$$s^2 \log\left(\frac{e^2}{s^2}\right) = s^{1+\beta}\left(s^{1-\beta}\log\left(\frac{e^2}{s^2}\right)\right) \le \sup_{s\in(0,1]}\left(s^{1-\beta}\log\left(\frac{e^2}{s^2}\right)\right)s^{1+\beta}$$
$$=: K_\beta s^{1+\beta} \quad \text{for } s \in (0, 1],$$

where K_β is a constant depending only on β. Since $s\log(s^{-1}e^2 \vee e^2)$ is concave on [0, 1] and nondecreasing and vanishes at 0, it follows from (5.5) that for $\|\phi(\theta) - \psi(\theta)\|_r \le 1$

$$\begin{aligned}
&2\langle \phi(0) - \psi(0), b(\phi) - b(\psi)\rangle_+ + \|\sigma(\phi) - \sigma(\psi)\|^2 \\
&\le \mu^{(2r)} A_1 \left(\frac{1}{\gamma_1} + 1\right)\|\phi - \psi\|_r^2 \log\left(\frac{e^2}{\|\phi(\theta) - \psi(\theta)\|_r^2} \vee e^2\right) \\
&\le K_\beta \mu^{(2r)} A_1 \left(\frac{1}{\gamma_1} + 1\right)\|\phi - \psi\|_r^{1+\beta},
\end{aligned}$$

which implies Assumption 3.1 holds for $K = K_\beta \mu^{(2r)} A_1 (1/\gamma_1 + 1)$ and $\delta = 1$. Hence, we only need to verify (3.2) to prove Theorem 3.1. By Itô's formula, for some $r_0 \in (0, 2r)$ and any initial data $X_t = \xi$, we have

$$\begin{aligned}
e^{r_0 t}|X(t)|^2 &= |\xi(0)|^2 + 2\int_0^t e^{r_0 s} X(s)\sigma(X_t)dW(t) \\
&\quad + \int_0^t e^{r_0 s}\left(r_0|X(s)|^2 + 2\langle X(s), b(X_s)\rangle + \|\sigma(X_s)\|^2\right)ds.
\end{aligned} \tag{5.6}$$

By using the Cauchy inequality and (5.4), we obtain

$$\begin{aligned}
&2\langle X(s), b(X_s)\rangle + \|\sigma(X_s)\|^2 \\
&\le -2\gamma_1|X(s)|^2 + 2\int_{-\infty}^0 |X(s)\Phi(X(s+\theta))|\mu(d\theta) \\
&\quad + \left|1 + \int_{-\infty}^0 1 \wedge |\Phi(X(t+\theta))|\mu(d\theta)\right|^2 \\
&\le -2\gamma_1|X(s)|^2 + |X(s)|^2 + \int_{-\infty}^0 |\Phi(X(s+\theta))|^2\mu(d\theta) + 4 \\
&\le -2\gamma_1|X(s)|^2 + |X(s)|^2 + 2(1+\varepsilon)A_1\int_{-\infty}^0 |X(s+\theta)|^2\mu(d\theta) + A_2 + 4.
\end{aligned} \tag{5.7}$$

Substituting (5.7) into (5.6) yields

$$\begin{aligned}
e^{r_0 t}|X(t)|^2 &\le \|\xi\|_r^2 + \frac{1}{r_0}(A_2 + 4)e^{r_0 t} + 2\int_0^t e^{r_0 s}X(s)\sigma(X_s)dW(s) \\
&\quad + (r_0 + 1 - 2\gamma_1)\int_0^t e^{r_0 s}|X(s)|^2 ds \\
&\quad + 2(1+\varepsilon)A_1\int_0^t\int_{-\infty}^0 e^{r_0 s}|X(s+\theta)|^2\mu(d\theta)ds.
\end{aligned} \tag{5.8}$$

Applying the Tonelli theorem and a substitution technique gives

$$\int_0^t\int_{-\infty}^0 e^{r_0 s}|X(s+\theta)|^2\mu(d\theta)ds$$

$$
\begin{aligned}
&\le \int_0^t \int_{-\infty}^{-s} e^{r_0 s}|X(s+\theta)|^2\mu(d\theta)ds + \int_0^t \int_{-s}^{0} e^{r_0 s}|X(s+\theta)|^2\mu(d\theta)ds \\
&\le \|\xi\|_r^2 \int_0^t \int_{-\infty}^{-s} e^{r_0 s}e^{-2r(s+\theta)}\mu(d\theta)ds + \int_{-t}^{0}\int_{-\theta}^{t} e^{r_0 s}|X(s+\theta)|^2 ds\mu(d\theta) \\
&\le \frac{1}{2r-r_0}\mu^{(2r)}\|\xi\|_r^2 + \int_{-\infty}^{0}\int_0^t e^{r_0(s-\theta)}|X(s)|^2 ds\mu(d\theta) \\
&\le \frac{1}{2r-r_0}\mu^{(2r)}\|\xi\|_r^2 + \mu^{(2r)}\int_0^t e^{r_0 s}|X(s)|^2 ds.
\end{aligned} \tag{5.9}
$$

By the Burkholder–Davis–Gundy inequality, we have

$$
\begin{aligned}
&2\mathbb{E}\sup_{0\le u\le t}\int_0^u e^{r_0 s}X(s)\sigma(X_s)dW(s) \\
&\le 2\sqrt{32}\mathbb{E}\left(\int_0^t e^{2r_0 s}\left|X(s)\big(1+\int_{-\infty}^{0} 1\wedge|\Phi(X(s+\theta))|\mu(d\theta)\big)\right|^2 ds\right)^{\frac{1}{2}} \\
&\le 2\sqrt{32}\mathbb{E}\left(\sup_{0\le s\le t} e^{r_0 s}|X(s)|^2\int_0^t 4e^{r_0 s}ds\right)^{\frac{1}{2}} \\
&\le \frac{1}{2}\mathbb{E}\sup_{0\le s\le t} e^{r_0 s}|X(s)|^2 + \frac{256}{r_0}e^{r_0 t}.
\end{aligned} \tag{5.10}
$$

Noting that $2\gamma_1 > 1+6(\beta_1+\beta_2+\beta_3)\mu^{(2r)}$, we can find constants $\varepsilon, r_0 > 0$ small enough such that

$$
r_0 + 1 - 2\gamma_1 + 2(1+\varepsilon)A_1\mu^{(r_0)} \le 0.
$$

Then substituting (5.9) and (5.10) into (5.8) yields that

$$
\mathbb{E}\sup_{0\le s\le t} e^{r_0 s}|X(s)|^2 \le 2\left(1+\frac{2(1+\varepsilon)A_1\mu^{(2r)}}{2r-r_0}\right)\|\xi\|_r^2 + \frac{520+2A_2}{r_0}e^{r_0 t}. \tag{5.11}
$$

Recall that for $r_0 \le 2r$,

$$
\mathbb{E}\|X_t\|_r^2 \le e^{-r_0 t}\|\xi\|_r^2 + e^{-r_0 t}\mathbb{E}\sup_{0\le s\le t} e^{r_0 s}|X(s)|^2.
$$

Then by (5.8), we have

$$
\mathbb{E}\|X_t\|_r^2 \le \left(3+\frac{4(1+\varepsilon)A_1\mu^{(2r)}}{2r-r_0}\right)e^{-r_0 t}\|\xi\|_r^2 + \frac{520+2A_2}{r_0}.
$$

This shows (3.2) holds for $V(\xi) = \|\xi\|_r^2$. Therefore Theorem 3.1 holds for (5.1). □

Remark 5.2. In fact, by a slight modification of the above proof, one can prove (3.2) still holds for $\sigma(\phi) = 1 + \int_{-\infty}^{0}|\Phi(\phi(\theta))|\mu(d\theta)$ when $\gamma_1 > 1 + 6(\beta_1+\beta_2+\beta_3)\mu^{(2r)}$. In general, it is not easy to verify the Lyapunov condition (3.2) for SFDEs (see, e.g., [3,8]).

Remark 5.3. We now show that there is Hölder-Dini continuous function Φ_2 such that (5.2) holds. Let

$$
f(x) = x^{1-a}\log\frac{e^2}{x^2}, \quad x \in (0,1], a \in (0,1).
$$

We compute

$$f'(x) = 2(1-a)x^{-a}\Big(-\log x - \frac{a}{1-a}\Big).$$

Then $f'(x) > 0$ for $x \in (0, \exp[-a/(1-a)])$ and $f'(x) \le 0$ for $x \in [\exp[-a/(1-a)], 1]$. Noting that $\lim_{x\to 0} f(x) = 0$ and $f(1) = 2$, there exists a unique constant $x_0 \in (0, \exp[-a/(1-a)])$ such that $f(x_0) = 2$. Then define

$$\varphi(x) = \begin{cases} \sqrt{f(x)}, & x \in (0, x_0) \\ \sqrt{2}, & x \in [x_0, \infty). \end{cases}$$

Clearly, $\varphi(x)$ is increasing and $\varphi(x) \le \sqrt{f(x)}$ for $x \in (0, 1]$. Moreover, it is easy to verify that

$$\int_0^1 \frac{\varphi(x)}{x} dx = \int_0^{x_0} \frac{\sqrt{x^{1-a}\log\frac{e^2}{x^2}}}{x} dx + \int_{x_0}^1 \frac{\sqrt{2}}{x} dx \le -\sqrt{2}\log x_0 + \int_0^{x_0} \sqrt{\frac{\log\frac{e^2}{x^2}}{x^{1+a}}} dx < \infty$$

where in the last step we have used the fact $\lim_{s\to 0} s^\beta \log(e^2/s^2) = 0$ for any given $\beta \in (0, 1)$. This implies the function $\varphi(x)$ is a Dini function. Hence, we can choose the function Φ_2 satisfying the following inequality

$$|\Phi_2(x) - \Phi_2(y)| \le |x-y|^{\frac{1+a}{2}} \varphi(|x-y|), \quad |x-y| \le 1,$$

which indicates Φ_2 is Hölder-Dini continuous. This, together with $\varphi(x) \le \sqrt{f(x)}$ for $x \in (0, 1]$, means that such Hölder-Dini continuous function Φ_2 satisfies (5.2).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

This section includes two auxiliary tail-estimates. The first one is to establish the Fernique-type inequality for Wiener processes, which is a direct consequence of the [21, Lemma 3.1]. The second one is a slight modification of [20, Lemma B.1].

Lemma A.1. *Let $W(t)$ be a 1-dimensional Wiener process. For any given $T > 0$, $\delta \in (0, 1/2)$, and for all $x > 0$ satisfying $x \ge (1 + 8\log 2)^{1/2}(T^{1-2\delta} + x^{-1})$,*

$$\mathbb{P}\left\{\sup_{t_1,t_2\in[0,T]} \frac{|W(t_1) - W(t_2)|}{|t_1 - t_2|^\delta} > x\right\}$$
$$\le C\big(Tx^{\frac{1-2\delta}{2}} + 1\big)^2 x^{-1}\big(T^{1-2\delta} + x^{-1}\big)\exp\left\{-\frac{x^2}{2(T^{1-2\delta} + x^{-1})^2}\right\},$$

where $C > 0$ depends only on δ.

Proof. Since $W(t)$ is a Wiener process, for any $t_1, t_2 \in [0, T]$,

$$\frac{W(t_1) - W(t_2)}{|t_1 - t_2|^\delta} \sim N(0, |t_1 - t_2|^{1-2\delta}).$$

Then we claim that

$$\gamma^2(\varepsilon) := \sup_{\substack{t_i, s_i \in [0,T] \\ |s_i - t_i| \le \varepsilon, i=1,2}} \mathbb{E}\left(\frac{W(t_1) - W(t_2)}{|t_1 - t_2|^\delta} - \frac{W(s_1) - W(s_2)}{|s_1 - s_2|^\delta} \right)^2 \le 5\varepsilon^{1-2\delta}. \tag{A.1}$$

We will prove (A.1) momentarily. As a consequence of (A.1), we obtain

$$\begin{aligned} Q(h) &:= (2+\sqrt{2}) \int_1^\infty \gamma(h 2^{-y^2}) dy \\ &\le \sqrt{5}(2+\sqrt{2}) \int_1^\infty 2^{-\frac{1-2\delta}{2} y^2} dy h^{\frac{1-2\delta}{2}} =: \bar{Q}(h). \end{aligned}$$

Note that

$$Q^{-1}(x) \ge \bar{Q}^{-1}(x) = C_1 x^{\frac{2}{1-2\delta}},$$

where $Q^{-1}(x) := \sup\{y : Q(y) \le x\}$, and C_1 depends only on δ. In addition,

$$\sigma^2 := \sup_{t_1, t_2 \in [0,T]} \mathbb{E}\left(\frac{W(t_1) - W(t_2)}{|t_1 - t_2|^\delta} \right)^2 = T^{1-2\delta}.$$

Hence the desired result follows from [21, Lemma 3.1].

We now prove the claim (A.1). First, we write $\gamma^2(\varepsilon) = \sup_{t_i, s_i \in [0,T], |s_i - t_i| \le \varepsilon, i=1,2} H(s_1, s_2, t_1, t_2)$, where

$$\begin{aligned} H(s_1, s_2, t_1, t_2) &:= \mathbb{E}\left(\frac{W(t_1) - W(t_2)}{|t_1 - t_2|^\delta} - \frac{W(s_1) - W(s_2)}{|s_1 - s_2|^\delta} \right)^2 \\ &= |t_1 - t_2|^{1-2\delta} + |s_1 - s_2|^{1-2\delta} - \frac{2\mathbb{E}\big[W(t_1) - W(t_2)\big]\big[W(s_1) - W(s_2)\big]}{|t_1 - t_2|^\delta |s_1 - s_2|^\delta}. \end{aligned}$$

We next derive an upper bound for $H(s_1, s_2, t_1, t_2)$ in different cases. Without loss of generality, we can assume $t_1 < t_2$. When $s_1 < s_2$, there are five cases to consider.

Case I1: $s_1 < s_2 \le t_1 < t_2$ or $t_1 < t_2 \le s_1 < s_2$.

$$H(s_1, s_2, t_1, t_2) = |t_1 - t_2|^{1-2\delta} + |s_1 - s_2|^{1-2\delta} \le |s_2 - t_2|^{1-2\delta} + |s_1 - t_1|^{1-2\delta}.$$

Case I2: $s_1 \le t_1 \le s_2 \le t_2$.

$$\begin{aligned} H(s_1, s_2, t_1, t_2) &= |t_1 - t_2|^{1-2\delta} + |s_1 - s_2|^{1-2\delta} - \frac{2(s_2 - t_1)}{|t_1 - t_2|^\delta |s_1 - s_2|^\delta} \\ &\le 2(|t_1 - t_2| \vee |s_1 - s_2|)^{1-2\delta} - \frac{2(s_2 - t_1)}{(|t_1 - t_2| \vee |s_1 - s_2|)^{2\delta}}. \end{aligned}$$

Then if $|t_1 - t_2| \le |s_1 - s_2|$, we have

$$H(s_1, s_2, t_1, t_2) \le 2|s_1 - s_2|^{1-2\delta} - \frac{2(s_2 - t_1)}{|s_1 - s_2|^{2\delta}} \le \frac{2(t_1 - s_1)}{|s_1 - s_2|^{2\delta}} \le 2|t_1 - s_1|^{1-2\delta}.$$

Similarly, when $|t_1 - t_2| > |s_1 - s_2|$, we have $H(s_1, s_2, t_1, t_2) \le 2|t_2 - s_2|^{1-2\delta}$.

Case I3: $s_1 \le t_1 < t_2 \le s_2$.

$$\begin{aligned} H(s_1, s_2, t_1, t_2) &= |t_1 - t_2|^{1-2\delta} + |s_1 - s_2|^{1-2\delta} - \frac{2(t_2 - t_1)}{|t_1 - t_2|^\delta |s_1 - s_2|^\delta} \\ &\le 2|s_1 - s_2|^{1-2\delta} - \frac{2(t_2 - t_1)}{|s_1 - s_2|^{2\delta}} \end{aligned}$$

$$\leq \frac{2(t_1 - s_1) + 2(s_2 - t_2)}{|s_1 - s_2|^{2\delta}}$$
$$\leq 2|t_1 - s_1|^{1-2\delta} + 2|s_2 - t_2|^{1-2\delta}$$

Case I4: $t_1 \leq s_1 < s_2 \leq t_2$. Similar to Case I3, we have

$$H(s_1, s_2, t_1, t_2) = |t_1 - t_2|^{1-2\delta} + |s_1 - s_2|^{1-2\delta} - \frac{2(s_2 - s_1)}{|t_1 - t_2|^{\delta}|s_1 - s_2|^{\delta}}$$
$$\leq 2|t_2 - s_2|^{1-2\delta} + 2|t_1 - s_1|^{1-2\delta}$$

Case I5: $t_1 \leq s_1 \leq t_2 \leq s_2$. Similar to Case I2, we obtain

$$H(s_1, s_2, t_1, t_2) = |t_1 - t_2|^{1-2\delta} + |s_1 - s_2|^{1-2\delta} - \frac{2(t_2 - s_1)}{|t_1 - t_2|^{\delta}|s_1 - s_2|^{\delta}}$$
$$\leq 2|s_1 - t_1|^{1-2\delta} + 2|s_2 - t_2|^{1-2\delta}.$$

When $s_1 \geq s_2$, we also have five cases as follows.

Case II1: $s_2 \leq s_1 \leq t_1 < t_2$ or $t_1 \leq t_2 \leq s_2 \leq s_1$. Similar to Case I1, we have

$$H(s_1, s_2, t_1, t_2) \leq |s_2 - t_2|^{1-2\delta} + |s_1 - t_1|^{1-2\delta}.$$

Case II2: $s_2 \leq t_1 \leq s_1 \leq t_2$.

$$H(s_1, s_2, t_1, t_2) = |t_1 - t_2|^{1-2\delta} + |s_1 - s_2|^{1-2\delta} - \frac{2(t_1 - s_1)}{|t_1 - t_2|^{\delta}|s_1 - s_2|^{\delta}}$$
$$\leq 2|t_2 - s_2|^{1-2\delta} + \frac{2(s_1 - t_1)}{|t_1 - t_2|^{\delta}|s_1 - s_2|^{\delta}}$$
$$\leq 2|t_2 - s_2|^{1-2\delta} + 2|t_1 - s_1|^{1-2\delta}.$$

Case II3: $s_2 \leq t_1 < t_2 \leq s_1$.

$$H(s_1, s_2, t_1, t_2) = |t_1 - t_2|^{1-2\delta} + |s_1 - s_2|^{1-2\delta} - \frac{2(t_1 - t_2)}{|t_1 - t_2|^{\delta}|s_1 - s_2|^{\delta}}$$
$$\leq 2|t_2 - s_2|^{1-2\delta} + |t_1 - s_1|^{1-2\delta} + \frac{2(t_2 - t_1)}{|t_1 - t_2|^{2\delta}}$$
$$\leq 4|t_2 - s_2|^{1-2\delta} + |t_1 - s_1|^{1-2\delta}.$$

Case II4: $t_1 \leq s_2 \leq s_1 \leq t_2$. Similar to Case II3, we arrive at

$$H(s_1, s_2, t_1, t_2) = |t_1 - t_2|^{1-2\delta} + |s_1 - s_2|^{1-2\delta} - \frac{2(s_2 - s_1)}{|t_1 - t_2|^{\delta}|s_1 - s_2|^{\delta}}$$
$$\leq 4|t_2 - s_2|^{1-2\delta} + |t_1 - s_1|^{1-2\delta}.$$

Case II5: $t_1 \leq s_2 \leq t_2 \leq s_1$. Similar to Case II2, we get

$$H(s_1, s_2, t_1, t_2) = |t_1 - t_2|^{1-2\delta} + |s_1 - s_2|^{1-2\delta} - \frac{2(s_2 - t_2)}{|t_1 - t_2|^{\delta}|s_1 - s_2|^{\delta}}$$
$$\leq 2|t_1 - s_1|^{1-2\delta} + 2|t_2 - s_2|^{1-2\delta}.$$

Summarizing the above observations, we have

$$\gamma^2(\varepsilon) = \sup_{t_i, s_i \in [0,T], |s_i - t_i| \leq \varepsilon, i=1,2} H(s_1, s_2, t_1, t_2) \leq 5\varepsilon^{1-2\delta}.$$

The claim (A.1) is established as desired. □

Lemma A.2. *Let $Y(t) \geq 0$ be an Itô process with*

$$dY(t) = U(t)dt + dM(t),$$

where M is a continuous local martingale with quadratic variation

$$\langle M\rangle(t) = \int_0^t m(s)ds, \quad t \geq 0.$$

If there exist constants $A \geq 0, B > 0, \lambda > 0, T > 0$ and random variable $\tau \in [0, T]$ such that

$$U(t) \leq -\lambda Y(t) + A, \quad m(t) \leq B, \quad \text{whenever } t \leq \tau.$$

Then for any $\delta \in (0, 1/2)$, there exist constants $C_1, C_2 > 0$, which depend only on δ, such that for any $R \geq 1$ satisfying

$$R \geq (1 + 8\log 2)^{1/2}(\Gamma(\delta) + \sup_{x>0} x^\delta e^{-x})\big(T^{1-2\delta} + R^{-1}(\Gamma(\delta) + \sup_{x>0} x^\delta e^{-x})\big),$$

$$\mathbb{P}\left\{\sup_{t\leq\tau}\big(Y(t) - e^{-\lambda t}Y(0)\big) \geq A\lambda^{-1} + B^{1/2}\lambda^{-\delta}R\right\}$$
$$\leq C_1(T^{\frac{5-2\delta}{2}} + 1)R^{-2\delta}\exp\left\{-\frac{C_2R^2}{2(T^{\frac{1-2\delta}{2}} + 1)^2}\right\}.$$

Proof. This result follows directly from Lemma A.1 and the same argument as that in the proof of [20, Lemma B.1]. □

Proposition A.3. *Under Assumptions 4.1 and 4.2, $\tilde{b}(\phi) := \lambda\sigma(\phi)\sigma^{-1}(\xi)\Gamma(\xi(0), \phi(0))$, $\phi \in \mathcal{C}_r$ as a drift satisfies Assumptions 2.1 and 2.2 for any fixed $\xi \in \mathcal{C}_r$, where $\Gamma(\cdot, \cdot)$ is given by (4.3).*

Proof. Since $su(s)$ is increasing, according to the definition of $\Gamma(\cdot, \cdot)$, we need only to verify $\tilde{b}$ satisfies Assumptions 2.1 and 2.2 for each component. Hence without any loss of generality, we can assume $d = 1$ in what follows and hence $\tilde{b}$ takes the following form:

$$\tilde{b}(\phi) := \lambda\sigma(\phi)\sigma^{-1}(\xi)(\xi(0) - \phi(0))u(|\xi(0) - \phi(0)|^2), \quad \phi \in \mathcal{C}_r.$$

In light of Assumptions 4.1 and 4.2, it is easy to see that $\tilde{b}$ satisfies the linear growth condition, which implies that $\tilde{b}$ satisfies Assumption 2.2 and is bounded on bounded subset of $\mathcal{C}_r$. Now it remains to show that $\tilde{b}$ satisfies (2.1). For any $\phi, \psi \in \mathcal{C}_r$ and fixed $\xi \in \mathcal{C}_r$ with $\|\phi\|_r \vee \|\psi\|_r < k$ for some $k > 0$, we compute

$$\begin{aligned}
&\langle \phi(0) - \psi(0), \tilde{b}(\phi) - \tilde{b}(\psi)\rangle \\
&\quad\leq \lambda\|\sigma^{-1}\|_\infty|\phi(0) - \psi(0)|\|\sigma(\phi) - \sigma(\psi)\||\xi(0) - \phi(0)|u(|\xi(0) - \phi(0)|^2) \\
&\qquad + \lambda\|\sigma\|_\infty\|\sigma^{-1}\|_\infty|\phi(0) - \psi(0)| \\
&\qquad \times \big|(\xi(0) - \phi(0))u(|\xi(0) - \phi(0)|^2) - (\xi(0) - \psi(0))u(|\xi(0) - \psi(0)|^2)\big| \\
&\quad =: \lambda\|\sigma^{-1}\|_\infty \Upsilon_1(\phi, \psi) + \lambda\|\sigma\|_\infty\|\sigma^{-1}\|_\infty \Upsilon_2(\phi, \psi).
\end{aligned}$$

Noting that $su(s)$ is increasing and concave, by Assumption 4.1, we obtain

$$\begin{aligned}
\|\sigma(\phi) - \sigma(\psi)\|^2 &\leq K_1\int_{-\infty}^0 e^{-2r\theta}e^{2r\theta}|\phi(\theta) - \psi(\theta)|^2u(e^{-2r\theta}e^{2r\theta}|\phi(\theta) - \psi(\theta)|^2)\mu(d\theta) \\
&\leq K_1\int_{-\infty}^0 e^{-2r\theta}\|\phi - \psi\|_r^2u(e^{-2r\theta}\|\phi - \psi\|_r^2)\mu(d\theta)
\end{aligned}$$

$$
\begin{aligned}
&\le K_1 \int_{-\infty}^{0} e^{-2r\theta} \|\phi - \psi\|_r^2 u(\|\phi(\theta) - \psi(\theta)\|_r^2)\mu(d\theta) \\
&\le K_1 \mu^{(2r)} \|\phi - \psi\|_r^2 u(\|\phi(\theta) - \psi(\theta)\|_r^2).
\end{aligned}
$$

Recalling that $u(\cdot) \ge 1$ and $su(s^2)$ is increasing, then we arrive at

$$
\Upsilon_1(\phi, \psi) \le \sqrt{K_1 \mu^{(2r)}} (k + \|\xi\|_r) u(k + \|\xi\|_r) \|\phi - \psi\|_r^2 u(\|\phi(\theta) - \psi(\theta)\|_r^2).
$$

We now estimate $\Upsilon_2(\phi, \psi)$ and assume $\phi(0) \le \psi(0)$ without any loss of generality.

Case (i): $\phi(0) \le \xi(0) \le \psi(0)$. In this case, we have $|\xi(0) - \phi(0)| \vee |\xi(0) - \psi(0)| \le |\phi(0) - \psi(0)|$. Since $su(s^2)$ is increasing, we obtain

$$
\Upsilon_2(\phi, \psi) \le 2|\phi(0) - \psi(0)|^2 u(|\phi(0) - \psi(0)|^2).
$$

Case (ii): $\xi(0) \le \phi(0) \le \psi(0)$. Since u is decreasing and $su(s^2)$ is increasing, we have

$$
\begin{aligned}
0 &\le (\xi(0) - \phi(0)) u(|\xi(0) - \phi(0)|^2) - (\xi(0) - \psi(0)) u(|\xi(0) - \psi(0)|^2) \\
&= (\xi(0) - \phi(0))(u(|\xi(0) - \phi(0)|^2) - u(|\xi(0) - \psi(0)|^2)) \\
&\quad + (\psi(0) - \psi(0)) u(|\xi(0) - \psi(0)|^2) \\
&\le (\psi(0) - \psi(0)) u(|\xi(0) - \psi(0)|^2) \le (\psi(0) - \psi(0)) u(|\phi(0) - \psi(0)|^2),
\end{aligned}
$$

which implies $\Upsilon_2(\phi, \psi) \le |\phi(0) - \psi(0)|^2 u(|\phi(0) - \psi(0)|^2)$.

Case (iii): $\phi(0) \le \psi(0) \le \xi(0)$. We compute

$$
\begin{aligned}
0 &\le (\xi(0) - \phi(0)) u(|\xi(0) - \phi(0)|^2) - (\xi(0) - \psi(0)) u(|\xi(0) - \psi(0)|^2) \\
&\le (\psi(0) - \phi(0)) u(|\xi(0) - \phi(0)|^2) \le (\psi(0) - \phi(0)) u(|\psi(0) - \phi(0)|^2),
\end{aligned}
$$

which implies $\Upsilon_2(\phi, \psi) \le |\phi(0) - \psi(0)|^2 u(|\phi(0) - \psi(0)|^2)$.

Summarizing the above estimations and noting that $su(s)$ is increasing, we have

$$
\begin{aligned}
\langle \phi(0) - \psi(0), \tilde{b}(\phi) - \tilde{b}(\psi) \rangle &\le \lambda \|\sigma^{-1}\|_\infty \sqrt{K_1 \mu^{(2r)}} (k + \|\xi\|_r) u(k + \|\xi\|_r) \|\phi - \psi\|_r^2 \\
&\quad \times u(\|\phi(\theta) - \psi(\theta)\|_r^2) \\
&\quad + \lambda \|\sigma\|_\infty \|\sigma^{-1}\|_\infty |\phi(0) - \psi(0)|^2 u(|\phi(0) - \psi(0)|^2) \\
&\le L_k \|\phi - \psi\|_r^2 u(\|\phi - \psi\|_r^2),
\end{aligned}
$$

for some $L_k > 0$. Hence, $\tilde{b}$ satisfies Assumptions 2.1 and 2.2 under Assumptions 4.1 and 4.2. □

References

[1] M. Arnaudon, A. Thalmaier, F.-Y. Wang, Gradient estimates and Harnack inequalities on non-compact Riemannian manifolds, Stochastic Process. Appl. 119 (10) (2009) 3653–3670.

[2] J. Bao, F.-Y. Wang, C. Yuan, Asymptotic log-Harnack inequality and applications for stochastic systems of infinite memory, Stochastic Process. Appl. 129 (11) (2019) 4576–4596.

[3] J. Bao, F.-Y. Wang, C. Yuan, Ergodicity for neutral type SDEs with infinite length of memory, Math. Nachr. 293 (9) (2020) 1675–1690.

[4] J. Bao, G. Yin, C. Yuan, Ergodicity for functional stochastic differential equations and applications, Nonlinear Anal. 98 (2014) 66–82.

[5] J. Bao, G. Yin, C. Yuan, L.Y. Wang, Exponential ergodicity for retarded stochastic differential equations, Appl. Anal. 93 (11) (2014) 2330–2349.

[6] O. Butkovsky, Subgeometric rates of convergence of Markov processes in the Wasserstein metric, Ann. Appl. Probab. 24 (2) (2014) 526–552.

[7] O. Butkovsky, A. Kulik, M. Scheutzow, Generalized couplings and ergodic rates for SPDEs and other Markov models, Ann. Appl. Probab. 30 (1) (2020) 1–39.
[8] O. Butkovsky, M. Scheutzow, Invariant measures for stochastic functional differential equations, Electron. J. Probab. 22 (2017) 23, Paper No. 98.
[9] J.C. Cox, J.E. Ingersoll Jr., S.A. Ross, A theory of the term structure of interest rates, Econometrica 53 (2) (1985) 385–407.
[10] G. Da Prato, J. Zabczyk, Ergodicity for Infinite-Dimensional Systems, in: London Mathematical Society Lecture Note Series, vol. 229, Cambridge University Press, Cambridge, 1996.
[11] A. Es-Sarhir, M.-K. von Renesse, M. Scheutzow, Harnack inequality for functional SDEs with bounded memory, Electron. Commun. Probab. 14 (2009) 560–565.
[12] A. Es-Sarhir, M. Scheutzow, O. van Gaans, Invariant measures for stochastic functional differential equations with superlinear drift term, Differential Integral Equations 23 (1–2) (2010) 189–200.
[13] S. Fang, T. Zhang, A study of a class of stochastic differential equations with non-Lipschitzian coefficients, Probab. Theory Related Fields 132 (3) (2005) 356–390.
[14] M. Hairer, J.C. Mattingly, Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing, Ann. of Math. (2) 164 (3) (2006) 993–1032.
[15] M. Hairer, J.C. Mattingly, M. Scheutzow, Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations, Probab. Theory Related Fields 149 (1–2) (2011) 223–259.
[16] Y. Hino, S. Murakami, T. Naito, Functional-Differential Equations with Infinite Delay, in: Lecture Notes in Mathematics, vol. 1473, Springer-Verlag, Berlin, 1991.
[17] N. Ikeda, S. Watanabe, Stochastic Differential Equations and Diffusion Processes, second ed., in: North-Holland Mathematical Library, vol. 24, North-Holland Publishing Co., Kodansha, Ltd., Amsterdam, Tokyo, 1989.
[18] I. Karatzas, S.E. Shreve, Brownian Motion and Stochastic Calculus, second ed., in: Graduate Texts in Mathematics, vol. 113, Springer-Verlag, New York, 1991.
[19] R. Khasminskii, Stochastic Stability of Differential Equations, second ed., in: Stochastic Modelling and Applied Probability, vol. 66, Springer, New York, 2012.
[20] A. Kulik, M. Scheutzow, Well-posedness, stability, and sensitivities for stochastic delay equations: A generalized coupling approach, Ann. Probab. 48 (6) (2020) 3041–3076.
[21] M.M. Meerschaert, W. Wang, Y. Xiao, Fernique-type inequalities and moduli of continuity for anisotropic Gaussian random fields, Trans. Amer. Math. Soc. 365 (2) (2013) 1081–1107.
[22] E. Odell, On the structure of separable infinite dimensional Banach spaces, 2007, http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.474.3251.
[23] M.-K. von Renesse, M. Scheutzow, Existence and uniqueness of solutions of stochastic functional differential equations, Random Oper. Stoch. Equ. 18 (3) (2010) 267–284.
[24] D. Revuz, M. Yor, Continuous Martingales and Brownian Motion, third ed., in: Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, Springer-Verlag, Berlin, 1999.
[25] M. Röckner, F.-Y. Wang, Log-harnack inequality for stochastic differential equations in Hilbert spaces and its consequences, Infin. Dimens. Anal. Quantum Probab. Relat. Top. 13 (1) (2010) 27–37.
[26] M. Scheutzow, Exponential growth rates for stochastic delay differential equations, Stoch. Dyn. 5 (2) (2005) 163–174.
[27] J. Shao, F.-Y. Wang, C. Yuan, Harnack inequalities for stochastic (functional) differential equations with non-Lipschitzian coefficients, Electron. J. Probab. 17 (100) (2012) 18.
[28] D.W. Stroock, S.R.S. Varadhan, Multidimensional Diffusion Processes, in: Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 233, Springer-Verlag, Berlin, 1979.
[29] F.-Y. Wang, Logarithmic Sobolev inequalities on noncompact Riemannian manifolds, Probab. Theory Related Fields 109 (3) (1997) 417–424.
[30] F.-Y. Wang, Harnack inequalities on manifolds with boundary and applications, J. Math. Pures Appl. 94 (3) (2010) 304–321.
[31] F.-Y. Wang, Harnack inequality for SDE with multiplicative noise and extension to Neumann semigroup on nonconvex manifolds, Ann. Probab. 39 (4) (2011) 1449–1467.
[32] F.-Y. Wang, C. Yuan, Harnack inequalities for functional SDEs with multiplicative noise and applications, Stochastic Process. Appl. 121 (11) (2011) 2692–2710.
[33] F. Wu, G. Yin, H. Mei, Stochastic functional differential equations with infinite delay: existence and uniqueness of solutions, solution maps, Markov properties, and ergodicity, J. Differential Equations 262 (3) (2017) 1226–1252.

[34] L. Xu, A modified log-Harnack inequality and asymptotically strong Feller property, J. Evol. Equ. 11 (4) (2011) 925–942.

[35] T. Yamada, S. Watanabe, On the uniqueness of solutions of stochastic differential equations, J. Math. Kyoto Univ. 11 (1971) 155–167.

[36] X. Zhang, Exponential ergodicity of non-Lipschitz stochastic differential equations, Proc. Amer. Math. Soc. 137 (1) (2009) 329–337.