Paper:
Bao, J., Wang, F. & Yuan, C. (2015). Hypercontractivity for functional stochastic differential equations. *Stochastic Processes and their Applications, 125*(9), 3636-3656.

http://dx.doi.org/10.1016/j.spa.2015.04.001

Hypercontractivity for functional stochastic differential equations

Jianhai Bao, Feng-Yu Wang, Chenggui Yuan

PII: S0304-4149(15)00102-7
DOI: http://dx.doi.org/10.1016/j.spa.2015.04.001
Reference: SPA 2773

To appear in: *Stochastic Processes and their Applications*

Received date: 5 September 2014
Revised date: 24 March 2015
Accepted date: 8 April 2015

Please cite this article as: J. Bao, F.-Y. Wang, C. Yuan, Hypercontractivity for functional stochastic differential equations, *Stochastic Processes and their Applications* (2015), http://dx.doi.org/10.1016/j.spa.2015.04.001

Hypercontractivity for Functional Stochastic Differential Equations*

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

[a)]School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China

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

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

jianhaibao13@gmail.com, wangfy@bnu.edu.cn, F.-Y.Wang@swansea.ac.uk, C.Yuan@swansea.ac.uk

March 24, 2015

Abstract

An explicit sufficient condition on the hypercontractivity is derived for the Markov semigroup associated with a class of functional stochastic differential equations. Consequently, the semigroup P_t converges exponentially to its unique invariant probability measure μ in both $L^2(\mu)$ and the totally variational norm $\|\cdot\|_{\rm var}$, and it is compact in $L^2(\mu)$ for sufficiently large $t > 0$. This provides a natural class of non-symmetric Markov semigroups which are compact for large time but non-compact for small time. A semi-linear model which may not satisfy this sufficient condition is also investigated. As the associated Dirichlet form does not satisfy the log-Sobolev inequality, the standard argument using functional inequalities does not work.

AMS subject Classification: 65G17, 65G60
Keywords: Hypercontractivity, compactness, exponential ergodicity, functional stochastic differential equation, Harnack inequality.

1 Introduction

The hypercontractivity, first found by Nelson [17] for the Ornstein-Ulenbeck semigroup, has been investigated intensively for various models of Markov semigroups, see, for instance, [3, 7, 11, 21, 23, 24] and references within. However, so far there is no any result on this property for the semigroup associated with functional stochastic differential equations (FSDEs, or SDEs with memory).

*Supported in part by Lab. Math. Com. Sys. and NNSFs of China (Nos. 11131003, 11401592).

It is well known by Gross (see [11]) that the log-Sobolev inequality implies the hypercontractivity. However, for SDEs with delay the log-Sobolev inequality for the associated Dirichlet form does not hold. Indeed, according to [21, Theorem 3.3.6], the super Poincaré inequality (and hence the log-Sobolev inequality) implies the uniform integrability of the associated Markov semigroup P_t for all $t > 0$, which is not the case for the Markov semigroup associated with SDEs with delay, since in this case P_t is not uniformly integrable for t smaller than the length of time delay, see Remark 1.1(2) for more details.

On the other hand, the dimension-free Harnack inequality introduced in [20] and further developed in numerous papers is a powerful tool in the study of the hypercontractivity, which works well even for non-linear SPDEs (see, e.g., [16, 22]). Recently, this type Harnack inequalities have been investigated in [25] for FSDEs. To derive the hypercontractivity and exponential ergodicity from the dimension-free Harnack inequality, the key point is to prove the Gauss-type concentration property of the unique invariant probability measure with respect to the uniform norm on the state space, which is, however, not easy for FSDEs. We will see that our proof of the exponential integrability is tricky (see the proof of Lemma 2.1).

Let $r_0 > 0$ be fixed, and let $\mathscr{C} = C([-r_0, 0]; \mathbb{R}^d)$ be equipped with the uniform norm $\|\cdot\|_\infty$. Let $\mathscr{B}_b(\mathscr{C})$ be the set of all bounded measurable functions from $\mathscr{C}$ to $\mathbb{R}$. Let $\{B(t)\}_{t\geq 0}$ be a d-dimensional Brownian motion defined on a complete filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$. Let σ be an invertible $d\times d$-matrix, $a \in C(\mathbb{R}^d; \mathbb{R}^d)$ and $b : \mathscr{C} \to \mathbb{R}^d$ be Lipschitz continuous. Consider the following FSDE on $(\mathbb{R}^d, \langle\cdot,\cdot\rangle, |\cdot|)$:

$$\mathrm{d}X(t) = \big\{a(X(t)) + b(X_t)\big\}\mathrm{d}t + \sigma \mathrm{d}B(t), \quad t > 0, \quad X_0 = \xi \in \mathscr{C}. \tag{1.1}$$

Herein, for each $t \geq 0$, $X_t \in \mathscr{C}$ is fixed by $X_t(\theta) := X(t+\theta), \theta \in [-r_0, 0]$, and is called the segment process of $X(t)$.

Assume that

$$2\langle a(\xi(0)) - a(\eta(0)) + b(\xi) - b(\eta), \xi(0) - \eta(0)\rangle \leq \lambda_2\|\xi - \eta\|_\infty^2 - \lambda_1|\xi(0) - \eta(0)|^2, \quad \xi, \eta \in \mathscr{C} \tag{1.2}$$

holds for some constants $\lambda_1, \lambda_2 \geq 0$. Then, the equation (1.1) has a unique non-explosive strong solution denoted by $\{X^\xi(t)\}_{t\geq -r_0}$ with the initial segment $X_0 = \xi$ (see, e.g., [19, Theorem 2.3]). The segment process is denoted by $\{X_t^\xi\}_{t\geq 0}$. Let P_t be the Markov semigroup corresponding to the segment (functional) solution $\{X_t^\xi\}_{t\geq 0}$, i.e.

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

To study the hypercontractivity, it is essential to know the existence and uniqueness of invariant probability measures of $\{X_t^\xi\}_{t\geq 0}$. For existence of invariant probability measures for FSDEs, we refer to Es-Sarhir et al. [8] and Kinnally–Williams [14] by adopting the Arzelá–Ascoli tightness characterization, and Reiβ et al. [18] by considering the semi-martingale characteristics; With regards to uniqueness of invariant probability measures for FSDEs, we refer to Hairer et al. [13] by using asymptotic coupling approach.

The following is the first main result of the paper.

Theorem 1.1. *If $\lambda := \sup_{s\in[0,\lambda_1]}\big(s-\lambda_2 e^{r_0 s}\big)>0$, then P_t has a unique invariant probability measure μ, and the following assertions hold.*

(1) *P_t is hypercontractive, i.e., $\|P_t\|_{2\to4}\le 1$ holds for large enough $t>0$, where $\|\cdot\|_{2\to4}$ is the operator norm from $L^2(\mu)$ to $L^4(\mu)$.*

(2) *P_t is compact on $L^2(\mu)$ for large enough $t>0$.*

(3) *There exists a constant $C>0$ such that*

$$\|P_t-\mu\|_2^2 := \sup_{\mu(f^2)\le 1}\mu\big((P_tf-\mu(f))^2\big)\le C e^{-\lambda t},\quad t\ge 0,$$

where $\mu(f):=\int_{\mathscr{C}} f(\xi)\mu(\mathrm{d}\xi), f\in\mathscr{B}_b(\mathscr{C})$.

(4) *There exist two constants $t_0, C>0$ such that*

$$\|P_t^\xi-P_t^\eta\|_{var}^2\le C\|\xi-\eta\|_\infty^2 e^{-\lambda t},\quad t\ge t_0,$$

where $\|\cdot\|_{var}$ is the total variational norm and P_t^ξ stands for the distribution of X_t^ξ for $(t,\xi)\in[0,\infty)\times\mathscr{C}$.

Remark 1.1 (1) We remark that an invariant probability measure μ of P_t must be shift-invariant provided; that is, letting $\phi_\theta(\xi)=\xi(\theta),\theta\in[-r_0,0]$, we have

$$\mu_\theta:=\mu\circ\phi_\theta^{-1}=\mu_0,\quad \theta\in[-r_0,0].$$

In fact, if μ is the law of $X_0=\xi$ which is independent of $(B(t))_{t\ge0}$, by [1, Lemma 1.1.9, p.14], the independence of $\xi\in\mathscr{C}$ and $\{B(t)\}_{t\ge0}$ and the double law of conditional expectation, one has

$$\pi(f)=\int_{\mathscr{C}}\mathbb{E}f(X_t^\eta)\pi(\mathrm{d}\eta)=\mathbb{E}(\mathbb{E}(f(X_t^\xi))|\mathscr{F}_0)=\mathbb{E}(f(X_t^\xi)),\quad t\ge0,\quad f\in\mathscr{B}_b(\mathscr{C}).$$

Then, $X_{-\theta}$ has the law μ for any $\theta\in[-r_0,0]$, so that $X_0(\theta)$ has the same distribution as $X_{-\theta}(\theta)=X_0(0)$; that is, $\mu_\theta=\mu_0$. Moreover, since the equation is non-degenerate, for any $t>0$, the law of $X(t)$ has a strictly positive density with respect to the Lebesgue measure (see, e.g., [15]). So, $\mu_\theta(\mathrm{d}x)=\rho(x)\mathrm{d}x$ holds for some measurable function $\rho>0$ on $\mathbb{R}^d$ and all $\theta\in[-r_0,0]$.

(2) It is well known that when P_t is symmetric in $L^2(\mu)$, the L^2-compactness of P_t for some $t>0$ implies the same property for all $t>0$ (see, e.g., [21, Theorem 0.3.9, p.13]). This assertion is wrong in the non-symmetric setting. In the present framework, P_t is not uniformly integrable (hence, non-compact) on $L^2(\mu)$ for $t\in[0,r_0]$, since, according to (1), $\mu_{-r_0}=\mu_{t-r_0}$ has full support on $\mathbb{R}^d$, and

$$P_tf(\xi)=\mathbb{E}f(X_t^\xi)=g(\xi(t-r_0)),\quad \xi\in\mathscr{C},t\in(0,r_0]$$

holds for $f(\xi) := g(\xi(-r_0)), g \in \mathscr{B}_b(\mathbb{R}^d)$. Therefore, Theorem 1.1 provides a class of Markov semigroups which are compact for large t but not uniformly integrable (hence, non-compact) for small $t \in (0, r_0]$. Moreover, when $r_0 = 0$, assertions in Theorem 1.1 reduce to the corresponding well known ones for SDEs without memory.

In applications, the following consequence of Theorem 1.1 is more convenient to use.

Corollary 1.2. *Let $k_1, k_2 > 0$ be two constants such that*

$$\langle a(x) - a(y), x - y\rangle \le -k_1|x-y|^2, \quad x, y \in \mathbb{R}^d, \tag{1.3}$$

$$|b(\xi) - b(\eta)| \le k_2\|\xi - \eta\|_\infty, \quad \xi, \eta \in \mathscr{C}. \tag{1.4}$$

If

$$k_2^2 \le \frac{2(\sqrt{k_1^2r_0^2+1}-1)}{r_0^2}\exp\Big[\sqrt{k_1^2r_0^2+1}-1-k_1r_0\Big], \tag{1.5}$$

then all assertions in Theorem 1.1 *hold for*

$$\lambda := \frac{r_0}{k_1r_0 - 1 + \sqrt{k_1^2r_0^2+1}}\bigg(\frac{2(\sqrt{k_1^2r_0^2+1}-1)}{r_0^2} - k_2^2\exp\Big[1+k_1r_0-\sqrt{k_1^2r_0^2+1}\Big]\bigg) > 0.$$

Next, we consider a semi-linear model which may not satisfy conditions in Theorem 1.1 and Corollary 1.2. Let $\mathbb{R}^d\otimes\mathbb{R}^d$ be the set of all real $d\times d$-matrices, and let ν be an $\mathbb{R}^d\otimes\mathbb{R}^d$-valued finite signed measure on $[-r_0, 0]$; that is, $\nu = (\nu_{ij})_{1\le i,j\le d}$, where every ν_{ij} is a finite signed measure on $[-r_0, 0]$. Consider the following semi-linear FSDE

$$\mathrm{d}X(t) = \Big\{\int_{-r_0}^0 \nu(\mathrm{d}\theta)X(t+\theta) + b(X_t)\Big\}\mathrm{d}t + \sigma\mathrm{d}B(t), \quad t > 0, \quad X_0 = \xi, \tag{1.6}$$

where $\sigma, B(t)$ are as in (1.1), and b satisfies (1.4). Let

$$\lambda_0 = \sup\bigg\{\mathrm{Re}(\lambda) : \ \lambda \in \mathbb{C}, \det\Big(\lambda I_{d\times d} - \int_{-r_0}^0 \mathrm{e}^{\lambda s}\nu(\mathrm{d}s)\Big) = 0\bigg\},$$

where $I_{d\times d} \in \mathbb{R}^d\otimes\mathbb{R}^d$ is the unitary matrix.

In particular, when $\nu = A\delta_0$, where $A \in \mathbb{R}^d\otimes\mathbb{R}^d$ and δ_0 is the Dirac measure at point 0, equation (1.6) reduces to the usual semi-linear FSDE:

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

and λ_0 is the largest real part of eigenvalues of A.

Let $\Gamma(0) = I_{d\times d}, \Gamma(\theta) = 0_{d\times d}$ for $\theta \in [-r_0, 0)$, and $\{\Gamma(t)\}_{t\geq 0}$ solve the following equation on $\mathbb{R}^d \otimes \mathbb{R}^d$:

$$\mathrm{d}\Gamma(t) = \bigg(\int_{-r_0}^0 \nu(\mathrm{d}\theta)\Gamma(t+\theta)\bigg)\mathrm{d}t. \tag{1.7}$$

According to [18, Theorem 3.1], the unique strong solution $\{X^\xi(t)\}_{t\geq 0}$ of (1.6) can be represented by

$$\begin{aligned} X^\xi(t) &= \Gamma(t)\xi(0) + \int_{-r_0}^0 \nu(\mathrm{d}\theta)\int_{-r_0}^\theta \Gamma(t+\theta-s)\xi(s)\mathrm{d}s \\ &\quad + \int_0^t \Gamma(t-s)b(X_s^\xi)\mathrm{d}s + \int_0^t \Gamma(t-s)\sigma \mathrm{d}B(s). \end{aligned} \tag{1.8}$$

In what follows, we assume $\lambda_0 < 0$. By [12, Theorem 3.2, p.271], for any $k \in (0, -\lambda_0)$, there exists a constant $c_k > 0$ such that

$$\|\Gamma(t)\| \leq c_k \mathrm{e}^{-kt}, \quad t \geq -r_0, \tag{1.9}$$

where $\|\cdot\|$ denotes the operator norm of the matrix $\cdot$. We remark that the optimal constant c_k is increasing in $k \in (0, -\lambda_0)$. If, in particular, $\nu = A\delta_0$ for a symmetric $d \times d$-matrix A, (1.9) holds for $c_k = 1$ and $k \in (0, -\lambda_0]$. In general, see Proposition 4.1 in the Appendix of the paper for an explicit estimate on c_k.

The second main result in this paper is stated as follows.

Theorem 1.3. *Let P_t be the Markov semigroup associated with the equation* (1.6) *such that ν satisfies $\lambda_0 < 0$ and b satisfies* (1.4)*. If $\lambda := \sup_{k\in(0,-\lambda_0)}(k - c_k k_2 \mathrm{e}^{kr_0}) > 0$, where c_k is in* (1.9)*, then all assertions in Theorem* 1.1 *hold.*

The following corollary follows immediately from Theorem 1.3 since $k_2 = 0$ for $b \equiv 0$, and $c_k = 1$ for $\nu = A\delta_0$ with some symmetric matrix A.

Corollary 1.4. *In the situation of Theorem* 1.3*.*

(1) *If $b \equiv 0$, then all assertions in Theorem* 1.1 *hold for all $\lambda \in (0, -\lambda_0)$.*

(2) *Let $\nu = A\delta_0$ for some symmetric $d \times d$-matrix A with largest eigenvalue $\lambda_0 < 0$. If $\lambda := \sup_{k\in(0,-\lambda_0]}(k - k_2\mathrm{e}^{kr_0}) > 0$, then all assertions in Theorem* 1.1 *hold.*

To conclude this section, let us compare Theorems 1.1 and 1.3. The framework of Theorem 1.1 is more general by the generality of $a(\cdot)$. On the other hand, the following example shows that Theorem 1.3 is not covered by Corollary 1.2, a comparable consequence of Theorem 1.1. Let $r_0 = 1$, $\nu(\cdot) = -I_{d\times d}\mathrm{e}^{-1}\delta_{-1}(\cdot)$, and $b \equiv 0$. Then,

$$\lambda_0 = \sup\big\{\mathrm{Re}(\lambda) : \ \lambda \in \mathbb{C}, \lambda + \mathrm{e}^{-\lambda-1} = 0\big\} = -1 < 0,$$

so that Corollary 1.4 applies for all $\lambda \in (0, -1)$; but Corollary 1.2 does not apply due to $a \equiv 0$.

The next section is devoted to the proofs of Theorem 1.1 and Corollary 1.2, while Theorem 1.3 is proved in Section 3. Finally, in Appendix we present an estimate on c_k in (1.9).

2 Proofs of Theorem 1.1 and Corollary 1.2

Since (1.2) still holds if we replace λ_1 by a smaller positive number, and $\lambda = \sup_{s\in[0,\lambda_1]}\big(s - \lambda_2 e^{r_0 s}\big) > 0$, there exists $\overline{\lambda}_1 \in (0, \lambda_1]$ such that $\lambda = \overline{\lambda}_1 - \lambda_2 e^{r_0\overline{\lambda}_1} > 0$. Hence, by using $\overline{\lambda}_1$ to replace λ_1, without loss of generality we assume that $\lambda = \lambda_1 - \lambda_2 e^{r_0\lambda_1} > 0$.

Lemma 2.1. *If $\lambda > 0$, then there exist two constants $c, \varepsilon > 0$ such that*

$$\sup_{t\geq 0} \mathbb{E}\, e^{\varepsilon\|X_t^\xi\|_\infty^2} \leq e^{c(1+\|\xi\|_\infty^2)}, \quad \xi \in \mathscr{C}.$$

Proof. Since in our proof we need to assume in advance that $\mathbb{E}\, e^{\varepsilon\|X_t^\xi\|_\infty^2} < \infty$ for some $\varepsilon > 0$ and each $t \geq 0$, we adopt an approximation argument. For each integer $n > \|\xi\|_\infty$, let

$$\tau_n = \inf\{t \geq 0 : \|X_t^\xi\|_\infty \geq n\}.$$

Then $\tau_n \uparrow \infty$ as $n \uparrow \infty$. Consider the following FSDE

$$\text{(2.1)}\quad dX^{(n)}(t) = \{a(X^{(n)}(t)) + b(X_t^{(n)})\}1_{[0,\tau_n]}(t)dt - \lambda_1 X^{(n)}(t)1_{(\tau_n,\infty)}(t)dt + \sigma dB(t), \quad t > 0$$

with the initial datum $X_0^{(n)} = \xi$. Then (2.1) has a unique strong solution $\{X^{(n)}(t)\}_{t\geq -r_0}$ (see, e.g., [19, Theorem 2.3]) such that $X_t^{(n)} = X_t^\xi$ for $t \leq \tau_n$. Therefore, for any $t > 0$,

$$\text{(2.2)}\qquad \lim_{n\to\infty} \|X_t^{(n)} - X_t^\xi\|_\infty = 0, \quad a.s.$$

Noting that there exists a constant $C(n) > 0$ such that

$$\big\langle \{a(X^{(n)}(t)) + b(X_t^{(n)})\}1_{[0,\tau_n]}(t) - \lambda_1 X^{(n)}(t)1_{(\tau_n,\infty)}(t),\, X^{(n)}(t)\big\rangle \leq C(n) - \lambda_1|X^{(n)}(t)|^2,$$

and utilizing Itô's formula, we infer that

$$\text{(2.3)}\qquad d|X^{(n)}(t)|^2 \leq 2\{C(n) - \lambda_1|X^{(n)}(t)|^2\}dt + 2\langle X^{(n)}(t), \sigma dB(t)\rangle.$$

For each integer $m > \|\xi\|_\infty$, define

$$\widetilde{\tau}_m = \inf\{t \geq 0 : |X^{(n)}(t)| \geq m\}.$$

Then $\widetilde{\tau}_m \uparrow \infty$ as $m \uparrow \infty$. In view of (2.3), for any $\alpha > 0$, we arrive at

$$\begin{aligned}\mathbb{E}\exp\Big(\alpha\int_0^{t\wedge\widetilde{\tau}_m}|X^{(n)}(s)|^2 ds\Big) &\leq \mathbb{E}\exp\Big(\frac{\alpha(|\xi(0)|^2 + C(n)t)}{\lambda_1} + \frac{\alpha}{\lambda_1}\int_0^{t\wedge\widetilde{\tau}_m}\langle X^{(n)}(s), \sigma dB(s)\rangle\Big)\\ &\leq \exp\Big(\frac{\alpha(|\xi(0)|^2 + C(n)t)}{\lambda_1}\Big)\Big(\mathbb{E}\exp\Big(\frac{2\alpha^2\|\sigma\|^2}{\lambda_1^2}\int_0^{t\wedge\widetilde{\tau}_m}|X^{(n)}(s)|^2 ds\Big)\Big)^{1/2},\end{aligned}$$

where in the last step we have used the fact that

$$\text{(2.4)}\qquad \mathbb{E}e^{N(s)} \leq (\mathbb{E}e^{2\langle N\rangle(s)})^{1/2}$$

for a $\mathbb{P}$-martingale $N(s)$. Choosing $\alpha = \frac{\lambda_1^2}{2\|\sigma\|^2}$, taking $m \to \infty$, and applying Fatou's lemma gives

$$\mathbb{E} \exp\Big(\alpha \int_0^t |X^{(n)}(s)|^2 \mathrm{d}s\Big) \le \exp\Big(\frac{2\alpha(|\xi(0)|^2 + C(n)t)}{\lambda_1}\Big). \tag{2.5}$$

Also, by the Itô formula, for any $\beta \le \frac{\sqrt{\alpha/2}}{2\|\sigma\|}$ we deduce from (2.3)-(2.5) that

$$\begin{aligned} \mathbb{E}\mathrm{e}^{\beta|X^{(n)}(t)|^2} &\le \mathbb{E}\exp\Big(\beta(|\xi(0)|^2 + 2C(n)t) + 2\beta\int_0^t \langle X^{(n)}(s), \sigma \mathrm{d}B(s)\rangle\Big) \\ &\le \exp(\beta(|\xi(0)|^2 + 2C(n)t))\Big(\mathbb{E}\exp\Big(8\beta^2\|\sigma\|^2\int_0^t |X^{(n)}(s)|^2\mathrm{d}s\Big)\Big)^{1/2} \\ &\le \exp(2(\beta + \alpha/\lambda_1)(|\xi(0)|^2 + C(n)t)). \end{aligned} \tag{2.6}$$

Next, by virtue of (2.4)-(2.6), and Hölder's inequality, for $\varepsilon_0 < \frac{1}{2}\Big(\beta \wedge \frac{\sqrt{\alpha/2}}{\|\sigma\|}\Big)$ we derive that

$$\begin{aligned} &\mathbb{E}\Big(\sup_{t-r_0\le s\le t} \mathrm{e}^{\varepsilon_0|X^{(n)}(s)|^2}\Big) \\ &= \mathbb{E}\Big(\sup_{(t-r_0)^+\le s\le t} \exp(\varepsilon_0(|X^{(n)}((t-r_0)^+)|^2 + \|\xi\|_\infty^2 + 2C(n)r_0)) + 2\varepsilon\int_{(t-r_0)^+}^s \langle X^{(n)}(u), \sigma\mathrm{d}B(u)\rangle\Big) \\ &\le \mathrm{e}\mathbb{E}\Big(\exp(\varepsilon_0(|X^{(n)}((t-r_0)^+)|^2 + \|\xi\|_\infty^2 + 2C(n)r_0)) + 2\varepsilon_0\int_{(t-r_0)^+}^t \langle X^{(n)}(s), \sigma\mathrm{d}B(s)\rangle\Big) \\ &\le \mathrm{e}(\mathbb{E}(\exp(2\varepsilon_0(|X^{(n)}((t-r_0)^+)|^2 + \|\xi\|_\infty^2 + 2C(n)r_0)))^{1/2} \\ &\quad \times \Big(\mathbb{E}\Big(\exp(8\varepsilon_0^2\|\sigma\|^2\int_{(t-r_0)^+}^t |X^{(n)}(s)|^2\mathrm{d}s\Big)\Big)^{1/2} \\ &< \infty, \end{aligned}$$

where $(t-r_0)^+ := (t-r_0) \vee 0$, and, in the first inequality, we have applied the fact that

$$\mathbb{E}\Big(\sup_{r\in[0,t]} \mathrm{e}^{M(r)}\Big) \le \mathrm{e}\mathbb{E}\mathrm{e}^{M(t)}$$

for a $\mathbb{P}$-submartingale $M(r)$. Consequently,

$$\mathbb{E}\mathrm{e}^{\varepsilon_0\|X_t^{(n)}\|_\infty^2} < \infty, \quad n \ge 1,\ t \ge 0 \tag{2.7}$$

holds for some constant $\varepsilon_0 > 0$.

Next, let $\xi_0(\theta) \equiv 0, \theta \in [-r_0, 0]$. By (1.2), we have

$$\begin{aligned} 2\langle a(\xi(0)) + b(\xi), \xi(0)\rangle &\le 2\langle a(\xi(0)) + b(\xi) - a(0) - b(\xi_0),\, \xi(0)\rangle + |a(0) + b(\xi_0)| \cdot |\xi(0)| \\ &\le c_0 + \lambda_2\|\xi\|_\infty^2 - \lambda_1'|\xi(0)|^2, \quad \xi \in \mathscr{C} \end{aligned}$$

for some constants $c_0 > 0$ and $\lambda_1' > 0$ such that $\lambda' := \lambda_1' - \lambda_2 e^{r_0\lambda_1'} > 0$ due to $\lambda > 0$. So, by Itô's formula,

$$\mathrm{d}|X^{(n)}(t)|^2 \le \big\{c_1 + \lambda_2\|X_t^{(n)}\|_\infty^2 - \lambda_1'|X^{(n)}(t)|^2\big\}\mathrm{d}t + \mathrm{d}M(t)$$

holds for $c_1 := c_0 + \|\sigma\|_{HS}^2$ and $\mathrm{d}M(t) := 2\langle\sigma \mathrm{d}B(t), X^{(n)}(t)\rangle$. This implies

$$e^{\lambda_1' t}|X^{(n)}(t)|^2 \le |\xi(0)|^2 + \int_0^t e^{\lambda_1' s}(c_1 + \lambda_2\|X_s^{(n)}\|_\infty^2)\mathrm{d}s + \int_0^t e^{\lambda_1' s}\mathrm{d}M(s).$$

Let $N(t) = \sup_{s\in[0,t]}\int_0^s e^{\lambda_1' r}\mathrm{d}M(r)$. We obtain

$$\begin{aligned}
e^{\lambda_1' t}\|X_t^{(n)}\|_\infty^2 &\le e^{r_0\lambda_1'}\sup_{\theta\in[-r_0,0]} e^{\lambda_1'(t+\theta)}|X^{(n)}(t+\theta)|^2\\
&\le e^{\lambda_1' r_0}\|\xi\|_\infty^2 + \int_0^t e^{\lambda_1'(s+r_0)}(c_1 + \lambda_2\|X_s^{(n)}\|_\infty^2)\mathrm{d}s + e^{\lambda_1' r_0}N(t)\\
&\le c_2(1+\|\xi\|_\infty^2)e^{\lambda_1' t} + e^{\lambda_1' r_0}N(t) + \lambda_2 e^{\lambda_1' r_0}\int_0^t e^{\lambda_1' s}\|X_s^{(n)}\|_\infty^2\mathrm{d}s
\end{aligned}$$

for some constant $c_2 > 0$. By Gronwall's inequality, one has

$$\begin{aligned}
e^{\lambda_1' t}\|X_t^{(n)}\|_\infty^2 \le & c_2(1+\|\xi\|_\infty^2)e^{\lambda_1' t} + e^{\lambda_1' r_0}N(t)\\
&+ \lambda_2 e^{\lambda_1' r_0}\int_0^t \big\{c_2(1+\|\xi\|_\infty^2)e^{\lambda_1' s} + e^{\lambda_1' r_0}N(s)\big\}\exp\big[\lambda_2 e^{\lambda_1' r_0}(t-s)\big]\mathrm{d}s.
\end{aligned}$$

Recalling that $\lambda' = \lambda_1' - \lambda_2 e^{\lambda_1' r_0} > 0$, we arrive at

$$\begin{aligned}
\|X_t^{(n)}\|_\infty^2 &\le c_2(1+\|\xi\|_\infty^2) + e^{\lambda_1'(r_0-t)}N(t)\\
&\quad + \lambda_2 e^{\lambda_1' r_0}\int_0^t \big\{c_2(1+\|\xi\|_\infty^2) + e^{\lambda_1'(r_0-s)}N(s)\big\}e^{-\lambda'(t-s)}\mathrm{d}s\\
&\le c_3\big(1+\|\xi\|_\infty^2 + e^{-\lambda_1' t}N(t)\big) + c_3\int_0^t e^{-\lambda_1' s-\lambda'(t-s)}N(s)\mathrm{d}s
\end{aligned}$$

for some constant $c_3 > 0$. Therefore, for any $\varepsilon \in (0,1)$,

$$\mathbb{E}\, e^{\varepsilon\|X_t^{(n)}\|_\infty^2} \le e^{c_3(1+\|\xi\|_\infty^2)}\sqrt{I_1\times I_2} \tag{2.8}$$

holds for

$$\begin{aligned}
I_1 &:= \mathbb{E}\exp\left[2c_3\varepsilon\int_0^t e^{-\lambda_1' s-\lambda'(t-s)}N(s)\mathrm{d}s\right],\\
I_2 &:= \mathbb{E}\exp\big[2c_3\varepsilon e^{-\lambda_1' t}N(t)\big].
\end{aligned}$$

To finish the proof, below we estimate I_1 and I_2, respectively.

(a) Estimate on I_1. Note that

$$I_1 = \mathbb{E}\exp\left[\frac{2c_3(1-\mathrm{e}^{-\lambda' t})\varepsilon}{\lambda'}\int_0^t \mathrm{e}^{-\lambda_1' s}N(s)\nu_0(\mathrm{d}s)\right],$$

where

$$\nu_0(\mathrm{d}s) := \frac{\lambda'}{1-\mathrm{e}^{-\lambda' t}}\mathrm{e}^{-\lambda'(t-s)}\mathrm{d}s, \quad [0,t].$$

To avoid the singularity of the reference probability measure $\nu_0(\cdot)$ above whenever $t \to 0$, we extend the integral to the larger interval $[-r_0, t]$. Define

$$\nu(\mathrm{d}s) = \frac{\lambda' \mathrm{e}^{\lambda' r_0}}{\mathrm{e}^{\lambda' r_0} - \mathrm{e}^{-\lambda' t}}\mathrm{e}^{-\lambda'(t-s)}\mathrm{d}s \ \text{ on } [-r_0, t],$$

Letting $N(s) = 0$ for $s \le 0$, and applying Jensen's inequality for the probability measure $\nu(\mathrm{d}s)$, we have

$$\begin{aligned}
&\exp\left[4c_3\varepsilon\int_0^t \mathrm{e}^{-\lambda_1' s - \lambda'(t-s)}N(s)\mathrm{d}s\right] \\
&= \exp\left[\frac{4c_3\varepsilon(\mathrm{e}^{\lambda' r_0} - \mathrm{e}^{-\lambda' t})}{\lambda' \mathrm{e}^{\lambda' r_0}}\int_{-r_0}^t \mathrm{e}^{-\lambda_1' s}N(s)\nu(\mathrm{d}s)\right] \\
&\le \int_{-r_0}^t \exp\left[\frac{4c_3\varepsilon(\mathrm{e}^{\lambda' r_0} - \mathrm{e}^{-\lambda' t})}{\lambda' \mathrm{e}^{\lambda' r_0}}\mathrm{e}^{-\lambda_1' s}N(s)\right]\nu(\mathrm{d}s) \\
&\le \frac{\lambda' \mathrm{e}^{\lambda' r_0}}{\mathrm{e}^{\lambda' r_0} - 1}\int_{-r_0}^t \exp\left[\frac{4c_3\varepsilon}{\lambda'}\mathrm{e}^{-\lambda_1' s}N(s)\right]\mathrm{e}^{-\lambda'(t-s)}\mathrm{d}s.
\end{aligned}$$

So, by Jensen's inequality and the Burkhold-Davis-Gundy inequality, there exist constants $c_4, c_5 > 0$ such that

$$\begin{aligned}
I_1^2 &\le \mathbb{E}\exp\left[4c_3\varepsilon\int_0^t \mathrm{e}^{-\lambda_1' s - \lambda'(t-s)}N(s)\mathrm{d}s\right] \\
&\le \frac{\lambda' \mathrm{e}^{\lambda' r_0}}{\mathrm{e}^{\lambda' r_0} - 1}\int_{-r_0}^t \mathrm{e}^{-\lambda'(t-s)}\mathbb{E}\exp\left[\frac{4c_3\varepsilon}{\lambda'}\mathrm{e}^{-\lambda_1' s}N(s)\right]\mathrm{d}s \\
&\le c_4\int_{-r_0}^t \mathrm{e}^{-\lambda'(t-s)}\left(\mathbb{E}\exp\left[c_4\varepsilon^2\mathrm{e}^{-2\lambda_1' s}\int_0^s \mathrm{e}^{2\lambda_1' u}\|X_u^{(n)}\|_\infty^2\,\mathrm{d}u\right]\right)^{\frac{1}{2}}\mathrm{d}s \\
&\le c_5\Big(1 + \mathbb{E}\int_{-r_0}^t \mathrm{e}^{-\lambda'(t-s)}\exp\left[c_5\varepsilon^2\int_{-r_0}^s \mathrm{e}^{-2\lambda_1'(s-u)}\|X_u^{(n)}\|_\infty^2\,\mathrm{d}u\right]\mathrm{d}s\Big),
\end{aligned}$$

where we set $X_s^{(n)} = \xi$ for $s \le 0$.

Now, using Jensen's inequality as above for the probability measure

$$\frac{2\lambda_1' \mathrm{e}^{\lambda_1' r_0}}{\mathrm{e}^{2\lambda_1' r_0} - \mathrm{e}^{-2\lambda_1' s}}\mathrm{e}^{-2\lambda_1'(s-u)}\mathrm{d}u \ \text{ on } [-r_0, s],$$

we arrive at

$$\begin{aligned} I_1^2 &\le c_6\Big(1+\mathbb{E}\int_{-r_0}^t \mathrm{e}^{-\lambda'(t-s)}\mathrm{d}s\int_{-r_0}^s \mathrm{e}^{c_6\varepsilon^2\|X_u^{(n)}\|_\infty^2-2\lambda_1'(s-u)}\mathrm{d}u\Big)\\ &= c_6\Big(1+\int_{-r_0}^t \mathbb{E}\,\mathrm{e}^{c_6\varepsilon^2\|X_u^{(n)}\|_\infty^2}\mathrm{d}u\int_u^t \mathrm{e}^{-\lambda'(t-s)-2\lambda_1'(s-u)}\mathrm{d}s\Big)\end{aligned}$$

for some constant $c_6>0$. Since $\lambda_1'\ge\lambda'>0$, we have

$$-2\lambda_1'(s-u)-\lambda'(t-s)\le-\lambda'(t-u)-\lambda_1'(s-u),$$

so that this implies

$$\begin{aligned} I_1^2 &\le c_6\Big(1+\int_{-r_0}^t \mathrm{e}^{-\lambda'(t-u)}\mathbb{E}\,\mathrm{e}^{c_6\varepsilon^2\|X_u^{(n)}\|_\infty^2}\mathrm{d}u\int_u^t \mathrm{e}^{-\lambda_1'(s-u)}\mathrm{d}s\Big)\\ &\le c_6\Big(1+\frac{1}{\lambda_1'}\int_{-r_0}^t \mathrm{e}^{-\lambda'(t-u)}\mathbb{E}\,\mathrm{e}^{c_6\varepsilon^2\|X_u^{(n)}\|_\infty^2}\,\mathrm{d}u\Big).\end{aligned} \tag{2.9}$$

(b) Estimate on I_2. A shown in (a), by the Burkhold-Davis-Gundy inequality and using Jensen's inequality for the probability measure

$$\frac{2\lambda_1'\mathrm{e}^{\lambda_1'r_0}}{\mathrm{e}^{2\lambda_1'r_0}-\mathrm{e}^{-2\lambda_1't}}\mathrm{e}^{-2\lambda_1'(t-s)}\mathrm{d}s\ \text{ on } [-r_0,t],$$

we conclude that

$$\begin{aligned} I_2^2 &\le c_7\Big(1+\mathbb{E}\exp\Big[c_7\varepsilon^2\mathrm{e}^{-2\lambda_1't}\int_{-r_0}^t \mathrm{e}^{2\lambda_1's}\|X_s^{(n)}\|_\infty^2\mathrm{d}s\Big]\Big)\\ &\le c_8\Big(1+\mathbb{E}\int_{-r_0}^t \mathrm{e}^{c_8\varepsilon^2\|X_s^{(n)}\|_\infty^2}\mathrm{e}^{-2\lambda_1'(t-s)}\mathrm{d}s\Big)\end{aligned} \tag{2.10}$$

holds for some constants $c_7,c_8>0$.

Now, combining (2.8), (2.9) with (2.10), and taking $\varepsilon=\varepsilon_0\wedge\frac{1}{c_6\vee c_8}$, we arrive at

$$\begin{aligned} &\mathbb{E}\mathrm{e}^{\varepsilon\|X_t^{(n)}\|_\infty^2}\le \mathrm{e}^{c_9(1+\|\xi\|_\infty^2)}\Big(1+\int_{-r_0}^t\big(\mathbb{E}\mathrm{e}^{\varepsilon\|X_s^{(n)}\|_\infty^2}\big)\mathrm{e}^{-\lambda'(t-s)}\mathrm{d}s\Big)^{\frac12}\\ &\le \mathrm{e}^{c_{10}(1+\|\xi\|_\infty^2)}+\frac{\lambda'}{2}\int_{-r_0}^t\big(\mathbb{E}\mathrm{e}^{\varepsilon\|X_s^{(n)}\|_\infty^2}\big)\mathrm{e}^{-\lambda'(t-s)}\mathrm{d}s\end{aligned}$$

for some constants $c_9,c_{10}>0$. Equivalently,

$$\mathrm{e}^{\lambda't}\mathbb{E}\mathrm{e}^{\varepsilon\|X_t^{(n)}\|_\infty^2}\le \mathrm{e}^{c_9(1+\|\xi\|_\infty^2)+\lambda't}+\frac{\lambda'}{2}\int_{-r_0}^t\big(\mathbb{E}\mathrm{e}^{\varepsilon\|X_s^{(n)}\|_\infty^2}\big)\mathrm{e}^{\lambda's}\mathrm{d}s.$$

By (2.7) and $\varepsilon\le\varepsilon_0$, we see that

$$\mathbb{E}\mathrm{e}^{\varepsilon\|X_t^{(n)}\|_\infty^2}<\infty,\quad t\ge0.$$

Then, by Gronwall's inequality,

$$\mathrm{e}^{\lambda' t}\mathbb{E}\mathrm{e}^{\varepsilon\|X_t^{(n)}\|_\infty^2} \le \mathrm{e}^{c_{10}(1+\|\xi\|_\infty^2)+\lambda' t} + \frac{\lambda'}{2}\int_{-r_0}^t \mathrm{e}^{c_{10}(1+\|\xi\|_\infty^2)+\lambda' s+\frac{\lambda'}{2}(t-s)}\mathrm{d}s.$$

Therefore,

$$\mathbb{E}\mathrm{e}^{\varepsilon\|X_t^{(n)}\|_\infty^2} \le \mathrm{e}^{c_{10}(1+\|\xi\|_\infty^2)} + \frac{\lambda'}{2}\int_{-r_0}^t \mathrm{e}^{c_{10}(1+\|\xi\|_\infty^2)-\frac{\lambda'}{2}(t-s)}\mathrm{d}s \le \mathrm{e}^{c(1+\|\xi\|_\infty^2)}$$

for some constant $c > 0$. According to (2.2), the proof is finished by applying Fatou's lemma. □

Lemma 2.2. *For any $t \ge 0$ and $\xi, \eta \in \mathscr{C}$, $\|X_t^\xi - X_t^\eta\|_\infty^2 \le \|\xi - \eta\|_\infty^2 \mathrm{e}^{\lambda_1 r_0 - \lambda t}$.*

Proof. By Itô's formula, we have

$$\mathrm{d}|X^\xi(t) - X^\eta(t)|^2 \le \big(\lambda_2\|X_t^\xi - X_t^\eta\|_\infty^2 - \lambda_1|X^\xi(t) - X^\eta(t)|^2\big)\mathrm{d}t.$$

Then

$$\mathrm{e}^{\lambda_1 t}|X^\xi(t) - X^\eta(t)|^2 \le |\xi(0) - \eta(0)|^2 + \lambda_2\int_0^t \mathrm{e}^{\lambda_1 s}\|X_s^\xi - X_s^\eta\|_\infty^2\mathrm{d}s.$$

So,

$$\mathrm{e}^{\lambda_1 t}\|X_t^\xi - X_t^\eta\|_\infty^2 \le \mathrm{e}^{r_0\lambda_1}\|\xi - \eta\|_\infty^2 + \lambda_2\mathrm{e}^{r_0\lambda_1}\int_0^t \mathrm{e}^{\lambda_1 s}\|X_s^\xi - X_s^\eta\|_\infty^2\mathrm{d}s.$$

Therefore, the proof is finished by Gronwall's inequality since we have assumed that $\lambda = \lambda_1 - \lambda_2\mathrm{e}^{r_0\lambda_1}$. □

Now, we introduce the dimension-free Harnack inequality in the sense of [20]. We are referred to [4, 9, 25] for more results on the Harnack inequality of FSDEs. Since results in these papers do not directly imply the following Lemma 2.3, we include a simple proof using coupling by change of measure introduced in [2]. By (1.2) and the Lipschitz property of b, (1.4) holds for some $k_2 \ge 0$ and

$$2\langle a(x) - a(y), x - y\rangle \le 2\langle b(\xi_y) - b(\xi_x), x - y\rangle + (\lambda_2 - \lambda_1)|x - y|^2 \le -k_1|x - y|^2, \quad x, y \in \mathbb{R}^d$$

holds for some constant $k_1 \in \mathbb{R}$ as required in Lemma 2.3, where $\xi_x(\theta) = x, \xi_y(\theta) = y$ for $\theta \in [-r_0, 0]$.

Lemma 2.3. *Let (1.3) and (1.4) hold for some constants $k_1 \in \mathbb{R}$ and $k_2 \ge 0$. Then, for any $p > 1, \delta > 0$, positive $f \in \mathscr{B}_b(\mathscr{C})$, and $\xi, \eta \in \mathscr{C}$,*

$$\begin{aligned}\big(P_{t+r_0}f(\xi)\big)^p \le &\big(P_{t+r_0}f^p(\eta)\big)\exp\bigg[\frac{p^2\|\sigma^{-1}\|^2(1+\delta)}{2(p-1)}\bigg\{\frac{2k_1|\xi(0)-\eta(0)|^2}{\mathrm{e}^{2k_1 t}-1}\\ &+\frac{k_2^2}{\delta}\Big(r_0\|\xi-\eta\|_\infty^2 + \frac{|\xi(0)-\eta(0)|^2(\mathrm{e}^{4k_1 t}-1-4k_1 t\mathrm{e}^{2k_1 t})}{2k_1(\mathrm{e}^{2k_1 t}-1)^2}\Big)\bigg\}\bigg].\end{aligned}$$

Proof. Let $X_s = X_s^\xi$ and $Y(s)$ solve the equation

$$\mathrm{d}Y(s) = \left(a(Y(s)) + b(X_s) + g(s)1_{[0,\tau)}(s)\cdot\frac{X(s)-Y(s)}{|X(s)-Y(s)|}\right)\mathrm{d}s + \sigma\mathrm{d}B(s), \quad Y_0 = \eta, \tag{2.11}$$

where

$$\tau := \inf\{s \ge 0 : X(s) = Y(s)\}$$

is the coupling time and $g \in C([0,\infty))$ is to be determined. It is easy to see that this equation has a unique solution up to the coupling time τ. Letting $Y(s) = X(s)$ for $s \ge \tau$, we obtain a solution $Y(s)$ for all $s \ge 0$. We will then choose g such that $\tau \le t$, i.e., $X_{t+r_0} = Y_{t+r_0}$. Obviously, we have

$$\mathrm{d}|X(s)-Y(s)| \le -\big\{k_1|X(s)-Y(s)| + g(s)\big\}\mathrm{d}s, \quad s < \tau.$$

Then

$$|X(s)-Y(s)| \le |\xi(0)-\eta(0)|\mathrm{e}^{-k_1 s} - \mathrm{e}^{-k_1 s}\int_0^s \mathrm{e}^{k_1 r} g(r)\mathrm{d}r, \quad s \le \tau. \tag{2.12}$$

In (2.12), take

$$g(s) = \frac{|\xi(0)-\eta(0)|\mathrm{e}^{k_1 s}}{\int_0^t \mathrm{e}^{2k_1 s}\mathrm{d}s}, \quad s \in [0,t]. \tag{2.13}$$

If $t < \tau$, we infer from (2.12) and (2.13) that

$$|X(s)-Y(s)| \le \frac{|\xi(0)-\eta(0)|(\mathrm{e}^{2k_1 t - k_1 s} - \mathrm{e}^{k_1 s})}{\mathrm{e}^{2k_1 t}-1}, \quad 0 \le s \le t. \tag{2.14}$$

This implies $X(t) = Y(t)$ and hence, it is contradictory to $t < \tau$. Hence we arrive at $\tau \le t$ and $X_{t+r_0} = Y_{t+r_0}$ as required. Moreover, note that (2.14) still holds for $t \ge \tau$. Now, let

$$h(s) = \sigma^{-1}\Big\{1_{[0,\tau)}g(s)\frac{X(s)-Y(s)}{|X(s)-Y(s)|} + b(X_s) - b(Y_s)\Big\}.$$

Noting that (2.14) implies

$$\|X_s - Y_s\|_\infty^2 \le 1_{[0,r_0]}(s)\|\xi-\eta\|_\infty^2 + 1_{(r_0,r_0+t]}(s)\frac{(\mathrm{e}^{2k_1 t - k_1(s-r_0)} - \mathrm{e}^{k_1(s-r_0)})^2|\xi(0)-\eta(0)|^2}{(\mathrm{e}^{2k_1 t}-1)^2},$$

we obtain from (1.4) and (2.13) that

$$\begin{aligned}
|h(s)|^2 &\le \|\sigma^{-1}\|^2(1_{[0,t]}(s)g(s) + k_2\|X_s - Y_s\|_\infty)^2\\
&\le \|\sigma^{-1}\|^2(1_{[0,t]}(s)(1+\delta)g(s)^2 + (1+\delta^{-1})k_2^2\|X_s-Y_s\|_\infty^2)\\
&\le 1_{[0,t]}(s)\frac{4k_1^2\mathrm{e}^{2k_1 s}\|\sigma^{-1}\|^2(1+\delta)|\xi(0)-\eta(0)|^2}{(\mathrm{e}^{2k_1 t}-1)^2}\\
&\quad + 1_{(r_0,r_0+t]}(s)\frac{\mathrm{e}^{2k_1 s}\|\sigma^{-1}\|^2(1+\delta)|\xi(0)-\eta(0)|^2k_2^2(\mathrm{e}^{2k_1 t - k_1(s-r_0)} - \mathrm{e}^{k_1(s-r_0)})^2}{\delta(\mathrm{e}^{2k_1 t}-1)^2}\\
&\quad + 1_{[0,r_0]}(s)\|\sigma^{-1}\|^2k_2^2(1+\delta^{-1})\|\xi-\eta\|_\infty^2 =: \gamma(s).
\end{aligned}$$

Then we have

$$\begin{aligned} \int_0^{t+r_0} |h(s)|^2 \mathrm{d}s &\le \int_0^{t+r_0} \gamma(s)\mathrm{d}s \\ &\le \exp\left[\|\sigma^{-1}\|^2(1+\delta)\left(\frac{(\mathrm{e}^{4k_1t}-1-4k_1t\mathrm{e}^{2k_1t})k_2^2|\xi(0)-\eta(0)|^2}{2k_1\delta(\mathrm{e}^{2k_1t}-1)^2}\right.\right. \\ &\qquad\left.\left. + \frac{r_0k_2^2\|\xi-\eta\|_\infty^2}{\delta} + \frac{2k_1|\xi(0)-\eta(0)|^2}{\mathrm{e}^{2k_1t}-1}\right)\right]. \end{aligned} \tag{2.15}$$

Hence we arrive at

$$\mathbb{E}\exp\left[\frac{1}{2}\int_0^{t+r_0}|h(s)|^2\mathrm{d}s\right] < \infty.$$

As a result, Novikov's condition holds so that, by the Girsanov theorem, $\{\widetilde{B}(s)\}_{t\in[0,t+r_0]}$ is a Brownian motion under the weighted probability measure $\mathrm{d}\mathbb{Q} := R\mathrm{d}\mathbb{P}$ with

$$R := \exp\left[-\int_0^{t+r_0}\langle h(s), \mathrm{d}B(s)\rangle - \frac{1}{2}\int_0^{t+r_0}|h(s)|^2\mathrm{d}s\right].$$

Observe that (2.11) can be rewritten as

$$\mathrm{d}Y(s) = \big(a(Y(s)) + b(Y_s)\big)\mathrm{d}s + \sigma\mathrm{d}\widetilde{B}(s), \quad Y_0 = \eta, \quad s \in [0, t+r_0].$$

By the weak uniqueness of solution and $X_{t+r_0} = Y_{t+r_0}$, we have

$$P_{t+r_0}f(\eta) = \mathbb{E}[Rf(Y_{t+r_0})] = \mathbb{E}[Rf(X_{t+r_0})]. \tag{2.16}$$

Then, by Jensen's inequality,

$$(P_{t+r_0}f(\eta))^p = \big(\mathbb{E}[Rf(X_{t+r_0})]\big)^p \le (P_{t+r_0}f^p(\xi))\big(\mathbb{E}R^{\frac{p}{p-1}}\big)^{p-1}.$$

Consequently, the desired assertion follows by noting that

$$\begin{aligned} \mathbb{E}R^{\frac{p}{p-1}} &\le \mathrm{e}^{\frac{p^2}{2(p-1)^2}\int_0^{t+r_0}\gamma(s)\mathrm{d}s}\mathbb{E}\mathrm{e}^{-\frac{p}{p-1}\int_0^{t+r_0}\langle h(s),\mathrm{d}B(s)\rangle - \frac{p^2}{2(p-1)^2}\int_0^{t+r_0}|h(s)|^2\mathrm{d}s} \\ &= \mathrm{e}^{\frac{p^2}{2(p-1)^2}\int_0^{t+r_0}\gamma(s)\mathrm{d}s}, \end{aligned} \tag{2.17}$$

and taking (2.15) into consideration. □

Lemma 2.4. *If $\lambda > 0$, then P_t has a unique invariant probability measure μ such that*

$$\lim_{t\to\infty} P_tf(\xi) = \mu(f), \quad f \in C_b(\mathscr{C}),\ \xi \in \mathscr{C}.$$

Proof. Let $\mathscr{P}(\mathscr{C})$ be the set of all probability measures on $\mathscr{C}$. Let W be the L^2-Wasserstein distance on $\mathscr{P}(\mathscr{C})$ induced by the distance $\rho(\xi,\eta) := 1 \wedge \|\xi-\eta\|_\infty$; that is,

$$W(\mu_1,\mu_2) := \inf_{\pi\in\mathscr{C}(\mu_1,\mu_2)} \big(\pi(\rho^2)\big)^{\frac{1}{2}}, \quad \mu_1,\mu_2 \in \mathscr{P}(\mathscr{C}),$$

where $\mathscr{C}(\mu_1,\mu_2)$ is the set of all couplings of μ_1 and μ_2. It is well known that $\mathscr{P}(\mathscr{C})$ is a complete metric space with respect to the distance W, and the convergence in W is equivalent to the weak convergence (see, e.g., [6, Theorems 5.4 and 5.6]). Let P_t^ξ be the law of X_t^ξ. Then it remains to prove the following two assertions:

(i) For any $\xi \in \mathscr{C}$, there exists $\mu_\xi \in \mathscr{P}(\mathscr{C})$ such that $\lim_{t\to\infty} W(P_t^\xi, \mu_\xi) = 0$;

(ii) For any $\xi, \eta \in \mathscr{C}$, $\mu_\xi = \mu_\eta$.

To show (i), it suffices to show that $\{P_t^\xi\}_{t\geq 0}$ is a Cauchy sequence with respect to W. To this end, for any $t_2 > t_1 > 0$, we consider the following FSDEs

$$\begin{aligned} &\mathrm{d}X(t) = \{a(X(t)) + b(X_t)\}\mathrm{d}t + \sigma \mathrm{d}B(t), \quad X_0 = \xi, t \in [0, t_2],\\ &\mathrm{d}\overline{X}(t) = \{a(\overline{X}(t)) + b(\overline{X}_t)\}\mathrm{d}t + \sigma \mathrm{d}B(t), \quad \overline{X}_{t_2-t_1} = \xi, t \in [t_2 - t_1, t_2]. \end{aligned}$$

Then the laws of X_{t_2} and $\overline{X}_{t_2}$ are $P_{t_2}^\xi$ and $P_{t_1}^\xi$, respectively. By (1.2), we have

$$\mathrm{d}|X(t) - \overline{X}(t)|^2 \leq \big(\lambda_2 \|X_t - \overline{X}_t\|_\infty^2 - \lambda_1 |X(t) - \overline{X}(t)|^2\big)\mathrm{d}t, \quad t \in [t_2 - t_1, t_2].$$

As in the proof of Lemma 2.2, this implies

$$\|X_t - \overline{X}_t\|_\infty^2 \leq \mathrm{e}^{\lambda_1 r_0} \|X_{t_2-t_1} - \xi\|_\infty^2 \mathrm{e}^{-\lambda(t+t_1-t_2)}, \quad t \in [t_2 - t_1, t_2].$$

In particular,

$$\|X_{t_2} - \overline{X}_{t_2}\|_\infty^2 \leq \mathrm{e}^{\lambda_1 r_0} \|X_{t_2-t_1} - \xi\|_\infty^2 \mathrm{e}^{-\lambda t_1}.$$

By Lemma 2.1, we have

$$\mathbb{E}\|X_{t_2-t_1} - \xi\|_\infty^2 \leq C := \sup_{t\geq 0} \mathbb{E}\|X_t - \xi\|_\infty^2 < \infty.$$

Then one has

$$\sup_{t_2\in[t_1,\infty)} W(P_{t_1}^\xi, P_{t_2}^\xi) \leq \sup_{t_2\in[t_1,\infty)} \mathbb{E}\{1 \wedge |X_{t_2} - \overline{X}_{t_2}\|_\infty^2\} \leq C\mathrm{e}^{\lambda_1 r_0 - \lambda t_1}$$

which tends to zero as t_1 goes to infinity, so that $\{P_t^\xi\}_{t\geq 0}$ is a Cauchy sequence with respect to W.

Next, (ii) follows by observing that

$$W(\mu_\xi, \mu_\eta) \leq W(P_t^\xi, \mu_\xi) + W(P_t^\eta, \mu_\eta) + W(P_t^\xi, P_t^\eta), \quad \xi, \eta \in \mathscr{C},$$

and taking (i) and Lemma 2.2 into account. □

Proof of Theorem 1.1. (a) We first prove that $\|P_t\|_{2\to 4} < \infty$ holds for large enough $t > 0$. Let $f \in \mathscr{B}_b(\mathscr{C})$ with $\mu(f^2) = 1$. By Lemma 2.3, for any $t_0 > r_0$ there exists a constant $c_0 > 0$ such that

$$(P_{t_0} f(\xi))^2 \leq (P_{t_0} f^2(\eta))\mathrm{e}^{c_0\|\xi-\eta\|_\infty^2}, \quad \xi, \eta \in \mathscr{C}.$$

By the Markov property and Schwartz's inequality,

$$\begin{aligned} |P_{t+t_0} f(\xi)|^2 &= |\mathbb{E}(P_{t_0} f)(X_t^\xi)|^2 \leq \Big(\mathbb{E}\sqrt{(P_{t_0} f^2(X_t^\eta)) \exp[c_0\|X_t^\xi - X_t^\eta\|_\infty^2]}\Big)^2\\ &\leq (\mathbb{E}(P_{t_0} f^2(X_t^\eta))\mathbb{E}\mathrm{e}^{c_0\|X_t^\xi - X_t^\eta\|_\infty^2} = (P_{t+t_0} f^2(\eta))\mathbb{E}\mathrm{e}^{c_0\|X_t^\xi - X_t^\eta\|_\infty^2}. \end{aligned}$$

Combining this with Lemma 2.2, we obtain

$$|P_{t+t_0}f(\xi)|^2 \le (P_{t+t_0}f^2(\eta))\exp\big[c_1 e^{-\lambda t}\|\xi-\eta\|_\infty^2\big].$$

Let $r>0$ such that $\mu(B_r)\ge \frac{1}{2}$, where $B_r := \{\|\cdot\|_\infty < R\}$. Then

$$|P_{t+t_0}f(\xi)|^2 \exp\big[-c_1 e^{-\lambda t}(\|\xi\|_\infty + r)^2\big] \le 2|P_{t+t_0}f(\xi)|^2\int_{B_r}\exp\big[-c_1e^{-\lambda t}\|\xi-\eta\|_\infty^2\big]\mu(d\eta)$$
$$\le 2\int_{\mathscr{C}} P_{t+t_0}f^2(\eta)\mu(d\eta) = 2.$$

Thus,

$$|P_{t+t_0}f(\xi)|^4 \le \exp\big[c_2(1+\|\xi\|_\infty^2 e^{-\lambda t})\big],\quad t\ge 0 \tag{2.18}$$

holds for some constant $c_2>0$. On the other hand, by Lemmas 2.1 and 2.4 we have

$$\mu(N\wedge e^{\varepsilon\|\cdot\|_\infty^2}) = \lim_{t\to\infty}\mathbb{E}(N\wedge e^{\varepsilon\|X_t^0\|_\infty^2}) \le e^c < \infty,\quad N>0$$

for some constant $c>0$. Taking $N\to\infty$, we obtain $\mu(e^{\varepsilon\|\cdot\|_\infty^2})<\infty$. Therefore, (2.18) implies $\|P_{t+t_0}\|_{2\to4}<\infty$ for large enough $t>0$.

(b) By, e.g., [25, Proposition 3.1 (2)], the Harnack inequality implies that P_t has a density with respect to μ for $t>r_0$. Thus, according to [26, Theorem 2.3], the hyperboundedness of P_t proved in (a) implies that P_t is compact in $L^2(\mu)$ for large enough $t>0$. Hence, Theorem 1.1(2) is proved.

(c) To prove Theorem 1.1(3), we let X_t, Y_t and R be given as in the proof of Lemma 2.3. By (2.16) and $P_{t+r_0}f(\xi)=\mathbb{E}f(X_{t+r_0})$, we have

$$|P_{t+r_0}f(\xi)-P_{t+r_0}f(\eta)| \le \mathbb{E}|f(X_{t+t_0})(R-1)| \le \sqrt{(P_{t+r_0}f^2(\xi))\mathbb{E}(R^2-1)}.$$

Take $p=2$ and $t=t_1>0$ such that $\|P_{t_1+r_0}\|_{2\to4}<\infty$ according to (a). By (2.17) there exists a constant $c_1>0$ such that $\mathbb{E}R^2\le e^{c_1\|\xi-\eta\|_\infty^2}$. So,

$$\begin{aligned}|P_{t_1+r_0}f(\xi)-P_{t_1+r_0}f(\eta)|^2 &\le (P_{t_1+r_0}f^2(\xi))(e^{c_1\|\xi-\eta\|_\infty^2}-1)\\ &\le (P_{t_1+r_0}f^2(\xi))c_1\|\xi-\eta\|_\infty^2 e^{c_1\|\xi-\eta\|_\infty^2}.\end{aligned} \tag{2.19}$$

Hence, for any $t>0$,

$$|P_{t+2(t_1+r_0)}f(\xi)-P_{t+2(t_1+r_0)}f(\eta)|^2 \le \big(\mathbb{E}|P_{t_1+r_0}(P_{t_1+r_0}f)(X_t^\xi) - P_{t_1+r_0}(P_{t_1+r_0}f)(X_t^\eta)|\big)^2$$
$$\le \Big(\mathbb{E}\sqrt{(P_{t_1+r_0}(P_{t_1+r_0}f)^2(X_t^\xi))c_1\|X_t^\xi - X_t^\eta\|_\infty^2 e^{c_1\|X_t^\xi-X_t^\eta\|_\infty^2}}\Big)^2$$
$$\le (P_{t+t_1+r_0}(P_{t_1+r_0}f)^2(\xi))\mathbb{E}\Big[c_1\|X_t^\xi-X_t^\eta\|_\infty^2 e^{c_1\|X_t^\xi-X_t^\eta\|_\infty^2}\Big].$$

Combining this with Lemma 2.2, we arrive at

$$\begin{aligned}&|P_{t+2(t_1+r_0)}f(\xi)-P_{t+2(t_1+r_0)}f(\eta)|^2\\ &\le (P_{t+t_1+r_0}(P_{t_1+r_0}f)^2(\xi))c_2e^{-\lambda t}\|\xi-\eta\|_\infty^2\exp\big[c_2e^{-\lambda t}\|\xi-\eta\|_\infty^2\big]\\ &\le (P_{t+t_1+r_0}(P_{t_1+r_0}f)^2(\xi))c_3e^{-\lambda t}\exp\Big[\frac{\varepsilon}{4}\|\xi-\eta\|_\infty^2\Big]\end{aligned}$$

for some constants $c_2, c_3 > 0$ and large enough $t > 0$, where $\varepsilon > 0$ such that $\mu(\mathrm{e}^{\varepsilon\|\cdot\|_\infty^2}) < \infty$ according to (a). So,

$$\begin{aligned}
&2\mu(|P_{t+2(t_1+r_0)}f - \mu(f)|^2) = \int_{\mathscr{C}\times\mathscr{C}} |P_{t+2(t_1+r_0)}f(\xi) - P_{t+2(t_1+r_0)}f(\eta)|^2\mu(\mathrm{d}\xi)\mu(\mathrm{d}\eta)\\
&\le c_3\mathrm{e}^{-\lambda t}\bigg(\int_{\mathscr{C}} \big\{P_{t+t_1+r_0}(P_{t_1+r_0}f)^2\big\}^2(\xi)\mu(\mathrm{d}\xi)\bigg)^{\frac{1}{2}}\bigg(\int_{\mathscr{C}\times\mathscr{C}} \exp\Big[\frac{\varepsilon}{2}\|\xi-\eta\|_\infty^2\Big]\mu(\mathrm{d}\xi)\mu(\mathrm{d}\eta)\bigg)^{\frac{1}{2}}\\
&\le C\mathrm{e}^{-\lambda t}\mu(f^2), \quad t \ge 0
\end{aligned}$$

holds for some constant $C > 0$, since by Jensen's inequality and the fact that μ is P_t-invariant, we have

$$\int_{\mathscr{C}} \big\{P_{t+t_1+r_0}(P_{t_1+r_0}f)^2\big\}^2\mathrm{d}\mu \le \int_{\mathscr{C}} (P_{t_1+r_0}f)^4\mathrm{d}\mu \le \|P_{t_1+r_0}\|_{2\to4}^4\mu(f^2)^2.$$

Therefore, the assertion in Theorem 1.1(3) holds for large enough $t > 0$. Since P_t is contractive in $L^2(\mu)$, it holds for all $t > 0$.

(d) We now go back to the proof of Theorem 1.1(1). This assertion follows from (a) and Theorem 1.1(3) by straightforward calculations. Let $f \in L^2(\mu)$ with $\mu(f^2) = 1$. Let $\widehat{f} = f - \mu(f)$. We have $\mu(P_t\widehat{f}) = \mu(\widehat{f}) = 0$. Let $t_0 > r_0$ such that $\|P_{t_0}\|_{2\to4} < \infty$, we obtain

$$\begin{aligned}
\mu((P_{t+t_0}f)^4) &= \mu(f)^4 + 4\mu(f)\mu((P_{t+t_0}\widehat{f})^3) + 6\mu(f)^2\mu((P_{t+t_0}\widehat{f})^2) + \mu((P_{t+t_0}\widehat{f})^4)\\
&\le \mu(f)^4 + 4|\mu(f)|\cdot\|P_{t_0}\|_{2\to3}^3\big\{\mu((P_t\widehat{f})^2)\big\}^{\frac{3}{2}}\\
&\quad + 6\mu(f)^2\mu((P_{t+t_0}\widehat{f})^2) + \|P_{t_0}\|_{2\to4}^4\mu((P_t\widehat{f})^2)^2\\
&\le \mu(f)^4 + c\mathrm{e}^{-\lambda t}\big\{|\mu(f)|\big(\mu(\widehat{f}^2)\big)^{\frac{3}{2}} + \mu(f)^2\mu(\widehat{f}^2) + \big(\mu(\widehat{f}^2)\big)^2\big\}
\end{aligned} \tag{2.20}$$

for some constant $c > 0$ according to Theorem 1.1(3). Since

$$|\mu(f)|\big(\mu(\widehat{f}^2)\big)^{\frac{3}{2}} \le \mu(f)^2\mu(\widehat{f}^2) + \big(\mu(\widehat{f}^2)\big)^2,$$

the relation (2.20) implies that for large $t > 0$,

$$\mu((P_{t+t_0}f)^4) \le \mu(f)^4 + 2\mu(f)^2\mu(\widehat{f}^2) + \big(\mu(\widehat{f}^2)\big)^2 = \big\{\mu(f)^2 + \mu(\widehat{f}^2)\big\}^2 = \mu(f^2) = 1.$$

(e) Finally, we prove Theorem 1.1(4). By the first inequality in (2.19), we have

$$\|\nabla_\eta P_{t_1+r_0}f\|^2(\xi) := \limsup_{s\to0} \frac{|P_{t_1+r_0}f(\xi+s\eta) - P_{t_1+r_0}f(\xi)|^2}{s^2} \le c_1\|\eta\|_\infty^2 P_{t_1+r_0}f^2(\xi).$$

Thus, $|P_{t_1+r_0}f(\xi) - P_{t_1+r_0}f(\eta)|^2 \le c_1\|f\|_\infty^2\|\xi-\eta\|_\infty^2$. Combining this with Lemma 2.2 and using the Markov property, we obtain

$$|P_{t+t_1+r_0}f(\xi) - P_{t+t_1+r_0}f(\eta)|^2 \le c_1\|f\|_\infty^2\mathbb{E}\|X_t^\xi - X_t^\eta\|_\infty^2 \le c_2\mathrm{e}^{-\lambda t}\|f\|_\infty^2$$

for some constants $c_2 > 0$. This completes the proof. □

Proof of Corollary 1.2. By (1.3) and (1.4), for any $s>0$ we have

$$\begin{aligned}&2\langle a(\xi(0))-a(\eta(0))+b(\xi)-b(\eta),\xi(0)-\eta(0)\rangle\\&\le -2k_1|\xi(0)-\eta(0)|^2+2k_2\|\xi-\eta\|_\infty\cdot|\xi(0)-\eta(0)|\\&\le -(2k_1-s)|\xi(0)-\eta(0)|^2+\frac{k_2^2}{s}\|\xi-\eta\|_\infty^2.\end{aligned}$$

Let $\lambda_1(s)=2k_1-s, \lambda_2(s)=\frac{k_2^2}{s}$. Then Theorem 1.1 applies if there exists $s\in(0,2k_1]$ such that

$$\lambda_2(s)<\lambda_1(s)\mathrm{e}^{-r_0\lambda_1(s)}=(2k_1-s)\mathrm{e}^{-r_0(2k_1-s)};$$

that is,

$$k_2^2<\sup_{s\in(0,2k_1)}(2k_1s-s^2)\mathrm{e}^{-r_0(2k_1-s)}, \tag{2.21}$$

where the sup is reached at

$$s_0=\frac{k_1r_0+\sqrt{k_1^2r_0^2+1}-1}{r_0},$$

such that (2.21) coincides with (1.5) and Theorem 1.1 applies with

$$\begin{aligned}&\lambda:=\lambda_1(s_0)-\lambda_2(s_0)\mathrm{e}^{r_0\lambda_1(s_0)}\\&=\frac{r_0}{k_1r_0-1+\sqrt{k_1^2r_0^2+1}}\left(\frac{2(\sqrt{k_1^2r_0^2+1}-1)}{r_0^2}-k_2^2\exp\left[1+k_1r_0-\sqrt{k_1^2r_0^2+1}\right]\right).\end{aligned}$$

□

3 Proof of Theorem 1.3

We first recall the following Fernique inequality [10] (see also [5]).

Lemma 3.1 (Fernique Inequality). *Let $(X(t))_{t\in D}$ be a family of centered Gaussian random variables on $\mathbb{R}^d$ with*

$$\sup_{t\in D}\mathbb{E}|X(t)|^2\le\sigma<\infty$$

for some constant $\sigma>0$, where $D:=\prod_{1\le i\le N}[a_i,b_i]$ is a cube in $\mathbb{R}^N$. Let $\phi\in C([0,\infty])$ be non-decreasing such that $\int_0^\infty\phi(\mathrm{e}^{-r^2})\mathrm{d}r<\infty$ and

$$\mathbb{E}|X(t)-X(s)|^2\le\phi(|t-s|),\quad s,t\in D.$$

Then there exist constants $C_1,C_2>0$ depending only on $(b_i-a_i)_{1\le i\le N},N,d,\phi$ and σ such that

$$\mathbb{P}\left(\sup_{t\in D}|X(t)|\ge r\right)\le C_1\mathrm{e}^{-C_2r^2},\quad r\ge1.$$

Proof of Theorem 1.3. Let P_t be the Markov semigroup associated with the equation (1.6) such that ν satisfies $\lambda_0 < 0$, b satisfies (1.4), and $\lambda = \sup_{k\in(0,-\lambda_0)}(k - c_k k_2 \mathrm{e}^{kr_0}) > 0$. Then, by following the proof of Lemma 2.4, (1.8) and (1.9) for some $k \in (0, -\lambda_0)$ imply that P_t has a unique invariant probability measure μ. Moreover, by taking $Z = 0$ and combining the linear drift with b, we see that Lemma 2.3 applies to the present equation for $k_1 = 0$ and some constant $k_2 > 0$. Thus, following the line in the proof of Theorem 1.1, we only need to show that Lemma 2.1 and Lemma 2.2 apply to the equation (1.6) as well.

Let $k \in (0, -\lambda_0)$ such that

$$\lambda = k - c_k k_2 \mathrm{e}^{kr_0} > 0. \tag{3.1}$$

It follows from (1.4), (1.8) and (1.9) that

$$\begin{aligned} |X^\xi(t) - X^\eta(t)| &\le \|\Gamma(t)\|\cdot|\xi(0)-\eta(0)| + \int_{-r_0}^0 \left\|\int_{-r_0}^\theta \Gamma(t+\theta-s)\nu(\mathrm{d}\theta)\right\|\cdot|\xi(s)-\eta(s)|\mathrm{d}s \\ &\quad + \int_0^t \|\Gamma(t-s)\|\cdot|b(X_s^\xi)-b(X_s^\eta)|\mathrm{d}s \\ &\le C_1\mathrm{e}^{-kt}\|\xi-\eta\|_\infty + c_k k_2\int_0^t \mathrm{e}^{-k(t-s)}\|X_s^\xi - X_s^\eta\|_\infty \mathrm{d}s \end{aligned}$$

for some constant $C_1 \ge 1$. Then

$$\begin{aligned} \mathrm{e}^{kt}\|X_t^\xi - X_t^\eta\|_\infty &\le \mathrm{e}^{kr_0}\sup_{t-r_0\le s\le t}(\mathrm{e}^{ks}|X^\xi(s)-X^\eta(s)|) \\ &\le C_1\mathrm{e}^{kr_0}\|\xi-\eta\|_\infty + c_k k_2 \mathrm{e}^{kr_0}\int_0^t \mathrm{e}^{ks}\|X_s^\xi - X_s^\eta\|_\infty\mathrm{d}s. \end{aligned}$$

This, together with Gronwall's inequality, gives that

$$\|X_t^\xi - X_t^\eta\|_\infty \le C_1\mathrm{e}^{kr_0}\|\xi-\eta\|_\infty \mathrm{e}^{-\lambda t}. \tag{3.2}$$

So, Lemma 2.2 applies.

Next, by (1.4), (1.8) and (1.9),

$$\begin{aligned} \mathrm{e}^{kt}\|X_t^\xi\|_\infty &\le C_2(\|\xi\|_\infty + \mathrm{e}^{kt}) + c_k\mathrm{e}^{kr_0}k_2\int_0^t \mathrm{e}^{ks}\|X_s^\xi\|_\infty\mathrm{d}s \\ &\quad + \mathrm{e}^{kr_0}\sup_{(t-r_0)^+\le s\le t}\left(\mathrm{e}^{ks}\left|\int_0^s \Gamma(s-r)\sigma\mathrm{d}B(r)\right|\right) \end{aligned}$$

holds for some constant $C_2 > 0$. By Gronwall's inequality, this implies

$$\begin{aligned}\|X_t^\xi\|_\infty &\le C_2(\|\xi\|_\infty + 1) + \mathrm{e}^{kr_0} \sup_{t-r_0\le s\le t}\Big|\int_0^s \Gamma(s-r)\sigma \mathrm{d}B(r)\Big| \\ &\quad + C_2(\|\xi\|_\infty+1)\int_0^t \mathrm{e}^{-\lambda(t-s)}\mathrm{d}s \\ &\quad + \mathrm{e}^{kr_0}\int_0^t \bigg(\sup_{(s-r_0)^+\le u\le s}\Big|\int_0^u \Gamma(u-r)\sigma\mathrm{d}B(r)\Big|\bigg)\mathrm{e}^{-\lambda(t-s)}\mathrm{d}s \\ &\le C_3(1+\|\xi\|_\infty^2) + C_3\int_0^t \mathrm{e}^{-\lambda(t-s)} \sup_{u\in[-r_0,0]}|Z_{s,u}|\mathrm{d}s\end{aligned}$$

for some constant $C_3 > 0$, where

$$Z_{s,u} := \int_0^{(s+u)^+} \Gamma(s+u-r)\sigma\mathrm{d}B(r), \quad s\ge 0, u\in[-r_0,0].$$

Then, by Jensen's inequality for the probability measure $\frac{\lambda \mathrm{e}^{\lambda r_0}}{\mathrm{e}^{\lambda r_0}-\mathrm{e}^{-\lambda t}}\mathrm{e}^{-\lambda(t-s)}\mathrm{d}s$ on $[-r_0,t]$, there exists a constant $C_4 > 0$ such that

$$\begin{aligned}\mathbb{E}\mathrm{e}^{\varepsilon\|X_t^\xi\|_\infty^2} &\le \mathrm{e}^{\varepsilon C_4(1+\|\xi\|_\infty^2)}\mathbb{E}\exp\bigg[\varepsilon C_4\bigg(\int_0^t \mathrm{e}^{-\lambda(t-s)}\sup_{u\in[-r_0,0]}|Z_{s,u}|\mathrm{d}s\bigg)^2\bigg] \\ &\le \mathrm{e}^{\varepsilon C_4(1+\|\xi\|_\infty^2)}\frac{\lambda\mathrm{e}^{\lambda r_0}}{\lambda \mathrm{e}^{\lambda r_0}-1}\int_{-r_0}^t \mathrm{e}^{-\lambda(t-s)}\bigg(\mathbb{E}\exp\Big[\frac{C_4\varepsilon}{\lambda}\sup_{u\in[-r_0,0]}|Z_{s,u}|^2\Big]\bigg)\mathrm{d}s\end{aligned} \tag{3.3}$$

for any $\varepsilon > 0$, where we set $Z_{s,u} = 0$ for $s \in [-r_0, 0]$. Note from Itô's isometry and (1.9) that

$$\sigma := \sup_{s\ge 0, u\in[-r_0,0]}\mathbb{E}|Z_{s,u}|^2 < \infty,$$

and that there exist constants $c_1, c_2, c_3 > 0$ such that

$$\begin{aligned}\mathbb{E}|Z_{s,u}-Z_{s,v}|^2 &\le 2\mathbb{E}\Big|\int_{(s+v)^+}^{(s+u)^+}\Gamma(s+u-r)\sigma\mathrm{d}B(r)\Big|^2 \\ &\quad + 2\mathbb{E}\Big|\int_0^{(s+v)^+}\big(\Gamma(s+u-r)-\Gamma(s+v-r)\big)\sigma\mathrm{d}B(r)\Big|^2 \\ &\le c_1|u-v| + 2\|\sigma\|^2\int_0^{(s+v)^+}\|\Gamma(s+u-r)-\Gamma(s+v-r)\|^2\mathrm{d}r \\ &\le c_1|u-v| + c_2|u-v|^2 \le c_3|u-v|, \quad s\ge 0, \quad -r_0\le v\le u\le 0.\end{aligned}$$

Thus, by Lemma 3.1 with $N = 1, D = [-r_0, 0]$ and $\phi(r) = cr$,

$$C(\varepsilon) := \sup_{s\ge 0}\mathbb{E}\exp\Big[\frac{C_4\varepsilon}{\lambda}\sup_{u\in[-r_0,0]}|Z_{s,u}|^2\Big] < \infty$$

holds for small enough $\varepsilon > 0$. Therefore, (3.3) implies the assertion in Lemma 2.1. □

4 Appendix

For application of Theorem 1.3, we aim to estimate the constant c_k in (1.9). Write $\nu = (\nu_{ij})_{1\le i,j\le d}$ for finite signed measures ν_{ij} on $[-r_0, 0]$. Let $|\nu_{ij}|$ be the total variation of ν_{ij}. For any $\lambda > \lambda_0$, define

$$\|\nu\| = \sup_{1\le i\le d}\sqrt{\sum_{1\le j\le d}|\nu_{ij}|([-r_0,0])^2},\quad T_\lambda = 2\mathrm{e}^{\lambda^- r_0}\|\nu\|,\quad \lambda^- = (-\lambda)\vee 0,$$

$$\rho_\lambda = \max_{\theta\in[-T_\lambda,T_\lambda]}\left\|\left((\lambda+\mathrm{i}\theta)I_{d\times d} - \int_{-r_0}^0 \mathrm{e}^{\lambda+\mathrm{i}\theta}\nu(\mathrm{d}s)\right)^{-1} - (\lambda+\mathrm{i}\theta-\lambda_0)^{-1}I_{d\times d}\right\|.$$

Proposition 4.1. For any $\lambda > \lambda_0$,

$$\|\Gamma(t)\| \le \left\{\frac{(\lambda-\lambda_0+1)\pi}{\lambda-\lambda_0} + \frac{4(|\lambda_0| + \mathrm{e}^{\lambda^- r_0}\|\nu\|)}{T_\lambda} + 2\rho_\lambda T_\lambda\right\}\mathrm{e}^{\lambda t},\quad t\ge 0.$$

Proof. For any $z \ne \lambda_0$, define

$$Q_z = zI_{d\times d} - \int_{r_0}^0 \mathrm{e}^{zs}\nu(\mathrm{d}s),\quad G_z = Q_z^{-1} - \frac{1}{z-\lambda_0}I_{d\times d}.$$

We have (see [12, Theorem 1.5.1])

$$\Gamma(t) = \lim_{T\to\infty}\int_{-T}^T Q_{\lambda+i\theta}^{-1}\mathrm{e}^{t(\lambda+\mathrm{i}\theta)}\mathrm{d}\theta = \lim_{T\to\infty}\int_{-T}^T\left(G_{\lambda+i\theta} + \frac{I_{d\times d}}{\lambda-\lambda_0+\mathrm{i}\theta}\right)\mathrm{e}^{t(\lambda+\mathrm{i}\theta)}\mathrm{d}\theta,\quad \lambda>\lambda_0. \tag{4.1}$$

Obviously, $\left\|\int_{-r_0}^0 \mathrm{e}^{(\lambda+\mathrm{i}T)s}\nu(\mathrm{d}s)\right\| \le \mathrm{e}^{r_0\lambda^-}\|v\|$ and

$$\sqrt{1+\lambda^2T^{-2}} - \frac{\mathrm{e}^{\lambda^- r_0}\|\nu\|}{|T|} \ge \frac{1}{2},\quad |T|\ge T_\lambda.$$

Then

$$\|Q_{\lambda+\mathrm{i}T}^{-1}\| \le \frac{1}{\sqrt{\lambda^2+T^2} - \mathrm{e}^{|\lambda|r_0}\|\nu\|} \le \frac{2}{|T|},\quad |T|\ge T_\lambda.$$

This yields

$$\begin{aligned}\|G_{\lambda+\mathrm{i}T}\| &\le \|Q_{\lambda+\mathrm{i}T}^{-1}\|\cdot\left\|\frac{\int_{-r_0}^0 \mathrm{e}^{(\lambda+\mathrm{i}T)s}\nu(\mathrm{d}s) - \lambda_0 I_{d\times d}}{\lambda+\mathrm{i}T-\lambda_0}\right\| \\ &\le \frac{2(|\lambda_0| + \mathrm{e}^{\lambda^- r_0}\|\nu\|)}{|T|\sqrt{(\lambda-\lambda_0)^2+T^2}} \le \frac{2(|\lambda_0| + \mathrm{e}^{\lambda^- r_0}\|\nu\|)}{T^2},\quad |T|\ge T_\lambda.\end{aligned}$$

Thus, for any $T \ge T_\lambda$,

$$\begin{aligned}\int_T^T \|G_{\lambda+i\theta}\mathrm{e}^{t(\lambda+\mathrm{i}\theta)}\|\mathrm{d}\theta &= \int_{-T_\lambda}^{T_\lambda}\|G_{\lambda+\mathrm{i}\theta}\mathrm{e}^{t(\lambda+\mathrm{i}\theta)}\|\mathrm{d}\theta + \int_{|\theta|>T_\lambda}\|G_{\lambda+\mathrm{i}\theta}\mathrm{e}^{t(\lambda+\mathrm{i}\theta)}\|\mathrm{d}\theta \\ &\le 2\rho_\lambda T_\lambda \mathrm{e}^{\lambda t} + \frac{4(|\lambda_0| + \mathrm{e}^{\lambda^- r_0}\|\nu\|)\mathrm{e}^{\lambda t}}{T_\lambda}.\end{aligned} \tag{4.2}$$

On the other hand,

$$\lim_{T\to\infty}\int_{-T}^{T}\frac{\mathrm{e}^{t(\lambda+\mathrm{i}\theta)}}{\lambda-\lambda_0+\mathrm{i}\theta}\,\mathrm{d}\theta = i\mathrm{e}^{\lambda t}\lim_{T\to\infty}\int_{-T}^{T}\frac{(\lambda-\lambda_0)\mathrm{e}^{it\theta}}{(\lambda-\lambda_0)^2+\theta^2}\mathrm{d}\theta - \mathrm{e}^{\lambda t}\lim_{T\to\infty}\int_{-T}^{T}\frac{\theta\mathrm{e}^{it\theta}}{(\lambda-\lambda_0)^2+\theta^2}\mathrm{d}\theta$$
$$=: \Theta_1+\Theta_2.$$

It is easy to see that

$$\|\Theta_1\| \le \frac{2\mathrm{e}^{\lambda t}}{\lambda-\lambda_0}\lim_{T\to\infty}\arctan\Big(\frac{\theta}{\lambda-\lambda_0}\Big)\Big|_0^T = \frac{\pi\mathrm{e}^{\lambda t}}{\lambda-\lambda_0}.$$

Moreover, by the residue theorem,

$$\begin{aligned}\|\Theta_2\| &= \Big|-2\pi\mathrm{e}^{\lambda t}i\mathrm{Res}\Big[\frac{z\mathrm{e}^{itz}}{(\lambda-\lambda_0)^2+z^2},(\lambda-\lambda_0)i\Big]\Big|\\ &= \Big|-2\pi\mathrm{e}^{\lambda t}i\lim_{z\to(\lambda-\lambda_0)i}(z-(\lambda-\lambda_0)i)\times\frac{z\mathrm{e}^{itz}}{(\lambda-\lambda_0)^2+z^2}\Big|\\ &= \Big|-2\pi\mathrm{e}^{\lambda t}i\lim_{z\to(\lambda-\lambda_0)i}\frac{z\mathrm{e}^{itz}}{2(\lambda-\lambda_0)i}\Big|\\ &= \Big|-2\pi\mathrm{e}^{\lambda t}i\frac{(\lambda-\lambda_0)i\mathrm{e}^{-t(\lambda-\lambda_0)}}{2(\lambda-\lambda_0)i}\Big|\\ &= \pi\mathrm{e}^{\lambda_0 t}\le\pi\mathrm{e}^{\lambda t}.\end{aligned}$$

Hence, we arrive at

$$\Big|\lim_{T\to\infty}\int_{-T}^{T}\frac{\mathrm{e}^{t(\lambda+\mathrm{i}\theta)}}{\lambda-\lambda_0+\mathrm{i}\theta}\mathrm{d}\theta\Big| \le \frac{(\lambda-\lambda_0+1)\pi\mathrm{e}^{\lambda t}}{\lambda-\lambda_0}.$$

Combing this with (4.2) and (4.1), we finish the proof.

□

Acknowledgement. The authors would like to thank the referee and the editor for helpful comments.

References

[1] D. Applebaum, *Lévy processes and stochastic calculus*, 2nd Ed., Cambridge University Press, Cambridge, 2009.

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

[3] D. Bakry, I. Gentil, M. Ledoux, *Analysis and Geometry of Markov Diffusion Operators,* Springer, 2013, Berlin.

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

[5] S. M. Berman, *An asymptotic bound for the tail of the distribution of maximum of a Gaussian process with stationary increaments,* Ann. Inst. H. Poinc. Probab. Stat. 21 (1985), 47–57.

[6] M.-F. Chen, *From Markov Chains to Non-Equilibrium Particle Systems,* World Scientific, 1992, Singapore.

[7] E. B. Davies, *Heat Kernels and Spectral Theory,* Cambridge: Cambridge Univ. Press, 1989.

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

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

[10] X. Fernique, *Régularité des trajectoires des fonctions aléatoires Gaussiennes,* In: Ecole d'été des Probabilités de St-Flour, IV-1974. Lect. Notes Math. 480, pp. 1–96, Springer, Berlin, 1975.

[11] L. Gross, *Logarithmic Sobolev inequalities and contractivity properties of semigroups,* Lecture Notes in Math. 1563, Springer-Verlag, 1993.

[12] J. K. Hale, S. M. Verduyn Lunel, *Introduction to Functional Differential Equations,* Springer, 1993, Berlin.

[13] 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 (2011), 223–259.

[14] M. S. Kinnally, R. J. Williams, *On existence and uniqueness of stationary distributions for stochastic delay differential equations with positivity constraints*, Electron. J. Probab. 15 (2010), 409–451.

[15] S. Kusuoka, D. Stroock, *Applications of the Malliavin calculus, I*, Tuniguchi Sympos. SA Kumta (1982), 271–306.

[16] W. Liu, *Harnack inequality and applications for stochastic evolution equations with monotone drifts,* J. Evol. Equ. 9 (2009), 747–770.

[17] E. Nelson, *The free Markov field,* J. Funct. Anal. 12 (1973), 211–227.

[18] Reiβ, M., Riedle, M., van Gaans, O., *Delay differential equations driven by Lévy processes: stationarity and Feller properties,* Stochastic Process. Appl. 116 (2006), 1409–1432.

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

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

[21] F.-Y. Wang, *Functional Inequalities, Markov Semigroups and Spectral Theory,* Science Press, 2005, Beijing.

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

[23] F.-Y. Wang, *Analysis for Diffusion Processes on Riemannian Manifolds,* World Scientific, 2013, Singapore.

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

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

[26] L. Wu, *Uniformly integrable operators and large deviations for Markov processes,* J. Funct. Anal. 172 (2000), 301–376.