Electronic Journal of Probability

Electron. J. Probab. **20** (2015), no. 93, 1–15.
ISSN: 1083-6489 DOI: 10.1214/EJP.v20-4108

Hypercontractivity for functional stochastic partial differential equations*

Jianhai Bao† Feng-Yu Wang‡ Chenggui Yuan§

Abstract

Explicitly sufficient conditions on the hypercontractivity are presented for two classes of functional stochastic partial differential equations driven by, respectively, non-degenerate and degenerate Gaussian noises. Consequently, these conditions imply that the associated Markov semigroup is L^2-compact and exponentially convergent to the stationary distribution in entropy, variance and total variational norm. As the log-Sobolev inequality is invalid under the present framework, we apply a criterion presented in the recent paper [15] using Harnack inequality, coupling property and Gaussian concentration property of the stationary distribution. To verify the concentration property, we prove a Fernique type inequality for infinite-dimensional Gaussian processes which might be interesting by itself.

Keywords: Hypercontractivity; functional stochastic partial differential equation; Harnack inequality; coupling .
AMS MSC 2010: 60H15; 60J60.
Submitted to EJP on February 13, 2015, final version accepted on September 9, 2015.

1 Introduction

The hypercontractivity was introduced in 1973 by Nelson [10] for the Ornstein-Ulenbeck semigroup. As applications, it implies the exponential convergence of the Markov semigroup in entropy (and hence, also in variance) to the associated stationary distribution, and it also implies the L^2-compactness of the semigroup subject to the existence of a density with respect to the stationary distribution, see [15] for more details. In the setting of symmetric Markov processes, Gross [9] proved that the hypercontractivity of the semigroup is equivalent to the log-Sobolev inequality for the associated Dirichlet form. This leads to an extensive study of the log-Sobolev inequality.

However, as explained in [3] the log-Sobolev inequality does not hold for the segment solution to a stochastic delay differential equation (SDDE). As the segment solution

*Supported in part by Lab. Math. Com. Sys., the 985 project and NNSFC(11131003, 11431014, 11401592).
†School of Mathematics and Statistics, Central South University, Changsha, China.
E-mail: jianhaibao13@gmail.com
‡School of Mathematical Sciences, Beijing Normal University, Beijing, China.
E-mail: wangfy@bnu.edu.cn,F.-Y.Wang@swansea.ac.uk
§Department of Mathematics, Swansea University, UK.
E-mail: C.Yuan@swansea.ac.uk

is a process on a functional space, the equation is also called a functional stochastic differential equation (FSDE). In this case, an efficient tool to prove the hypercontractivity is the dimension-free Harnack inequality introduced in [11], where diffusion semigroups on Riemannian manifolds are concerned. By using the coupling by change of measures, this type Harnack inequality has been established for various stochastic equations, see the recent monograph [14] and references within. The aim of the present paper is to prove the hypercontractivity for functional stochastic partial differential equations (FSPDEs) in Hilbert spaces. We will consider non-degenerate noise and degenerate noise, respectively, so that the corresponding results derived in [3] for finite-dimensional FSDEs as well as in [15] for degenerate SPDEs are extended.

In the recent paper [15], the second named author developed a general criterion on the hypercontractivity by using the Harnack inequality of the semigroup, the concentration property of the underlying probability measure, and the coupling property. In general, let P_t be a Markov semigroup on $L^2(\mu)$ for a probability space $(E, \mathscr{F}, \mu)$ such that μ is P_t-invariant. By definition, P_t is hypercontractive if $\|P_t\|_{2\to4} = 1$ holds for large enough $t > 0$, where $\|\cdot\|_{2\to4}$ is the operator norm from $L^2(\mu)$ to $L^4(\mu)$. For any $(x, y) \in E \times E$, a process (X_t, Y_t) on $E \times E$ is called a coupling for the Markov semigroup with initial point (x, y) if

$$P_t f(x) = \mathbb{E} f(X_t), \quad P_t f(y) = \mathbb{E} f(Y_t), \quad t \geq 0, f \in \mathscr{B}_b(E),$$

where $\mathscr{B}_b(E)$ stands for the set of all bounded measurable functions defined on E.

The general criterion due to Wang [15] is stated as follows.

Theorem 1.1 ([14]). Assume that the following three conditions hold for some measurable functions $\rho : E \times E \mapsto (0, \infty)$ and $\phi : [0, \infty) \mapsto (0, \infty)$ such that $\lim_{t\to\infty} \phi(t) = 0$:

(i) **(Harnack Inequality)** There exist constants $t_0, c_0 > 0$ such that

$$(P_{t_0} f(\xi))^2 \leq (P_{t_0} f^2(\eta)) \mathrm{e}^{c_0 \rho(\xi,\eta)^2}, \quad f \in \mathscr{B}_b(E),\ \xi, \eta \in E;$$

(ii) **(Coupling Property)** For any $(\xi, \eta) \in E \times E$, there exists a coupling (X_t, Y_t) for the Markov semigroup P_t such that

$$\rho(X_t, Y_t) \leq \phi(t) \rho(\xi, \eta), \quad t \geq 0;$$

(iii) **(Concentration Property)** There exists $\varepsilon > 0$ such that $(\mu \times \mu)(\mathrm{e}^{\varepsilon \rho(\cdot,\cdot)^2}) < \infty$.

Then P_t is hypercontractive and compact in $L^2(\mu)$ for large enough $t > 0$, and

$$\begin{aligned} &\mu((P_t f) \log P_t f) \leq c \mathrm{e}^{-\alpha t} \mu(f \log f), \quad t \geq 0, f \geq 0, \mu(f) = 1; \\ &\|P_t - \mu\|_2^2 := \sup_{\mu(f^2) \leq 1} \mu\big((P_t f - \mu(f))^2\big) \leq c \mathrm{e}^{-\alpha t}, \quad t \geq 0 \end{aligned} \tag{1.1}$$

hold for some constants $c, \alpha > 0$.

We will apply the previous criterion to non-degenerate and degenerate FSPDEs, respectively. To state our main results, we first introduce some notation.

For two separable Hilbert spaces $\mathbb{H}_1, \mathbb{H}_2$, let $\mathscr{L}(\mathbb{H}_1, \mathbb{H}_2)$ (respectively, $\mathscr{L}_{HS}(\mathbb{H}_1, \mathbb{H}_2)$) be the set of all bounded (respectively, Hilbert-Schmidt) linear operators from $\mathbb{H}_1$ to $\mathbb{H}_2$. We will use $|\cdot|$ and $\langle \cdot, \cdot \rangle$ to denote the norm and the inner product on a Hilbert space, and let $\|\cdot\|$ and $\|\cdot\|_{HS}$ stand for the operator norm and the Hilbert-Schmidt norm for a linear operator. Below we introduce our main results for non-degenerate FSPDEs and degenerate FSPDEs, respectively.

1.1 Non-Degenerate FSPDEs

Let $\mathbb{H}$ be a separable Hilbert space. For a fixed constant $r_0 > 0$, let $\mathscr{C} = C([-r_0, 0]; \mathbb{H})$ be equipped with the uniform norm $\|f\|_\infty := \sup_{-r_0 \le \theta \le 0} |f(\theta)|$. For $t \ge 0$ and $h \in C([-r_0, \infty); \mathbb{H})$, let $h_t \in \mathscr{C}$ be such that $h_t(\theta) = h(t+\theta), \theta \in [-r_0, 0]$.

Let $W(t)$ be a cylindrical Wiener process on $\mathbb{H}$ under a complete filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\ge0}, \mathbb{P})$; that is,

$$W(t) = \sum_{i=1}^{\infty} B_i(t) e_i, \quad t \ge 0$$

for an orthonormal basis $\{e_i\}_{i\ge1}$ on $\mathbb{H}$ and a sequence of independent one-dimensional Wiener processes $\{B_i(t)\}_{i\ge1}$ on $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\ge0}, \mathbb{P})$.

Consider the following FSPDE on $\mathbb{H}$:

$$\mathrm{d}X(t) = \{AX(t) + b(X_t)\}\mathrm{d}t + \sigma \mathrm{d}W(t), \quad t > 0, \quad X_0 = \xi \in \mathscr{C}, \tag{1.2}$$

where $(A, \mathscr{D}(A))$ is a densely defined closed operator on $\mathbb{H}$ generating a C_0-semigroup e^{tA}, $b : \mathscr{C} \mapsto \mathbb{H}$ is measurable, $(\sigma, \mathscr{D}(\sigma))$ is a densely defined linear operator on $\mathbb{H}$. We assume that A, b and σ satisfy the following conditions.

(A1) $(-A, \mathscr{D}(A))$ is self-adjoint with discrete spectrum $0 < \lambda_1 \le \lambda_2 \le \cdots$ counting multiplicities such that $\lambda_i \uparrow \infty$. Moreover, there exists a constant $\delta \in (0,1)$ such that, for every $t > 0$, $\mathrm{e}^{-t(-A)^{1-\delta}}\sigma$ extends to a unique Hilbert-Schmidt operator on $\mathbb{H}$ which is denoted again by $\mathrm{e}^{-t(-A)^{1-\delta}}\sigma$ and satisfies

$$\int_0^1 \|\mathrm{e}^{-t(-A)^{1-\delta}}\sigma\|_{HS}^2 \mathrm{d}t < \infty. \tag{1.3}$$

(A2) There exists a constant $L > 0$ such that $|b(\xi) - b(\eta)| \le L\|\xi - \eta\|_\infty, \ \xi, \eta \in \mathscr{C}$.

(A3) σ is invertible, i.e., there exists $\sigma^{-1} \in \mathscr{L}(\mathbb{H}, \mathbb{H})$ such that $\sigma^{-1}\mathbb{H} \subset \mathscr{D}(\sigma)$ and $\sigma\sigma^{-1} = I$, the identity operator.

We first observe that assumptions **(A1)** and **(A2)** imply the existence and uniqueness of continuous mild solutions to (1.2); that is, for any $\mathscr{F}_0$-measurable random variable $X_0 = \xi \in \mathscr{C}$, there exists a unique continuous adapted process $\{X(t)\}_{t\ge r_0}$ on $\mathbb{H}$ such that $\mathbb{P}$-a.s.

$$X(t) = \mathrm{e}^{tA}\xi(0) + \int_0^t \mathrm{e}^{(t-s)A} b(X_s)\mathrm{d}s + \int_0^t \mathrm{e}^{(t-s)A}\sigma \mathrm{d}W(s), \quad t \ge 0. \tag{1.4}$$

To this end, it suffices to show that (1.3) implies

$$\int_0^1 \|\mathrm{e}^{tA}\sigma\|_{HS}^{2(1+\varepsilon)} \mathrm{d}t < \infty \tag{1.5}$$

for some $\varepsilon > 0$, see, for instance, [14, Theorem 4.1.3]. To prove (1.5), we reformulate condition (1.3) using the eigenbasis $\{e_i\}_{i\ge1}$ of A, i.e., $\{e_i\}_{i\ge1}$ is an orthonormal basis of $\mathbb{H}$ such that $Ae_i = -\lambda_i e_i, i \ge 1$. By noting that

$$\|\mathrm{e}^{-t(-A)^{1-\delta}}\sigma\|_{HS}^2 = \|(\mathrm{e}^{-t(-A)^{1-\delta}}\sigma)^*\|_{HS}^2 = \sum_{j=1}^{\infty} |(\mathrm{e}^{-t(-A)^{1-\delta}}\sigma)^* e_j|^2 = \sum_{j=1}^{\infty} \mathrm{e}^{-2\lambda_j^{1-\delta}t} |\sigma^* e_j|^2,$$

(1.3) is equivalent to

$$\sum_{j=1}^{\infty} \frac{|\sigma^* e_j|^2}{\lambda_j^{1-\delta}} < \infty. \tag{1.6}$$

This implies that $\mu_j := \frac{|\sigma^* e_j|^2}{\lambda_j^{1-\delta}}(j \geq 1)$ gives rise to a finite measure on $\mathbb{N}$, so that by Hölder's inequality,

$$\begin{aligned}
&\int_0^1 \|\mathrm{e}^{tA}\sigma\|_{HS}^{2(1+\varepsilon)}\mathrm{d}t = \int_0^1 \Big(\sum_{j=1}^\infty \mathrm{e}^{-2\lambda_j t}|\sigma^* e_j|^2\Big)^{1+\varepsilon}\mathrm{d}t \\
&= \int_0^1 \Big(\sum_{j=1}^\infty \mu_j \mathrm{e}^{-2\lambda_j t}\lambda_j^{1-\delta}\Big)^{1+\varepsilon}\mathrm{d}t \leq C\int_0^1 \Big(\sum_{j=1}^\infty \mu_j \lambda_j^{(1+\varepsilon)(1-\delta)}\mathrm{e}^{-2(1+\varepsilon)\lambda_j t}\Big)\mathrm{d}t \\
&\leq C\sum_{j=1}^\infty \frac{|\sigma^* e_j|^2}{\lambda_j^{1-\varepsilon(1-\delta)}} < \infty, \ \ \varepsilon \leq \frac{\delta}{1-\delta},
\end{aligned}$$

where $C := (\sum_{i=1}^\infty \mu_i)^\varepsilon$. Thus, (1.3) implies (1.5) for $\varepsilon \in (0, \frac{\delta}{1-\delta}]$.

To emphasize the initial datum $X_0 = \xi \in \mathscr{C}$, we denote the solution and the segment solution by $\{X^\xi(t)\}_{t\geq -r_0}$ and $\{X_t^\xi\}_{t\geq 0}$, respectively. Then the Markov semigroup for the segment solution is defined as

$$P_t f(\xi) = \mathbb{E} f(X_t^\xi), \ f \in \mathscr{B}_b(\mathscr{C}), \ \xi \in \mathscr{C}, \ t \geq 0. \tag{1.7}$$

We are ready to state the main result in this part.

Theorem 1.2. Let **(A1)**-**(A3)** hold. If $\lambda := \sup_{s\in(0,\lambda_1]}\big(s - L\mathrm{e}^{sr_0}\big) > 0$, then the following assertions hold.

(1) P_t has a unique invariant probability measure μ such that $\mu(\mathrm{e}^{\varepsilon\|\cdot\|_\infty^2}) < \infty$ for some $\varepsilon > 0$.

(2) P_t is hypercontractive and compact in $L^2(\mu)$ for large enough $t > 0$, and (1.1) holds for some constants $c, \alpha > 0$.

(3) For any $t_0 > r_0$, there exists a constant $c > 0$ such that

$$\|\mu_t^\xi - \mu_t^\eta\|_{\mathrm{var}} \leq c\|\xi - \eta\|_\infty \mathrm{e}^{-\lambda t}, \ \ t \geq t_0,$$

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

To illustrate the above result, we present below an example, where $\mathbb{H} = L^2(D; \mathrm{d}x)$ for a bounded domain in $\mathbb{R}^d$.

Example 1.1. For a bounded domain $D \subset \mathbb{R}^d$, let $\mathbb{H} = L^2(D; \mathrm{d}x)$ and $A = -(-\Delta)^\alpha$, where Δ is the Dirichlet Laplacian on D and $\alpha > \frac{d}{2}$ is a constant. Let $\sigma = I$ be the identity operator on $\mathbb{H}$, and $b(\xi) = L\int_{-r_0}^0 \xi(r)\nu(\mathrm{d}r)$ for a signed measure ν on $[-r_0, 0]$ with total variation 1; or $b(\xi) = \sup_{r\in[-r_0,0]}\langle \xi(r), g(r)\rangle$ for some measurable $g : [-r_0, 0] \to \mathbb{H}$ with $\|g\|_\infty \leq L$. Then assertions in Theorem 1.2 hold provided

$$\lambda := \sup_{s\in(0,(d\pi^2)^\alpha R(D)^{-2\alpha})} (s - L\mathrm{e}^{sr_0}) > 0,$$

where $R(D)$ is the diameter of D.

Proof. Since $A = -(-\Delta)^\alpha$, it is well known that the eigenvalues $\{\lambda_i\}_{i\geq 1}$ of A satisfy $\lambda_i \geq c i^{\frac{2\alpha}{d}}$ $(i \geq 1)$ for some constant $c > 0$. So, for $\alpha > \frac{d}{2}$ assumptions **(A1)**-**(A3)** hold for the above choices of $\mathbb{H}, A, \sigma$ and b. By Theorem 1.2, it remains to prove $\lambda_1 \geq \frac{(d\pi^2)^\alpha}{R(D)^{2\alpha}}$. Letting $\overline{\lambda}_1$ be the first eigenvalue of $-\Delta$, by the definition of A this is equivalent to $\overline{\lambda}_1 \geq \frac{d\pi^2}{R(D)^2}$. As D is covered by a cube of edge length $R(D)$, by the domain-monotonicity and the shift-invariance of the first Dirichlet eigenvalue of $-\Delta$, $\overline{\lambda}_1$ is bounded below by the first Dirichlet eigenvalue of $-\Delta$ on the cube $[0, R(D)]^d$, which is equal to $\frac{d\pi^2}{R(D)^2}$ with eigenfunction $u(x) := \prod_{i=1}^d \sin\big(\frac{\pi x_i}{R(D)}\big)$. Then the proof is finished. □

1.2 Degenerate FSPDEs

Let $\mathbb{H} = \mathbb{H}_1 \times \mathbb{H}_2$ for two separable Hilbert spaces $\mathbb{H}_1$ and $\mathbb{H}_2$, and let $\mathscr{C} = C([-r_0, 0]; \mathbb{H})$ as in Subsection 1.1. Consider the following degenerate FSPDE on $\mathbb{H}$:

$$\begin{cases} \mathrm{d}X(t) = \{A_1 X(t) + BY(t)\}\mathrm{d}t, \\ \mathrm{d}Y(t) = \{A_2 Y(t) + b(X_t, Y_t)\}\mathrm{d}t + \sigma \mathrm{d}W(t), \end{cases} \tag{1.8}$$

where $(A_i, \mathscr{D}(A_i))$ is a densely defined closed linear operator on $\mathbb{H}_i$ generating a C_0-semigroup e^{tA_i} $(i = 1, 2)$, $B \in \mathscr{L}(\mathbb{H}_2, \mathbb{H}_1)$, $b : \mathscr{C} \mapsto \mathbb{H}_2$ is measurable, $(\sigma, \mathscr{D}(\sigma))$ is a densely defined closed operator on $\mathbb{H}_2$, and $W(t)$ is the cylindrical Wiener process on $\mathbb{H}_2$. Corresponding to **(A1)**-**(A3)** in the non-degenerate case, we make the following assumptions (see [15] for the case without delay, i.e., $b(X_t, Y_t)$ depends only on $X(t)$ and $Y(t)$).

(B1) $(-A_2, \mathscr{D}(A_2))$ is self-adjoint with discrete spectrum $0 < \lambda_1 \leq \lambda_2 \leq \cdots$ counting multiplicities such that $\lambda_i \uparrow \infty$, σ is invertible, and

$$\int_0^1 \|\mathrm{e}^{-t(-A_2)^{1-\delta_0}} \sigma\|_{HS}^2 \mathrm{d}t < \infty$$

holds for some constant $\delta_0 \in (0, 1)$.

(B2) There exist constants $K_1, K_2 > 0$ such that

$$|b(\xi_1, \eta_1) - b(\xi_2, \eta_2)| \leq K_1 \|\xi_1 - \eta_1\|_\infty + K_2 \|\xi_2 - \eta_2\|_\infty, \quad (\xi_1, \eta_1),\ (\xi_2, \eta_2) \in \mathscr{C}.$$

(B3) $A_1 \leq \delta - \lambda_1$ for some constant $\delta \geq 0$; i.e., $\langle A_1 x, x\rangle \leq (\delta - \lambda_1)|x|^2$ holds for all $x \in \mathscr{D}(A_1)$.

(B4) There exists $A_0 \in \mathscr{L}(\mathbb{H}_1, \mathbb{H}_1)$ such that $B\mathrm{e}^{tA_2} = \mathrm{e}^{tA_1}\mathrm{e}^{tA_0}B$ holds for $t \geq 0$, and

$$Q_t := \int_0^t \mathrm{e}^{sA_0} BB^* \mathrm{e}^{sA_0^*} \mathrm{d}s, \quad t \geq 0$$

is invertible on $\mathbb{H}_1$.

Obviously, when $\mathbb{H}_1 = \mathbb{H}_2, \sigma = B = I$ and $A_1 = A_2$ with discrete spectrum $\{-\lambda_i\}_{i\geq 1}$ such that $\sum_{i=1}^\infty \frac{1}{\lambda_i^{1-\delta}} < \infty$ holds for some constant $\delta \in (0, 1)$, then assumptions **(B1)**, **(B3)** and **(B4)** hold. See [15] for more examples, where $\mathbb{H}_2$ might be a subspace of $\mathbb{H}_1$.

Similarly to the case without delay considered in [15], assumptions **(B3)** and **(B4)** will be used to prove the Harnack inequality. Moreover, as explained in Subsection 1.1 for the non-degenerate case, from [14, Theorem 4.1.3] we conclude that assumptions **(B1)** and **(B2)** imply the existence, uniqueness and non-explosion of the continuous mild solution $(X^{\xi,\eta}(t), Y^{\xi,\eta}(t))$ for any initial point $(\xi, \eta) \in \mathscr{C}$. Let P_t be the Markov semigroup generated by the segment solution. We have

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

Theorem 1.3. Assume **(B1)**-**(B4)**. If

$$\lambda' := \frac{1}{2}\Big(\delta + K_2 + \sqrt{(K_2 - \delta)^2 + 4K_1\|B\|}\Big) < \sup_{s\in(0,\lambda_1]} s\mathrm{e}^{-sr_0}, \tag{1.9}$$

then all assertions in Theorem 1.1 hold with $\lambda := \sup_{s\in(0,\lambda_1]} \big(s - \mathrm{e}^{sr_0}\lambda'\big)$.

Examples 1.2. Let $\mathbb{H}_1 = \mathbb{H}_2 = L^2(D; \mathrm{d}x)$ and $\sigma = B = I, A_1 = A_2 = -(-\Delta)^\alpha$ for some $\alpha > \frac{d}{2}$ as in Example 1.1. Then assumptions **(B1)**, **(B3)** and **(B4)** hold. See [15] for more examples, where $\mathbb{H}_2$ might be a subspace of $\mathbb{H}_1$. To verify **(B2)** we take, for instances,

$$b(\xi, \eta) = K_1 \int_{-r_0}^{0} \xi \mathrm{d}\nu_1 + K_2 \int_{-r_0} \eta \mathrm{d}\nu_2$$

for some signed measures ν_1, ν_2 on $[-r_0, 0]$ with total variations not larger than 1; or simply $b(\xi, \eta) = \|K_1\xi + K_2\eta\|_\infty$, where interactions exist between ξ and η.

The remainder of this paper is organized as follows. In Section 2 we present a Fernique type inequality for infinite-dimensional Gaussian processes, which will be used to prove the concentration condition required in Theorem 1.1(3). Theorems 1.2 and 1.3 are proved in Sections 3 and 4, respectively.

2 Infinite-dimensional Fernique's inequality

In [8], Fernique introduced an inequality for the distribution of the maximum of Gaussian processes. To prove the exponential integrability of $\|X_t\|_\infty$ for FSPDEs, one needs an infinite-dimensional version of this inequality. However, as the dimension goes to infinity, existing Fernique's inequality for multi-dimensional Gaussian processes becomes invalid. So, we modify the inequality so that it holds also in infinite-dimensions. To this end, we first recall the inequality for one-dimensional Gaussian processes (see, e.g., [4, page 49] for the multi-dimensional case).

Lemma 2.1 (Fernique's inequality)**.** Let $\{\gamma(t)\}_{t\in[0,1]}$ be a continuous Gaussian process on $\mathbb{R}$ with zero mean and $\Gamma = \sup_{t\in[0,1]}(\mathbb{E}\gamma(t)^2)^{\frac{1}{2}} < \infty$. Let

$$\phi(r) = \sup_{s,t\in[0,1],|s-t|\le r} \left(\mathbb{E}|\gamma(s) - \gamma(t)|^2\right)^{\frac{1}{2}}, \quad r \in [0,1].$$

If $\theta := \int_1^\infty \phi(\mathrm{e}^{-s^2})\mathrm{d}s < \infty$, then

$$\mathbb{P}\Big(\max_{t\in[0,1]} |\gamma(t)| \ge r\big(\Gamma + \big(2+\sqrt{2}\big)\theta\big)\Big) \le \frac{5\mathrm{e}}{2} \int_r^\infty \mathrm{e}^{-\frac{1}{2}s^2}\mathrm{d}s, \quad r \ge \sqrt{5}.$$

Now, we call a process $\{\gamma(t)\}_{t\in[0,1]}$ on the Hilbert space $\mathbb{H}$ a cylindrical continuous Gaussian process, if, for an orthonormal basis $\{e_i\}_{i\ge1}$, every one-dimensional process $\gamma_i(t) := \langle\gamma(t), e_i\rangle$ is a continuous Gaussian process. For a cylindrical continuous Gaussian process $\gamma(t)$ with zero mean, let

$$\phi_i(r) = \sup_{s,t\in[0,1],|s-t|\le r} \left(\mathbb{E}|\gamma_i(t) - \gamma_i(s)|^2\right)^{\frac{1}{2}}, \quad r \in [0,1],$$
$$\Gamma_i = \sup_{t\in[0,1]} (\mathbb{E}\gamma_i(t)^2)^{\frac{1}{2}}, \quad \delta_i = \Gamma_i + \big(2+\sqrt{2}\big) \int_1^\infty \phi_i(\mathrm{e}^{-s^2})\mathrm{d}s, \quad i \ge 1.$$

Theorem 2.2. Let $\gamma(t)$ be a cylindrical continuous Gaussian process on $\mathbb{H}$ with zero mean such that

$$\theta := \sum_{i=1}^\infty \delta_i^2 \log(\mathrm{e} + \delta_i^{-1}) < \infty. \tag{2.1}$$

Then, for any positive constant $\lambda < \min_{i\ge1} \frac{\log(\mathrm{e}+\delta_i^{-1})}{2\theta}$, there exists a constant $c > 0$ such that

$$\mathbb{P}\Big(\max_{t\in[0,1]} |\gamma(t)| \ge r\Big) \le c\mathrm{e}^{-\lambda r^2}, \quad r \ge 0. \tag{2.2}$$

Proof. Let $\widetilde{\lambda} = \min_{i\geq 1} \frac{\log(\mathrm{e}+\delta_i^{-1})}{2\theta}$. Obviously, (2.1) implies $\lim_{i\to\infty} \delta_i = 0$ so that $\widetilde{\lambda} > 0$. For any $\lambda \in (0, \widetilde{\lambda})$, it suffices to prove (2.2) for some constant $c > 0$ and large enough $r > 0$. Below, we assume that

$$r^2 \geq \frac{5\theta\widetilde{\lambda}}{\widetilde{\lambda} - \lambda}. \tag{2.3}$$

In this case,

$$r_i := \Big(\frac{r^2 \log(\mathrm{e} + \delta_i^{-1})}{\theta}\Big)^{\frac{1}{2}} \geq \frac{r}{\sqrt{\theta}} \geq \sqrt{5},$$

so that Lemma 2.1 implies

$$\mathbb{P}\Big(\max_{t\in[0,1]} |\gamma_i(t)| \geq r_i\delta_i\Big) \leq \frac{5\mathrm{e}}{2}\int_{r_i}^{\infty} \mathrm{e}^{-\frac{1}{2}s^2}\mathrm{d}s \leq c_1\mathrm{e}^{-\frac{1}{2}r_i^2}, \;\; i \geq 1$$

for some constant $c_1 > 0$. Then

$$\begin{aligned}
&\mathbb{P}\Big(\max_{t\in[0,1]} |\gamma(t)| \geq r\Big) \leq \mathbb{P}\Big(\sum_{i=1}^{\infty} \max_{t\in[0,1]} |\gamma_i(t)|^2 \geq r^2\Big)\\
&\leq \sum_{i=1}^{\infty} \mathbb{P}\Big(\max_{t\in[0,1]} |\gamma_i(t)|^2 \geq \frac{r^2\delta_i^2 \log(\mathrm{e} + \delta_i^{-1})}{\theta}\Big) = \sum_{i=1}^{\infty} \mathbb{P}\Big(\max_{t\in[0,1]} |\gamma_i(t)| \geq r_i\delta_i\Big) \\
&\leq c_1 \sum_{i=1}^{\infty} \mathrm{e}^{-\frac{1}{2}r_i^2} \leq c_1\mathrm{e}^{-\lambda r^2} \sum_{i=1}^{\infty} \exp\Big[-r^2\Big(\frac{\log(\mathrm{e} + \delta_i^{-1})}{2\theta} - \lambda\Big)\Big].
\end{aligned} \tag{2.4}$$

Since, by (2.3) and the definition of $\widetilde{\lambda}$, we have

$$r^2\Big(\frac{\log(\mathrm{e} + \delta_i^{-1})}{2\theta} - \lambda\Big) \geq \frac{r^2 \log(\mathrm{e} + \delta_i^{-1})}{2\theta}\Big(1 - \frac{\lambda}{\widetilde{\lambda}}\Big) \geq \frac{5}{2}\log(\mathrm{e} + \delta_i^{-1}),$$

it follows from (2.1) that

$$\sum_{i=1}^{\infty} \exp\Big[-r^2\Big(\frac{\log(\mathrm{e} + \delta_i^{-1})}{2\theta} - \lambda\Big)\Big] \leq \sum_{i=1}^{\infty} \delta_i^{\frac{5}{2}} < \infty.$$

Combining this with (2.4), we finish the proof. □

3 Proof of Theorem 1.2

We will verify conditions (i)-(iii) in Theorem 1.1. Firstly, according to [14, Theorem 4.2.4], assumptions **(A1)**-**(A3)** implies that, for any $t_0 > r_0$, there exists a constant $c_0 > 0$ such that the following Harnack inequality holds:

$$\big(P_{t_0} f(\eta)\big)^2 \leq (P_{t_0} f^2(\xi))\mathrm{e}^{c_0\|\xi-\eta\|_\infty^2}, \quad \xi, \eta \in \mathscr{C}, f \in \mathscr{B}_b(\mathscr{C}). \tag{3.1}$$

That is, condition (i) holds for $\rho(\xi, \eta) := \|\xi - \eta\|_\infty$.

To verify (ii) and (iii), we will need the condition that $\lambda := \sup_{s\in(0,\lambda_1]}(s - L\mathrm{e}^{sr_0}) > 0$. Without loss of generality, we may and do assume that the maximum is attained at the point λ_1; otherwise, in the following it suffices to replace λ_1 by $\lambda_1' \in (0, \lambda_1]$ which attains the maximum. By **(A1)**, **(A2)**, and (1.4), one has

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

Then, we obtain that

$$\begin{aligned}\mathrm{e}^{\lambda_1 t}\|X_t^\xi - X_t^\eta\|_\infty &\le \mathrm{e}^{\lambda_1 r_0}\sup_{-r_0\le\theta\le 0}(\mathrm{e}^{\lambda_1(t+\theta)}|X^\xi(t+\theta)-X^\eta(t+\theta)|)\\ &\le \mathrm{e}^{\lambda_1 r_0}\Big(\|\xi-\eta\|_\infty + L\int_0^t \mathrm{e}^{\lambda_1 s}\|X_s^\xi - X_s^\eta\|_\infty \mathrm{d}s\Big).\end{aligned} \tag{3.2}$$

Thus, by Gronwall's inequality we derive that

$$\|X_t^\xi - X_t^\eta\|_\infty \le \mathrm{e}^{\lambda_1 r_0}\mathrm{e}^{-\lambda t}\|\xi-\eta\|_\infty,\quad t\ge 0,\quad \xi,\eta\in\mathscr{C}. \tag{3.3}$$

That is, condition (ii) holds.

To show condition (iii) in Theorem 1.1, we need to prove the exponential integrability of the segment solution.

Lemma 3.1. Assume **(A1)** and **(A2)**. If $\lambda > 0$, then there exists an $r > 0$ such that

$$\sup_{t\ge 0}\mathbb{E}\mathrm{e}^{r\|X_t^\xi\|_\infty^2} < \infty,\quad \xi\in\mathscr{C}. \tag{3.4}$$

Proof. (a) We first use Theorem 2.2 to prove

$$\sup_{t\ge 0}\mathbb{E}\mathrm{e}^{\varepsilon\|Z_t\|_\infty^2} < \infty \tag{3.5}$$

for some $\varepsilon > 0$, where

$$Z_t(\theta) := \int_0^{(t+\theta)^+}\mathrm{e}^{(t+\theta-s)A}\sigma\mathrm{d}W(s),\quad t\ge 0,\ \theta\in[-r_0,0]. \tag{3.6}$$

To this end, for fixed $t_0 > 0$ let

$$\gamma(t) = \int_0^{(t_0-tr_0)^+}\mathrm{e}^{(t_0-tr_0-s)A}\sigma\mathrm{d}W(s),\quad t\in[0,1].$$

Then, (3.6) implies

$$\|Z_{t_0}\|_\infty^2 = \sup_{t\in[0,1]}|\gamma(t)|^2. \tag{3.7}$$

Letting $\{e_i\}_{i\ge 1}$ be the eigenbasis of A, we have

$$\gamma_i(t) := \langle\gamma(t),e_i\rangle = \int_0^{(t_0-tr_0)^+}\mathrm{e}^{-\lambda_i(t_0-tr_0-s)}\langle\sigma^* e_i,\mathrm{d}W(s)\rangle,\quad t\in[0,1]. \tag{3.8}$$

Obviously,

$$\Gamma_i := \sup_{t\in[0,1]}\big(\mathbb{E}\gamma_i(t)^2\big)^{\frac12} \le |\sigma^* e_i|\bigg(\int_0^\infty \mathrm{e}^{-2\lambda_i s}\mathrm{d}s\bigg)^{\frac12} = \frac{|\sigma^* e_i|}{\sqrt{2\lambda_i}},\quad i\ge 1. \tag{3.9}$$

Moreover, note that, for any $r\in(0,1)$, there exists a constant $c(r) > 0$ such that $|\mathrm{e}^{-s}-\mathrm{e}^{-t}|\le c(r)|s-t|^r$ holds for all $s,t\ge 0$. Then, (3.8), implies that for any $0\le t'\le t\le 1$,

$$\begin{aligned}&\mathbb{E}|\gamma_i(t)-\gamma_i(t')|^2\\ &= |\sigma^* e_i|^2\bigg(\int_0^{(t_0-tr_0)^+}\mathrm{e}^{-2\lambda_i(t_0-tr_0-s)}\big(1-\mathrm{e}^{-\lambda_i(t-t')r_0}\big)^2\mathrm{d}s + \int_{(t_0-tr_0)^+}^{(t_0-t'r_0)^+}\mathrm{e}^{-2\lambda_i(t_0-t'r_0-s)}\mathrm{d}s\bigg)\\ &\le |\sigma^* e_i|^2\bigg(\frac{c(\frac{\delta}{4})^2[r_0(t-t')]^{\frac{\delta}{2}}}{2\lambda_i^{1-\frac{\delta}{2}}} + \frac{c(\frac{\delta}{2})[2r_0(t-t')]^{\frac{\delta}{2}}}{2\lambda_i^{1-\frac{\delta}{2}}}\bigg)\\ &=: \frac{c_1(t-t')^{\frac{\delta}{2}}|\sigma^* e_i|^2}{\lambda_i^{1-\frac{\delta}{2}}},\quad i\ge 1,\end{aligned}$$

where the constant $c_1 > 0$ is independent of t, t', t_0 and i. So, by the definition of ϕ_i,

$$\phi_i(r) \le \frac{c_1^{1/2} r^{\frac{\delta}{4}} |\sigma^* e_i|}{\lambda_i^{\frac{1}{2}-\frac{\delta}{4}}}, \quad r \in [0,1].$$

Combining this with (3.9), we deduce from the definition of δ_i that

$$\delta_i \le \frac{c_2 |\sigma^* e_i|}{\lambda_i^{\frac{1}{2}-\frac{\delta}{4}}}, \quad i \ge 1$$

holds for some constant $c_2 > 0$ independent of t_0. This and (1.6) lead to (2.1). Therefore, according to Theorem 2.2 and (3.7), we prove (3.5) for some constant $\varepsilon \in (0,1)$.

(b) Next, we prove (3.4) for small $r > 0$. By (3.3), it suffices to prove for $\xi \equiv 0$. We simply denote $X(t) = X^0(t)$. It follows from **(A1)**, **(A2)**, and (1.4) that

$$\mathrm{e}^{\lambda_1 t}|X(t)| \le \int_0^t \mathrm{e}^{\lambda_1 s}\{c_0 + L\|X_s\|_\infty\}\mathrm{d}s + \mathrm{e}^{\lambda_1 t}\Big|\int_0^t \mathrm{e}^{(t-s)A}\sigma \mathrm{d}W(s)\Big|, \quad t \ge 0$$

holds for some constant $c_0 > 0$. This implies

$$\begin{aligned}
\mathrm{e}^{\lambda_1 t}\|X_t\|_\infty &\le \mathrm{e}^{\lambda_1 r_0} \sup_{-r_0 \le \theta \le 0} (\mathrm{e}^{\lambda_1 (t+\theta)}|X(t+\theta)|) \\
&\le c_1 \mathrm{e}^{\lambda_1 t}(1 + \|Z_t\|_\infty) + L\mathrm{e}^{\lambda_1 r_0}\int_0^t \mathrm{e}^{\lambda_1 s}\|X_s\|_\infty \mathrm{d}s
\end{aligned}$$

for some constant $c_1 > 0$, where Z_t is defined in (3.6). So, by Gronwall's formula,

$$\begin{aligned}
\|X_t\|_\infty &\le c_1(1 + \|Z_t\|_\infty) + c_1 L \mathrm{e}^{\lambda_1 r_0}\mathrm{e}^{-\lambda_1 t}\int_0^t \Big\{\mathrm{e}^{\lambda_1 s} + \mathrm{e}^{\lambda_1 s}\|Z_s\|_\infty\Big\}\mathrm{e}^{L\mathrm{e}^{\lambda_1 r_0}(t-s)}\mathrm{d}s \\
&\le c_2(1 + \|Z_t\|_\infty) + c_2 \int_0^t \|Z_s\|_\infty \mathrm{e}^{-\lambda(t-s)}\mathrm{d}s
\end{aligned}$$

holds for some constant $c_2 > 0$, where $\lambda = \lambda_1 - L\mathrm{e}^{\lambda_1 r_0} > 0$ as assumed above. Thus, using Hölder's inequality and applying Jensen's inequality for the probability measure $\nu(\mathrm{d}s) := \frac{\lambda \mathrm{e}^{-\lambda(t-s)}}{1-\mathrm{e}^{-\lambda t}}\mathrm{d}s$ on $[0,t]$, we obtain

$$\begin{aligned}
\mathbb{E}\mathrm{e}^{r\|X_t\|_\infty^2} &\le \mathrm{e}^{c_3}\big(\mathbb{E}\mathrm{e}^{c_3 r\|Z_t\|_\infty^2}\big)^{\frac{1}{2}}\left(\mathbb{E}\exp\left[c_3 r\left(\frac{(1-\mathrm{e}^{-\lambda t})}{\lambda}\int_0^t \|Z_s\|_\infty \nu(\mathrm{d}s)\right)^2\right]\right)^{\frac{1}{2}} \\
&\le \mathrm{e}^{c_3}\big(\mathbb{E}\mathrm{e}^{c_3 r\|Z_t\|_\infty^2}\big)^{\frac{1}{2}}\left(\int_0^t \mathbb{E}\exp\left[\frac{c_3 r}{\lambda^2}\|Z_s\|_\infty^2\right]\nu(\mathrm{d}s)\right)^{\frac{1}{2}} \\
&\le \mathrm{e}^{c_3}\sup_{s\ge 0}\mathbb{E}\exp\left[\frac{c_3 r}{1\wedge\lambda^2}\|Z_s\|_\infty^2\right], \quad t \ge 0, r > 0
\end{aligned} \tag{3.10}$$

for some constant $c_3 > 0$. Thus, when $r > 0$ is small enough, (3.4) follows from (3.5). □

Now, we are in position check condition (iii) in theorem 1.1.

Lemma 3.2. Assume **(A1)** and **(A2)**. If $\lambda > 0$, then P_t admits a unique invariant measure μ. Moreover, $\mu(\mathrm{e}^{\varepsilon\|\cdot\|_\infty^2}) < \infty$ for some $\varepsilon > 0$.

Proof. The proof is similar to that of [3, Lemma 2.4]. Let μ_t^ξ be the law of X_t^ξ. Note that if μ_t^ξ converges weakly to a probability measure μ^ξ as $t \to \infty$, then μ^ξ is an invariant probability measure of P_t (see, e.g., [6, Theorem 3.1.1]. Let $\mathscr{P}(\mathscr{C})$ be the set of all probability measures on $\mathscr{C}$. Consider the L^1-Wasserstein distance W induced by $\rho(\xi,\eta) := 1 \wedge \|\xi - \eta\|_\infty$, i.e.,

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

where $\mathscr{C}(\mu_1,\mu_2)$ is the set of all couplings for 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 (see, e.g., [5, Lemma 5.3 and Lemma 5.4]), and the topology induced by W coincides with the weak topology (see, e.g., [5, Theorem 5.6]). So, to show existence of an invariant measure, it is sufficient to prove that μ_t^ξ is a W-Cauchy sequence as $t\to\infty$, i.e.,

$$\lim_{t_1,t_2\to\infty} W(\mu_{t_1}^\xi,\mu_{t_2}^\xi)=0. \tag{3.11}$$

For any $t_2>t_1>0$, consider the following SPDEs

$$\mathrm{d}X(t)=\{AX(t)+b(X_t)\}\mathrm{d}t+\sigma\mathrm{d}W(t),\ t\in[0,t_2],\quad X_0=\xi,$$

and

$$\mathrm{d}Y(t)=\{AY(t)+b(Y_t)\}\mathrm{d}t+\sigma\mathrm{d}W(t),\ t\in[t_2-t_1,t_2],\quad Y_{t_2-t_1}=\xi.$$

Then, the laws of $X_{t_2}(\xi)$ and $Y_{t_2}(\xi)$ are $\mu_{t_2}^\xi$ and $\mu_{t_1}^\xi$, respectively. Also, following an argument leading to derive (3.2), we obtain

$$\mathrm{e}^{\lambda_1 t}\mathbb{E}\|X_t-Y_t\|_\infty^2\le c_1\mathbb{E}\|X_{t_2-t_1}-\xi\|_\infty^2+L\mathrm{e}^{\lambda_1 r_0}\int_{t_2-t_1}^t \mathrm{e}^{\lambda_1 s}\mathbb{E}\|X_s-Y_s\|_\infty^2\mathrm{d}s,\ t\in[t_2-t_1,t_2]$$

for some constant $c_1>0$. By Gronwall's inequality and $\lambda=\lambda_1-L\mathrm{e}^{\lambda_1 r_0}>0$ as assumed above, this implies

$$\mathbb{E}\|X_t-Y_t\|_\infty^2\le c_1\mathrm{e}^{-\lambda(t-t_2+t_1)}\mathbb{E}\|X_{t_2-t_1}-\xi\|_\infty^2,\quad t\in[t_2-t_1,t_2].$$

Combining this with (3.4) yields

$$\mathbb{E}\|X_{t_2}-Y_{t_2}\|_\infty^2\le c_2\mathrm{e}^{-\lambda t_1}$$

so that

$$W(\mu_{t_1}^\xi,\mu_{t_2}^\xi)\le\mathbb{E}\|X_{t_2}-Y_{t_2}\|_\infty\le\sqrt{c_2}\mathrm{e}^{-\frac{\lambda t_1}{2}}.$$

Therefore, (3.11) holds, and, by the completeness of W, there exists $\mu^\xi\in\mathscr{P}(\mathscr{C})$ such that

$$\lim_{t\to\infty}W(\mu_t^\xi,\mu^\xi)=0. \tag{3.12}$$

To prove the uniqueness, it suffices to show that μ^ξ is independent of $\xi\in\mathscr{C}$. This follows since, by the triangle inequality, (3.3) and (3.12),

$$W(\mu^\xi,\mu^\eta)\le\lim_{t\to\infty}\big\{W(\mu_t^\xi,\mu^\xi)+W(\mu_t^\eta,\mu^\eta)+W(\mu_t^\xi,\mu_t^\eta)\big\}=0,\quad \xi,\eta\in\mathscr{C}.$$

Finally, since $\mu_t^0\to\mu$ weakly as $t\to\infty$, by (3.4) we have

$$\mu(\mathrm{e}^{r\|\cdot\|_\infty^2})=\lim_{N\to\infty}\mu(N\wedge\mathrm{e}^{r\|\cdot\|_\infty^2})=\lim_{N\to\infty}\lim_{t\to\infty}\mathbb{E}(N\wedge\mathrm{e}^{r\|X_t^0\|_\infty^2})<\infty.$$

Thus, the proof is finished. □

With the above preparations, we present below a proof of Theorem 1.2.

Proof of Theorem 1.2. According to Theorem 1.1, the first two assertions follow from (3.1), (3.2) and Lemma 3.2. It remains to prove the last assertion. According to [13, Proposition 2.2], the Harnack inequality (3.1) implies the log-Harnack inequality

$$P_{t_0}\log f(\xi)\le\log P_{t_0}f(\eta)+\frac{c_0}{2}\|\xi-\eta\|_\infty^2,\quad 0<f\in\mathscr{B}_b(\mathscr{C}),\ \xi,\eta\in\mathscr{C}.$$

By [2, Proposition 2.3], this implies

$$|P_{t_0}f(\xi)-P_{t_0}f(\eta)|^2\le c_0\|\xi-\eta\|_\infty^2\|f\|_\infty^2,\ \ f\in\mathscr{B}_b(\mathscr{C}),\ \xi,\eta\in\mathscr{C}.$$

Combining this with the Markov property, we obtain

$$\|\mu_{t_0+t}^\xi-\mu_{t_0+t}^\eta\|\text{var}\le 2\sup_{\|f\|_\infty\le 1}\mathbb{E}|P_{t_0}f(X_t^\xi)-P_{t_0}f(X_t^\eta)|\le 2\sqrt{c_0}\,\mathbb{E}\|X_t^\xi-X_t^\eta\|_\infty,\ \ t\ge 0.$$

Therefore, the last assertion follows from (3.3). □

4 Proof of Theorem 1.3

According to what we have done in the last section for the proof of Theorem 1.2, it suffices to verify the existence and uniqueness of the invariant probability measure, as well as conditions (i)-(iii) in Theorem 1.1. In the present setting we have to pay more attention on the degenerate part. In particular, the known Harnack inequality (see [14, Corollary 4.4.4]) does not meet our requirement as the exponential term in the upper bound is not integrable with respect to the invariant probability measure. So, we first establish the following Harnack inequality which extends the corresponding one in [15] for the case without delay. The proof is modified from [15] using the coupling by change measures. This method was introduced in [1] on manifolds and further developed in [12] for SPDEs and in [7] for SDDEs, see [14] for a self-contained account on coupling by change of measures and applications.

Lemma 4.1. Assume **(B1)**-**(B4)**. Then, for any $t_0>r_0$, there exists a constant $c>0$ such that

$$(P_{t_0}f(\overline{\xi},\overline{\eta}))^2\le \mathrm{e}^{c(\|\xi-\overline{\xi}\|_\infty^2+\|\eta-\overline{\eta}\|_\infty^2)}P_{t_0}f^2(\xi,\eta),\quad (\xi,\eta),\ (\overline{\xi},\overline{\eta})\in\mathscr{C},\ f\in\mathscr{B}_b(\mathscr{C}).\tag{4.1}$$

Proof. Let $(X(t),Y(t))=(X^{\xi,\eta}(t),Y^{\xi,\eta}(t))$ for $t\ge 0$, and let $(\overline{X}(t),\overline{Y}(t))$ solve the following equation for $(\overline{X}_0,\overline{Y}_0)=(\overline{\xi},\overline{\eta})$:

$$\begin{cases}\mathrm{d}\overline{X}(t)=\{A_1\overline{X}(t)+B\overline{Y}(t)\}\mathrm{d}t,\\ \mathrm{d}\overline{Y}(t)=\Big\{A_2\overline{Y}(t)+b(X_t,Y_t)+\frac{1_{[0,t_0-r_0]}(t)}{t_0-r_0}\mathrm{e}^{tA_2}(\eta(0)-\overline{\eta}(0))+\mathrm{e}^{tA_2}h'(t)\Big\}\mathrm{d}t+\sigma\mathrm{d}W(t),\end{cases}$$

where

$$h(t):=t(t_0-r_0-t)^+B^*\mathrm{e}^{-tA_0^*}e,\quad t\in[0,t_0]\tag{4.2}$$

for A_0 in **(B4)** and some $e\in\mathbb{H}_1$ to be determined. Obviously,

$$\overline{Y}(t)-Y(t)=\mathrm{e}^{tA_2}\Big\{\frac{(\overline{\eta}(0)-\eta(0))(t_0-r_0-t)^+}{t_0-r_0}+h(t)\Big\},\quad t\in[0,t_0].\tag{4.3}$$

In particular, we have $\overline{Y}_{t_0}=Y_{t_0}$. Next, the equations of $X(t)$ and $\overline{X}(t)$ yield

$$\overline{X}(t)-X(t)=\mathrm{e}^{tA_1}(\overline{\xi}(0)-\xi(0))+\int_0^t\mathrm{e}^{(t-s)A_1}B(\overline{Y}(s)-Y(s))\mathrm{d}s.\tag{4.4}$$

Substituting (4.3) into (4.4), we find that

$$\overline{X}(t)-X(t)=\mathrm{e}^{tA_1}(\overline{\xi}(0)-\xi(0))+\int_0^t\mathrm{e}^{(t-s)A_1}B\mathrm{e}^{sA_2}\Big\{\frac{(\overline{\eta}(0)-\eta(0))(t_0-r_0-s)^+}{t_0-r_0}+h(s)\Big\}\mathrm{d}s.$$

By virtue of **(B4)** and the definition of h, this implies

$$\begin{aligned}
&\overline{X}(t) - X(t)\\
&= \mathrm{e}^{tA_1}(\bar{\xi}(0) - \xi(0)) + \int_0^t \mathrm{e}^{(t-s)A_1}\mathrm{e}^{sA_1}\mathrm{e}^{sA_0}B\Big\{\frac{(\overline{\eta}(0)-\eta(0))(t_0-r_0-s)^+}{t_0-r_0} + h(s)\Big\}\mathrm{d}s\\
&= \mathrm{e}^{tA_1}\Big(\bar{\xi}(0) - \xi(0) + \int_0^t \mathrm{e}^{sA_0}B\Big\{\frac{(\overline{\eta}(0)-\eta(0))(t_0-r_0-s)^+}{t_0-r_0} + h(s)\Big\}\mathrm{d}s\Big)\\
&= \mathrm{e}^{tA_1}\Big(\bar{\xi}(0) - \xi(0) + \int_0^{t_0-r_0} \mathrm{e}^{sA_0}B\Big\{\frac{(\overline{\eta}(0)-\eta(0))(t_0-r_0-s)}{t_0-r_0} + h(s)\Big\}\mathrm{d}s\Big)
\end{aligned} \tag{4.5}$$

for any $t \in [t_0 - r_0, t_0]$. Moreover, **(B4)** implies that

$$\widetilde{Q}_{t_0-r_0} := \int_0^{t_0-r_0} s(t_0-r_0-s)\mathrm{e}^{sA_0}BB^*\mathrm{e}^{sA_0^*}\mathrm{d}s$$

is invertible on $\mathbb{H}_1$. In (4.2), in particular, take

$$e = -\widetilde{Q}_{t_0-r_0}^{-1}\Big\{\bar{\xi}(0) - \xi(0) + \int_0^{t_0-r_0} \frac{t_0-r_0-s}{t_0-r_0}\mathrm{e}^{sA_0}B(\overline{\eta}(0)-\eta(0))\mathrm{d}s\Big\}.$$

Then, inserting $h(\cdot)$ back into (4.5) leads to $\overline{X}(t) = X(t)$ for arbitrary $t \in [t_0 - r_0, t_0]$, i.e., $\overline{X}_{t_0} = X_{t_0}$. Therefore, we arrive at $(X_{t_0}, Y_{t_0}) = (\overline{X}_{t_0}, \overline{Y}_{t_0})$.

Let

$$\widetilde{W}(t) = W(t) + \int_0^t \phi(s)\mathrm{d}s, \quad t \in [0, t_0],$$

where

$$\phi(t) := \sigma^{-1}\Big(b(X_t, Y_t) - b(\overline{X}_t, \overline{Y}_t) + \frac{1_{[0,t_0-r_0]}(t)}{t_0-r_0}\mathrm{e}^{tA_2}(\eta(0)-\overline{\eta}(0)) + \mathrm{e}^{tA_2}h'(t)\Big).$$

By (4.3) and (4.5), for some constant $C > 0$ we have

$$\|X_t - \overline{X}_t\|_\infty^2 + \|Y_t - \overline{Y}_t\|_\infty^2 \le C(\|\xi - \bar{\xi}\|_\infty^2 + \|\eta - \overline{\eta}\|_\infty^2), \quad t \in [0, t_0]. \tag{4.6}$$

Thus, by the Girsanov theorem (see, e.g., [6, Theorem 10.14]), $\{\widetilde{W}(s)\}_{t\in[0,T]}$ is a cylindrical Wiener process under the weighted probability measure $\mathrm{d}\mathbb{Q} := R\mathrm{d}\mathbb{P}$ with

$$R := \exp\Big(-\int_0^{t_0}\langle\phi(s), \mathrm{d}W(s)\rangle - \frac{1}{2}\int_0^{t_0}|\phi(s)|^2\mathrm{d}s\Big).$$

Now, we reformulate the equation for $(\overline{X}(t), \overline{Y}(t))$ as

$$\begin{cases}\mathrm{d}\overline{X}(t) = \{A_1\overline{X}(t) + B\overline{Y}(t)\}\mathrm{d}t,\\ \mathrm{d}\overline{Y}(t) = \{A_2\overline{Y}(t) + b(X_t, Y_t)\}\mathrm{d}t + \sigma\mathrm{d}\widetilde{W}(t), \quad t \in [0, t_0].\end{cases}$$

Then, invoking the weak uniqueness of the equation and using $(\overline{X}_{t_0}, \overline{Y}_{t_0}) = (X_{t_0}, Y_{t_0})$, we derive that

$$\begin{aligned}(P_{t_0}f(\bar{\xi}, \overline{\eta}))^2 &= \big\{\mathbb{E}_{\mathbb{Q}}f(\overline{X}_{t_0}, \overline{Y}_{t_0})\big\}^2 = \big\{\mathbb{E}(Rf(X_{t_0}, Y_{t_0})\big\}^2\\ &\le (\mathbb{E}R^2)\mathbb{E}f^2(X_{t_0}, Y_{t_0}) = (\mathbb{E}R^2)P_{t_0}f^2(\xi, \eta).\end{aligned}$$

Combining this with (4.6) and the definitions of R and ϕ, we prove (4.1) for some constant $c > 0$. □

Next, the following lemma verifies condition (ii) in Theorem 1.1. As explained in Section 3 that we may and do assume $\lambda = \lambda_1 - \lambda' e^{\lambda_1 r_0} > 0$; otherwise in the sequel it suffices to replace λ_1 by $\lambda_1' \in (0, \lambda_1]$ which attains the maximum in the definition of λ.

Lemma 4.2. Assume **(B1)**-**(B3)** and let (1.9) hold. Then there exists $c > 0$ such that for $\lambda := \sup_{s\in(0,\lambda_1]}(s - \lambda' e^{sr_0}) > 0$,

$$\|X_t^{\xi,\eta} - X_t^{\bar\xi,\bar\eta}\|_\infty + \|Y_t^{\xi,\eta} - Y_t^{\bar\xi,\bar\eta}\|_\infty \le c(\|\xi - \bar\xi\|_\infty + \|\eta - \bar\eta\|_\infty)e^{-\lambda t} \tag{4.7}$$

for any $t \ge 0, (\xi,\eta), (\bar\xi,\bar\eta) \in \mathscr{C}$.

Proof. By **(B1)**-**(B3)**, we have

$$\begin{aligned} & e^{\lambda_1 t}|X^{\xi,\eta}(t) - X^{\bar\xi,\bar\eta}(t)| - |\xi(0) - \bar\xi(0)| \\ & \le \int_0^t e^{\lambda_1 s}\{\delta|X^{\xi,\eta}(s) - X^{\bar\xi,\bar\eta}(s)| + \|B\|\cdot|Y^{\xi,\eta}(s) - Y^{\bar\xi,\bar\eta}(s)|\}ds, \\ & e^{\lambda_1 t}|Y^{\xi,\eta}(t) - Y^{\bar\xi,\bar\eta}(t)| - |\eta(0) - \bar\eta(0)| \\ & \le \int_0^t e^{\lambda_1 s}\{K_1\|X_s^{\xi,\eta} - X_s^{\bar\xi,\bar\eta}\|_\infty + K_2\|Y_s^{\xi,\eta} - Y_s^{\bar\xi,\bar\eta}\|_\infty\}ds. \end{aligned} \tag{4.8}$$

Next, let

$$\alpha = \frac{\delta - K_2 + \sqrt{(K_2-\delta)^2 + 4K_1\|B\|}}{2\|B\|}. \tag{4.9}$$

It is easy to see that $\alpha > 0$ and, for $\lambda' > 0$ defined in (1.9), we have

$$\alpha\delta + K_1 = \lambda'\alpha, \quad \alpha\|B\| + K_2 = \lambda'. \tag{4.10}$$

Combining (4.8), (4.9) with (4.10), we derive

$$\begin{aligned} & e^{\lambda_1 t}(\alpha\|X_t^{\xi,\eta} - X_t^{\bar\xi,\bar\eta}\|_\infty + \|Y_t^{\xi,\eta} - Y_t^{\bar\xi,\bar\eta}\|_\infty) \\ & \quad \le e^{\lambda_1 r_0}\Big\{\alpha\|\xi - \bar\xi\|_\infty + \|\eta - \bar\eta\|_\infty \\ & \quad + \int_0^t e^{\lambda_1 s}((\delta\alpha + K_1)\|X_s^{\xi,\eta} - X_s^{\bar\xi,\bar\eta}\|_\infty + (\alpha\|B\| + K_2)\|Y_s^{\xi,\eta} - Y_s^{\bar\xi,\bar\eta}\|_\infty)ds\Big\} \\ & \quad \le e^{\lambda_1 r_0}\Big\{\alpha\|\xi - \bar\xi\|_\infty + \|\eta - \bar\eta\|_\infty \\ & \quad + \lambda'\int_0^t e^{\lambda_1 s}(\alpha\|X_s^{\xi,\eta} - X_s^{\bar\xi,\bar\eta}\|_\infty + \|Y_s^{\xi,\eta} - Y_s^{\bar\xi,\bar\eta}\|_\infty)ds\Big\}. \end{aligned}$$

Therefore, we complete the proof by using Gronwall's inequality and $\lambda = \lambda_1 - \lambda' e^{\lambda_1 r_0} > 0$ as assumed above. ☐

Moreover, corresponding to Lemma 3.1 for the non-degenerate case, we have the following result on the exponential integrability of the segment solution.

Lemma 4.3. Assume **(B1)**-**(B3)** and let (1.9) hold. Then there exists a constant $\varepsilon > 0$ such that

$$\sup_{t\ge 0} \mathbb{E} e^{\varepsilon(\|X_t^{\xi,\eta}\|_\infty^2 + \|Y_t^{\xi,\eta}\|_\infty^2)} < \infty, \quad (\xi,\eta) \in \mathscr{C}.$$

Proof. By Lemma 4.2, it suffices to prove for $(\xi,\eta) \equiv (0,0)$. Simply denote $(X_t, Y_t) = (X_t^{0,0}, Y_t^{0,0})$. We have

$$X(t) = \int_0^t e^{(A_1-\delta)(t-s)}\big(\delta X(s) + BY(s)\big)ds, \quad t \ge 0.$$

Then, **(B3)** yields

$$e^{\lambda_1 t}|X(t)| \le \int_0^t e^{\lambda_1 s}\{\|B\|\cdot|Y(s)| + \delta|X(s)|\}ds. \tag{4.11}$$

Next, according to **(B1)** and **(B2)**, it follows that

$$e^{\lambda_1 t}|Y(t)| \le \int_0^t e^{\lambda_1 s}\{c_0 + K_1\|X_s\|_\infty + K_2\|Y_s\|_\infty\}ds + e^{\lambda_1 t}\Big|\int_0^t e^{A_2(t-s)}\sigma dW(s)\Big| \tag{4.12}$$

holds for $c_0 := |b(0,0)|$. Obviously, using $(\mathbb{H}_2, A_2)$ to replace $(\mathbb{H}, A)$, we see that (3.5) holds for

$$Z_t(\theta) := \int_0^{(t+\theta)^+} e^{A_2(t-s)}\sigma dW(s), \quad \theta \in [-r_0, 0].$$

Combining (4.10), (4.11) with (4.12), for the present Z_t we have

$$\begin{aligned}
&e^{\lambda_1 t}(\alpha\|X(t)\|_\infty + \|Y(t)\|_\infty) \\
&\le e^{\lambda_1 r_0}\Big(\alpha \sup_{-r_0\le\theta\le 0}(e^{\lambda_1(t+\theta)}|X(t+\theta)|) + \sup_{-r_0\le\theta\le 0}(e^{\lambda_1(t+\theta)}|Y(t+\theta)|)\Big) \\
&\le e^{\lambda_1 r_0}\Big(\int_0^t e^{\lambda_1 s}\{c_0 + (\alpha\delta + K_1)\|X_s\|_\infty + (\alpha\|B\| + K_2)\|Y_s\|_\infty\}ds + e^{\lambda_1 t}\|Z_t\|_\infty\Big) \\
&\le c_1 e^{\lambda_1 t}\big(1 + \|Z_t\|_\infty\big) + \lambda' e^{\lambda_1 r_0}\int_0^t e^{\lambda_1 s}(\alpha\|X_s\|_\infty + \|Y_s\|_\infty)ds
\end{aligned}$$

for some constant $c_1 > 0$. By Gronwall's inequality and $\lambda = \lambda_1 - \lambda' e^{\lambda_1 r_0} > 0$ as assumed above, this yields

$$\begin{aligned}
\alpha\|X_t\|_\infty + \|Y_t\|_\infty &\le c_1\big(1 + \|Z_t\|_\infty\big) + c_1\lambda' e^{\lambda_1 r_0}\int_0^t \big(1 + \|Z_s\|_\infty\big)e^{-\lambda''(t-s)}ds \\
&\le c_2\Big(1 + \|Z_t\|_\infty + \int_0^t \|Z_s\|_\infty e^{-\lambda''(t-s)}ds\Big)
\end{aligned}$$

for some constant $c_2 > 0$. Hence, by using Hölder's and Jensen's inequalities as in (3.10) and applying (3.5) for the present Z_t, we finish the proof. □

Finally, the following lemma ensures the existence and uniqueness of invariant probability measure and verifies condition (iii) in Theorem 1.1, so that the proof of Theorem 1.3 is finished.

Lemma 4.4. Assume **(B1)**- **(B3)** and (1.9). Then P_t has a unique invariant measure μ. Moreover, $\mu(e^{\varepsilon\|\cdot\|_\infty^2}) < \infty$ holds for some constant $\varepsilon > 0$.

Proof. Let $\mu_t^{\xi,\eta}$ be the distribution of $(X_t^{\xi,\eta}, Y_t^{\xi,\eta})$ and let

$$\rho((\xi,\eta),(\overline{\xi},\overline{\eta})) = 1 \wedge (\|\xi - \overline{\xi}\|_\infty + \|\eta - \overline{\eta}\|_\infty).$$

Making using of Lemmas 4.2 and 4.3 and carrying out an argument of Lemma 3.2, we only need to prove that $\{\mu_t^{\xi,\eta}\}_{t\ge 0}$ is W-Cauchy as $t \to \infty$.

For any $t_2 > t_1 > 0$, let $(\widetilde{X}(t), \widetilde{Y}(t))$ solve equation (1.8) for $t \in [t_2 - t_1, t_2]$ with $(\widetilde{X}_{t_2-t_1}, \widetilde{Y}_{t_2-t_1}) = (\xi, \eta)$. Then, the laws of $(\widetilde{X}_{t_2}, \widetilde{Y}_{t_2})$ is $\mu_{t_1}^{\xi,\eta}$. So,

$$W(\mu_{t_1}^{\xi,\eta}, \mu_{t_2}^{\xi,\eta}) \le \mathbb{E}\big(\|X_{t_2}^{\xi,\eta} - \widetilde{X}_{t_2}\|_\infty + \|Y_{t_2}^{\xi,\eta} - \widetilde{Y}_{t_2}\|_\infty\big). \tag{4.13}$$

Next, repeating the proof of Lemma 4.2 for $t \in [t_2 - t_1, t_2]$ and $(\widetilde{X}_t, \widetilde{Y}_t)$ in place of $(X_t^{\overline{\xi},\overline{\eta}}, Y_t^{\overline{\xi},\overline{\eta}})$, we obtain

$$\|X_{t_2}^{\xi,\eta} - \widetilde{X}_{t_2}\|_\infty + \|Y_{t_2}^{\xi,\eta} - \widetilde{Y}_{t_2}\|_\infty \le c(\|\xi - X_{t_2-t_1}^{\xi,\eta}\|_\infty + \|\eta - Y_{t_2-t_1}^{\xi,\eta}\|_\infty)e^{-\lambda t_1}$$

for some constant $c > 0$ independent of t_1 and t_2. Combining this with (4.13) and using Lemma 4.3, we prove $\lim_{t_1,t_2\to\infty} W(\mu_{t_1}^{\xi,\eta}, \mu_{t_2}^{\xi,\eta}) = 0$. □

Acknowledgments. We are indebted to the referee for his/her valuable comments which have greatly improved our earlier version.

References

[1] Arnaudon, M., Thalmaier A., Wang, F.-Y., Harnack inequality and heat kernel estimate on manifolds with curvature unbounded below, *Bull. Sci. Math.* **130** (2006), 223–233. MR-2215664

[2] Arnaudon, M., Thalmaier A., Wang, F.-Y., Equivalent log-Harnack and gradient for point-wise curvature lower bound, *Bull. Math. Sci.* **138** (2014), 643–655. MR-3224277

[3] Bao, J., Wang, F.-Y., Yuan, C., Hypercontractivity for Functional Stochastic Differential Equations, *Stochastic Process. Appl.*, **125** (2015), 3636–3656. MR-3357623

[4] Berman, S. M., An asymptotic bound for the tail of the distribution of the maximum of a Gaussian process, *Ann. Inst. Henri Poincaré*, **21** (1985), 47–57. MR-0791269

[5] Chen, M.-F., *From Markov Chains to Non-Equilibrium Particle Systems,* World Scientific, 1992, Singapore. MR-2091955

[6] Da Prato, G., Zabczyk, J., *Ergodicity for infinite-dimensional systems*, In: London Mathematical Society, Lecture Note Series, vol. 229, Cambridge University Press, Cambridge, 1996. MR-1417491

[7] Es-Sarhir, A., von Renesse, M.-K., Scheutzow, M., *Harnack inequality for functional SDEs with bounded memory, Electron. Commun. Probab.*, **14** (2009), 560–565. MR-2570679

[8] Ferniqu, X., *Régularité des trajectoires des fonctions aléatoires Gaussiennes,* École d'été. Calcul des Probab. St-Flour, 1974. MR-0413238

[9] Gross, L., *Logarithmic Sobolev inequalities and contractivity properties of semigroups,* Lecture Notes in Math. 1563, Springer-Verlag, 1993. MR-1292277

[10] Nelson, E., The free Markov field, *J. Funct. Anal.*, **12** (1973), 211–227. MR-0343816

[11] Wang, F.-Y., Logarithmic Sobolev inequalities on noncompact Riemannian manifolds, *Probab. Theory Related Fields*, **109** (1997), 417–424. MR-1481127

[12] Wang, F.-Y., Harnack inequality and applications for stochastic generalized porous media equations, *Ann. Probab.*, **35** (2007), 1333–1350. MR-2330974

[13] Wang, F.-Y., Harnack inequalities on manifolds with boundary and applications, *J. Math. Pures Appl.*, **94** (2010), 304–321. MR-2679029

[14] Wang, F.-Y., *Harnack inequalities and Applications for Stochastic Partial Differential Equations*, Springer, 2013, Berlin. MR-3099948

[15] Wang, F.-Y., Hypercontractivity for Stochastic Hamiltonian Systems, arXiv:1409.1995v2.