Hypercontractivity for Functional Stochastic Partial Differential Equations

Jianhai Bao\textsuperscript{b)}, Feng-Yu Wang\textsuperscript{a),c)}, Chenggui Yuan\textsuperscript{c)}

\textsuperscript{a)}School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China
\textsuperscript{b)}School of Mathematics and Statistics, Central South University, Changsha 410083, China
\textsuperscript{c)}Department of Mathematics, Swansea University, Singleton Park, SA2 8PP, UK

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

March 10, 2015

Abstract

Explicit 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 framework, we apply a criterion presented in the recent paper \cite{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.

AMS subject Classification: 60H15, 60J60

Keywords: Hypercontractivity, functional stochastic partial differential equation, Harnack inequality, coupling.

1 Introduction

The hypercontractivity was introduced in 1973 by Nelson \cite{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

\textsuperscript{*}Supported in part by Lab. Math. Com. Sys., the 985 project and NNSFC(11131003, 11431014, 11401592).
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 intensive study of the log-Sobolev inequality.

However, as explained in [3] that the log-Sobolev inequality does not hold for the segment solution to a stochastic delay differential equation (SDDE). As the segment solution 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 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 a process on a functional space, the equation is also called a functional stochastic differential equation (FSDE).

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

$$
(P_{t_0} f(\xi))^2 \leq (P_{t_0} f(\eta)) e^{c \rho(\xi, \eta)^2}, \quad f \in B_b(E), \quad \xi, \eta \in E;
$$

(ii) (Coupling) 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) There exists a constant $\varepsilon > 0$ such that $(\mu \times \mu)(e^{\varepsilon \rho(\cdot, \cdot)}) < \infty$.

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

$$
\mu((P_{t} f) \log P_{t} f) \leq c e^{-\alpha t} \mu(f \log f), \quad t \geq 0, f \geq 0, \mu(f) = 1;
$$

$$
\|P_t - \mu\|_2^2 := \sup_{\mu(f) \leq 1} \mu((P_t f - \mu(f))^2) \leq c e^{-\alpha t}, \quad t \geq 0
$$

(1.1)
hold for some constants \(c, \alpha > 0\).

We will apply this result 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 \(\mathcal{L}(\mathbb{H}_1, \mathbb{H}_2)\) (respectively, \(\mathcal{L}_{HS}(\mathbb{H}_1, \mathbb{H}_2)\)) be the set of all bounded (respectively, Hilbert-Schmidt) linear operators from \(\mathbb{H}_1\) to \(\mathbb{H}_2\). We will use \(|·|\) and \(⟨·, ·⟩\) to denote the norm and inner product on a Hilbert space, and let \(∥·∥\) and \(∥·∥_{HS}\) 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 \(\mathcal{C} = C([−r_0, 0]; \mathbb{H})\) be equipped with the uniform norm \(∥f∥\infty \coloneqq \sup_{−r_0≤θ≤0} |f(θ)|\). For \(t ≥ 0\) and \(h \in C([−r_0, ∞); \mathbb{H})\), let \(h_t \in \mathcal{C}\) be such that \(h_t(θ) = h(t + θ), θ ∈ [−r_0, 0]\).

Let \(W(t)\) be a cylindrical Brownian motion on \(\mathbb{H}\) under a complete filtered probability space \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t≥0}, \mathbb{P})\); that is,

\[
W(t) = \sum_{i=1}^{∞} B_i(t)e_i, \quad t ≥ 0
\]

for an orthonormal basis \(\{e_i\}_{i≥1}\) on \(\mathbb{H}\) and a sequence of independent one-dimensional Brownian motions \(\{B_i(t)\}_{i≥1}\) on \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t≥0}, \mathbb{P})\).

Consider the following FSPDE on \(\mathbb{H}\):

\[
(1.2) \quad dX(t) = \{AX(t) + b(X_t)\}dt + σdW(t), \quad t > 0, \quad X_0 = ξ ∈ \mathcal{C},
\]

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

**A1** \((-A, \mathcal{D}(A))\) is self-adjoint with discrete spectrum \(0 < λ_1 ≤ λ_2 ≤ · · ·\) counting multiplicities such that \(λ_i ↑ ∞\). Moreover, there exists a constant \(δ ∈ (0, 1)\) such that for every \(t > 0\), \(e^{−t(-A)^{1−δ}}σ\) extends to a unique Hilbert-Schmidt operator on \(\mathbb{H}\) which is denoted again by \(e^{−t(-A)^{1−δ}}σ\) and satisfies

\[
(1.3) \quad \int_0^1 ∥e^{−t(-A)^{1−δ}}σ∥^2_{HS}dt < ∞.
\]

**A2** There exists a constant \(L > 0\) such that \(∥b(ξ) − b(η)∥ ≤ L∥ξ − η∥_∞, ξ, η ∈ \mathcal{C}\).

**A3** \(σ\) is invertible, i.e. there exists \(σ^{-1} ∈ \mathcal{L}(\mathbb{H}, \mathbb{H})\) such that \(σ^{-1}\mathbb{H} ⊂ \mathcal{D}(σ)\) and \(σσ^{-1} = I\), 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 $\mathcal{F}_0$-measurable random variable $X_0$ on $\mathcal{C}$, there exists a unique continuous adapted process $X_t$ on $\mathbb{H}$ such that $\mathbb{P}$-a.s.

\begin{equation}
X(t) = e^{tA}\xi(0) + \int_0^t e^{(t-s)A}b(X_s)ds + \int_0^t e^{(t-s)A}\sigma dW(s), \quad t \geq 0.
\end{equation}

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

\begin{equation}
\int_0^1 \|e^{tA}\sigma\|_{HS}^{2(1+\epsilon)} dt < \infty
\end{equation}

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

By noting that
\[\|e^{-t(-A)^{1-\delta}}\sigma\|^2_{HS} = \sum_{i,j=1}^{\infty} \langle e^{-t(-A)^{1-\delta}}\sigma e_i, e_j \rangle^2 = \sum_{j=1}^{\infty} e^{-2\lambda_j t}\|\sigma^* e_j\|^2,\]

(1.3) is equivalent to

\begin{equation}
\sum_{j=1}^{\infty} \frac{|\sigma^* e_j|^2}{\lambda_j^{1-\delta}} < \infty.
\end{equation}

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{align*}
\int_0^1 \|e^{tA}\sigma\|_{HS}^{2(1+\epsilon)} dt &= \int_0^1 \left( \sum_{j=1}^{\infty} e^{-2\lambda_j t}|\sigma^* e_j|^2 \right)^{1+\epsilon} dt \\
&= \int_0^1 \left( \sum_{j=1}^{\infty} \mu_j e^{-2\lambda_j t}\lambda_j^{1-\delta} \right)^{1+\epsilon} dt \leq C \int_0^1 \left( \sum_{j=1}^{\infty} \mu_j \lambda_j^{(1+\epsilon)(1-\delta)} e^{-2(1+\epsilon)\lambda_j t} \right) dt \\
&\leq C \sum_{j=1}^{\infty} \frac{|\sigma^* e_j|^2}{\lambda_j^{1-\epsilon(1-\delta)}} < \infty, \quad \epsilon \leq \frac{\delta}{1-\delta},
\end{align*}

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

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

\begin{equation}
P_t f(\xi) := \mathbb{E} f(X_t^\xi), \quad f \in \mathcal{B}_b(\mathcal{C}), \quad \xi \in \mathcal{C}.
\end{equation}

We are ready to state the main result in this part.
Theorem 1.2. Let (A1)-(A3) hold. If \( \lambda := \sup_{s \in (0, \lambda_1)} (s - L e^{st_0}) > 0 \), then the following assertions hold.

1. \( P_t \) has a unique invariant probability measure \( \mu \) such that \( \mu(e^{||\cdot||_\infty}) < \infty \) 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

\[
\| \mu_t^\xi - \mu_t^\eta \|_{\text{var}} \leq c \| \xi - \eta \|_\infty e^{-\lambda t}, \quad t \geq t_0,
\]

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

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 \( \mathcal{C} = C([-r_0, 0]; \mathbb{H}) \) as in Subsection 1.1. Consider the following degenerate FSPDE on \( \mathbb{H} \):

\[
\begin{aligned}
&dX(t) = \{ A_1 X(t) + BY(t) \} dt, \\
&dY(t) = \{ A_2 Y(t) + b(X_t, Y_t) \} dt + \sigma dW(t),
\end{aligned}
\]

where \((A_i, \mathcal{D}(A_i)) \) is a densely defined closed linear operator on \( \mathbb{H}_i \) generating a \( C_0 \)-semigroup \( e^{tA_i} \) \((i = 1, 2)\), \( B \in \mathcal{L}(\mathbb{H}_2; \mathbb{H}_1) \), \( b: \mathcal{C} \to \mathbb{H}_2 \) is measurable, \((\sigma, \mathcal{D}(\sigma)) \) is a densely defined closed linear operator on \( \mathbb{H}_2 \), and \( W(t) \) is the cylindrical Brownian motion 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) \((-A_2, \mathcal{D}(A_2)) \) is self-adjoint with discrete spectrum \( 0 < \lambda_1 \leq \lambda_2 \leq \cdots \) counting multiplicities such that \( \lambda_i \uparrow \infty \), \( \sigma \) is invertible, and

\[
\int_0^1 \| e^{-t(-A_2)^{1-\delta}} \sigma \|_{HS}^2 dt < \infty
\]

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

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

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

(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 \( x \in \mathcal{D}(A_1) \).
(B4) There exists $A_0 \in \mathcal{L}(\mathbb{H}_1; \mathbb{H}_1)$ such that $Be^{tA_2} = e^{tA_0}e^{tA_1}B$ holds for $t \geq 0$, and

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

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 $\sum_{i=1}^{\infty} \frac{1}{\lambda_i} < \infty$ 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 $(\xi, \eta) \in \mathcal{C}$. Let $P_t$ be the Markov semigroup for the segment solution. We have

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

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

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

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

The remainder of this paper is organized as follows. In Section 2 we present a Fernique 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 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, known Fernique inequality for multi-dimensional Gaussian processes becomes invalid. So, we modify the inequality such that it holds also in infinite-dimensions. To this end, we first recall the inequality for one-dimensional Gaussian processes (see e.g. [11, page 49] for the multi-dimensional case).  

**Lemma 2.1** (Fernique 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

$$\phi(r) := \sup_{s,t \in [0, 1], |s-t| \leq r} (\mathbb{E}|\gamma(s) - \gamma(t)|^2)^{\frac{1}{2}}, \quad r \in [0, 1].$$
If \( \theta := \int_1^\infty \phi(e^{-s^2})ds < \infty \), then

\[
\mathbb{P}\left( \max_{t \in [0,1]} |\gamma(t)| \geq r(\Gamma + (2 + \sqrt{2})\theta) \right) \leq \frac{5e}{2} \int_r^\infty e^{-\frac{1}{2}s^2}ds, \quad r \geq \sqrt{5}.
\]

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 \( \{e_i\}_{i \geq 1} \), 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

\[
\phi_i(r) = \sup_{s,t \in [0,1], |s-t| \leq 1} (\mathbb{E} |\gamma_i(t) - \gamma_i(s)|^2)^{\frac{1}{2}}, \quad r \in [0,1],
\]

\[
\Gamma_i = \sup_{t \in [0,1]} (\mathbb{E} \gamma_i(t)^2)^{\frac{1}{2}}, \quad \delta_i = \Gamma_i + (2 + \sqrt{2}) \int_1^\infty \phi_i(e^{-s^2})ds, \quad i \geq 1.
\]

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

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

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

(2.2) \[ \mathbb{P}\left( \max_{t \in [0,1]} |\gamma(t)| \geq r \right) \leq ce^{-\lambda r^2}, \quad r \geq 0. \]

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

(2.3) \[ r^2 \geq \frac{5\theta \tilde{\lambda}}{\lambda - \tilde{\lambda}}. \]

In this case,

\[
r_i := \left( \frac{r^2 \log(e + \delta_i^{-1})}{\theta} \right)^{\frac{1}{2}} \geq \frac{r}{\sqrt{\theta}} \geq \sqrt{5},
\]

so that Lemma 2.1 implies

\[
\mathbb{P}\left( \max_{t \in [0,1]} |\gamma_i(t)| \geq r_i \delta_i \right) \leq \frac{5e}{2} \int_{r_i}^\infty e^{-\frac{1}{2}s^2}ds \leq c_1 e^{-\frac{1}{2}r_i^2}, \quad i \geq 1
\]

for some constant \( c_1 > 0 \). Then

\[
\mathbb{P}\left( \max_{t \in [0,1]} |\gamma(t)| \geq r \right) \leq \mathbb{P}\left( \sum_{i=1}^\infty \max_{t \in [0,1]} |\gamma_i(t)|^2 \geq r^2 \right)
\]

\[
\leq \sum_{i=1}^\infty \mathbb{P}\left( \max_{t \in [0,1]} |\gamma_i(t)|^2 \geq \frac{r^2 \log(e + \delta_i^{-1})}{\theta} \right) = \sum_{i=1}^\infty \mathbb{P}\left( \max_{t \in [0,1]} |\gamma_i(t)| \geq r_i \delta_i \right)
\]

\[
\leq c_1 \sum_{i=1}^\infty e^{-\frac{1}{2}r_i^2} \leq c_1 e^{-\lambda r^2} \sum_{i=1}^\infty \exp\left[ -r^2 \left( \frac{\log(e + \delta_i^{-1})}{2\theta} - \lambda \right) \right].
\]
Since, by (2.3) and the definition of \( \lambda \), we have
\[
\frac{r^2}{2} \left( \log(e + \delta_i^{-1}) - \lambda \right) \geq \frac{r^2}{2} \log(e + \delta_i^{-1}) (1 - \frac{\lambda}{\lambda}) \geq \frac{5}{2} \log(e + \delta_i^{-1}),
\]
it follows from (2.1) that
\[
\sum_{i=1}^{\infty} \exp \left[ -r^2 \left( \log(e + \delta_i^{-1}) - \lambda \right) \right] \leq \sum_{i=1}^{\infty} \delta_i^{\frac{5}{2}} < \infty.
\]
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:
\[
(3.1) \quad (P_t f(t))^2 \leq (P_t f(X(t))) e^{c_0 ||X - Y||^2}, \quad \xi, \eta \in \mathcal{C}, f \in \mathcal{B}_b(\mathcal{C}).
\]
That is, condition (i) holds for \( \rho(\xi, \eta) := ||X - Y||^2 \).

To verify (ii) and (iii), we will need the condition that \( \lambda := \sup_{s \in (0, \lambda_1]} (s - Le^{sr_0}) > 0 \). Without loss of generality, we may and do assume that the maximum is attend by \( \lambda_1 \); otherwise, in the following it suffices to replace \( \lambda_1 \) by \( \lambda_1' \in (0, \lambda_1] \) which attends the maximum.

By (A1), (A2), and (1.4), one has
\[
e^{\lambda_1 t} |X_t^\xi - X_t^\eta| \leq |\xi(0) - \eta(0)| + L \int_0^t e^{\lambda_1 s} ||X_s^\xi - X_s^\eta||_\infty ds.
\]
Then, we obtain that
\[
e^{\lambda_1 t} ||X_t^\xi - X_t^\eta||_\infty \leq e^{\lambda_1 r_0} \sup_{-r_0 \leq \theta \leq 0} \left( e^{\lambda_1 (t+\theta)} |X_t^\xi(t+\theta) - X_t^\eta(t+\theta)| \right)
\]
\[
\leq e^{\lambda_1 r_0} \left( ||\xi - \eta||_\infty + L \int_0^t e^{\lambda_1 s} ||X_s^\xi - X_s^\eta||_\infty ds \right).
\]
Thus, by Gronwall’s inequality we obtain
\[
(3.3) \quad ||X_t^\xi - X_t^\eta||_\infty \leq e^{\lambda_1 r_0} e^{-\lambda t} ||\xi - \eta||_\infty, \quad t \geq 0, \quad \xi, \eta \in \mathcal{C}.
\]
That is, condition (ii) holds.

To verify condition (iii) in Theorem 1.1 we prove the exponential integrability of the solution.

Lemma 3.1. Assume (A1) and (A2). If \( \lambda > 0 \) then there exists a constant \( r > 0 \) such that
\[
(3.4) \quad \sup_{t \geq 0} \mathbb{E} e^{r||X_t^\xi||^2_\infty} < \infty, \quad \xi \in \mathcal{C}.
\]
Proof. (a) We first use Theorem 2.2 to prove

\[(3.5) \quad \sup_{t \geq 0} \mathbb{E} \exp \left[ \varepsilon \|Z_t\|^2 \right] < \infty\]

for some \( \varepsilon > 0 \), where

\[(3.6) \quad Z_t(\theta) := \int_0^{(t+\theta)^+} e^{(t+\theta-s)A} \sigma dW(s), \quad t \geq 0, \theta \in [-r_0, 0].\]

To this end, for fixed \( t_0 > 0 \) let

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

Then (3.6) implies

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

Letting \( \{e_i\}_{i \geq 1} \) be the eigenbasis of \( A \), we have

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

Obviously,

\[(3.9) \quad \Gamma_i := \sup_{t \in [0,1]} \left( \mathbb{E} |\gamma(t)|^2 \right)^{\frac{1}{2}} \leq |\sigma^* e_i| \left( \int_0^\infty e^{-2\lambda_i s} ds \right)^{\frac{1}{2}} = \frac{|\sigma^* e_i|}{\sqrt{2\lambda_i}}, \quad i \geq 1.\]

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

\[
\begin{align*}
\mathbb{E}|\gamma_i(t) - \gamma_i(t')|^2 & = |\sigma^* e_i|^2 \left( \int_0^{(t-t_0)^+} e^{-2\lambda_i(t-t_0-s)} (1 - e^{-\lambda_i(t-t')r_0})^2 ds + \int_0^{(t-t_0)^+} e^{-2\lambda_i(t-t'-s)} ds \right) \\
& \leq |\sigma^* e_i|^2 \left( \frac{c(\frac{r}{2})^2 [r_0(t-t')]^\frac{r}{2}}{2\lambda_i^{1-\frac{r}{2}}} + \frac{c(\frac{r}{2})^2 [2r_0(t-t')]^\frac{r}{2}}{2\lambda_i^{1-\frac{r}{2}}} \right) \\
& =: \frac{c_1(t-t')^\frac{r}{2} |\sigma^* e_i|^2}{\lambda_i^{1-\frac{r}{2}}}, \quad i \geq 1,
\end{align*}
\]

where the constant \( c_1 > 0 \) is independent of \( t, t', t_0 \) and \( i \). So, by the definition of \( \phi_i \),

\[\phi_i(r) \leq \frac{c_1 r^\frac{r}{2} |\sigma^* e_i|}{\lambda_i^{1-\frac{r}{2}}}, \quad r \in [0, 1].\]
Combining this with (3.9) we obtain from the definition of $\delta_i$ that

$$\delta_i \leq \frac{c_2|\sigma^*c_i|}{\lambda_i^{1-\frac{1}{2}}}, \quad i \geq 1$$

holds for some constant $c_2 > 0$ independent of $t_0$. This and (1.6) imply (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 $\xi = 0$. We simply denote $X(t) = X^0(t)$. It follows from (A1), (A2), and (1.4) that

$$e^{\lambda t}|X(t)| \leq \int_0^t e^{\lambda s}\{c_0 + L\|X_s\|_\infty\}ds + e^{\lambda t}\left|\int_0^t e^{(t-s)\sigma}dW(s)\right|, \quad t \geq 0$$

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

$$e^{\lambda t}\|X_t\|_\infty \leq e^{\lambda t}\sup_{-r_0 \leq \theta \leq 0} (e^{\lambda (t+\theta)}|X(t+\theta)|)$$

$$\leq c_1 e^{\lambda t}(1 + \|Z_t\|_\infty) + L e^{\lambda t}\int_0^t e^{\lambda s}\|X_s\|_\infty ds$$

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

$$\|X_t\|_\infty \leq c_1(1 + \|Z_t\|_\infty) + c_1 L e^{\lambda t}\int_0^t \{e^{\lambda s} + e^{\lambda s}\|Z_s\|_\infty\}e^{L e^{\lambda r}(t-s)}ds$$

$$\leq c_2(1 + \|Z_t\|_\infty) + c_2 \int_0^t \|Z_s\|_\infty e^{-\lambda(t-s)}ds$$

holds for some constant $c_2 > 0$, where $\lambda = \lambda_1 - L e^{\lambda t_0} > 0$ as assumed above. Thus, using Hölder's inequality and applying Jensen's inequality for the probability measure $\nu(ds) := \frac{e^{-\lambda(t-s)}}{1 - e^{-\lambda t}}$ on $[0, t]$, we obtain

$$\mathbb{E}e^{r\|X_t\|_\infty^2} \leq e^{c_3\left(\mathbb{E}e^{c_3r\|Z_t\|_\infty^2}\right)^{\frac{1}{2}}\left(\mathbb{E}\exp\left[c_3r\left(\frac{(1 - e^{-\lambda t})}{\lambda}\int_0^t \|Z_s\|_\infty \nu(ds)\right)^2\right]\right)^{\frac{1}{2}}}$$

$$\leq e^{c_3\left(\mathbb{E}e^{c_3r\|Z_t\|_\infty^2}\right)^{\frac{1}{2}}\left(\int_0^t \mathbb{E}\exp\left[c_3r\|Z_s\|_\infty^2\right] \nu(ds)\right)^{\frac{1}{2}}}$$

$$\leq e^{c_3\sup_{s \geq 0} \mathbb{E}\exp\left[c_3r\left(\frac{1}{1 + \lambda_2}\|Z_s\|_\infty^2\right)\right], \quad t \geq 0, r > 0$$

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

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

10
Proof. The proof is similar to that of [3] Lemma 2.4. Let $\mu^*_t$ be the law of $X^*_t$. It is easy to see that if $\mu^*_t$ converges weakly to a probability measure $\mu^*$ as $t \to \infty$, then $\mu^*$ is an invariant probability measure of $P_t$. Let $\mathcal{P}(\mathcal{C})$ be the set of all probability measures on $\mathcal{C}$. Consider the $L^1$-Wasserstein distance induced by $\rho(\xi, \eta) := 1 \wedge \|\xi - \eta\|_{\infty}$:

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

where $\mathcal{F}(\mu_1, \mu_2)$ is the set of all couplings for $\mu_1$ and $\mu_2$. It is well known that $\mathcal{P}(\mathcal{C})$ is a complete metric space with respect to the distance $W$ (see, e.g., [3] 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 $\mu^*_t$ is a $W$-Cauchy sequence as $t \to \infty$, i.e.,

$$\lim_{t_1, t_2 \to \infty} W(\mu^*_{t_1}, \mu^*_{t_2}) = 0. \tag{3.11}$$

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

$$dX(t) = \{AX(t) + b(X_t)\}dt + \sigma dW(t), \quad t \in [0, t_2], \quad X_0 = \xi,$$

and

$$dY(t) = \{AY(t) + b(Y_t)\}dt + \sigma dW(t), \quad t \in [t_2 - t_1, t_2], \quad Y_{t_2 - t} = \xi.$$ 

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

$$e^{\lambda t}\mathbb{E}\|X_t - Y_t\|^2_{\infty} \leq c_1\mathbb{E}\|X_{t_2-t_1} - \xi\|^2_{\infty} + Lc_1e^{\lambda t_0}\int_{t_2-t_1}^t e^{\lambda s}\mathbb{E}\|X_s - Y_s\|^2_{\infty}ds, \quad t \in [t_2 - t_1, t_2]$$

for some constant $c_1 > 0$. By Gronwall’s inequality and $\lambda = \lambda_1 - Le^{\lambda_1t_0}$ as assumed above, this implies

$$\mathbb{E}\|X_t - Y_t\|^2_{\infty} \leq c_1e^{-\lambda(t-t_2+t_1)}\mathbb{E}\|X_{t_2-t_1} - \xi\|^2_{\infty}, \quad t \in [t_2 - t_1, t_2].$$

Combining with (3.4) yields

$$\mathbb{E}\|X_{t_2} - Y_{t_2}\|^2_{\infty} \leq c_2e^{-\lambda t_1},$$

so that

$$W(\mu^*_{t_1}, \mu^*_{t_2}) \leq \mathbb{E}\|X_{t_2} - Y_{t_2}\|_{\infty} \leq \sqrt{c_2}e^{-\lambda t_1/2}.$$ 

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

$$\lim_{t \to \infty} W(\mu^*_{t_1}, \mu^*_{t_2}) = 0. \tag{3.12}$$

To prove the uniqueness, it suffices to show that $\mu^*$ is independent of $\xi \in \mathcal{C}$. This follows since, by the triangle inequality, (3.3) and (4.2),

$$W(\mu^*, \mu^0) \leq \lim_{t \to \infty} \{W(\mu^*_{t_1}, \mu^*) + W(\mu^*_{t_2}, \mu^0) + W(\mu^*_{t_1}, \mu^0)\} = 0, \quad \xi, \eta \in \mathcal{C}.$$
Finally, since $\mu_t^0 \rightarrow \mu$ weakly as $t \rightarrow \infty$, by (3.4) we have

$$\mu(e^{r\|\cdot\|_2}e^{r\|\cdot\|_\infty}) = \lim_{N \rightarrow \infty} \mu(N \wedge e^{r\|X_0\|_2}) = \lim_{N \rightarrow \infty} \lim_{t \rightarrow \infty} \mathbb{E}(N \wedge e^{r\|X_0\|_\infty}) < \infty.$$ 

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

$$P_t \log f(\xi) \leq \log P_t f(\eta) + \frac{c_0}{2} \|\xi - \eta\|_\infty^2, \quad 0 < f \in B_b(\mathcal{C}), \xi, \eta \in \mathcal{C}.$$ 

By [2, Proposition 2.3], this implies

$$|P_t f(\xi) - P_t f(\eta)|^2 \leq c_0 \|\xi - \eta\|_\infty^2 \|f\|_\infty^2, \quad f \in B_b(\mathcal{C}), \xi, \eta \in \mathcal{C}.$$ 

Combining this with the Markov property we obtain

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

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 of 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

$$P_{t_0} f(\xi, \eta))^2 \leq e^{c(\|\xi - \eta\|_\infty^2 + \|\eta - \sigma\|_\infty^2)} P_{t_0} f^2(\xi, \eta), \quad (\xi, \eta), (\xi, \eta) \in \mathcal{C}, \ f \in B_b(\mathcal{C}).$$
Proof. Let \((X(t), Y(t)) = (X^{t,η}(t), Y^{t,η}(t))\) for \(t \geq 0\), and let \((\bar{X}(t), \bar{Y}(t))\) solve the following equation for \((\bar{X}_0, \bar{Y}_0) = (\xi, \eta)\):

\[
\begin{aligned}
\frac{d\bar{X}}{dt}(t) &= \{A_1\bar{X}(t) + B\bar{Y}(t)\}dt, \\
\frac{d\bar{Y}}{dt}(t) &= \{A_2\bar{Y}(t) + b(X_t, Y_t) + \frac{1_{[0,t_0-r_0]}(t)}{t_0-r_0}e^{tA_2}(\eta(0) - \eta(0)) + e^{tA_2}h(t)\}dt + \sigma dW(t),
\end{aligned}
\]

where

\[h(t) := (t_0 - r_0 - t)^+ e^{-sA_0^*}e, \quad t \in [0, t_0]\]

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

\[\frac{d\bar{Y}}{dt}(t) - Y(t) = e^{tA_2}\left\{\frac{(\bar{Y}(0) - \eta(0))(t_0-r_0-t)^+}{t_0-r_0} + h(t)\right\}, \quad t \in [0, t_0].\]

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

\[\bar{X}(t) - X(t) = e^{tA_1}(\bar{\xi}(0) - \xi(0)) + \int_0^t e^{(t-s)A_1}B(\bar{Y}(s) - Y(s))ds.\]

Substituting (4.2) into this equation we obtain

\[\bar{X}(t) - X(t) = e^{tA_1}(\bar{\xi}(0) - \xi(0)) + \int_0^t e^{(t-s)A_1}B e^{sA_2}\left\{\frac{(\bar{Y}(0) - \eta(0))(t_0-r_0-s)^+}{t_0-r_0} + h(s)\right\}ds.\]

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

\[
\begin{aligned}
\bar{X}(t) - X(t) &= e^{tA_1}(\bar{\xi}(0) - \xi(0)) + \int_0^t e^{(t-s)A_1}e^{sA_1}e^{sA_0}B\left\{\frac{(\bar{Y}(0) - \eta(0))(t_0-r_0-s)^+}{t_0-r_0} + h(s)\right\}ds \\
&= e^{tA_1}(\bar{\xi}(0) - \xi(0)) + \int_0^t e^sA_0B\left\{\frac{(\bar{Y}(0) - \eta(0))(t_0-r_0-s)^+}{t_0-r_0} + h(s)\right\}ds \\
&= e^{tA_1}(\bar{\xi}(0) - \xi(0)) + \int_0^{t_0-r_0} e^sA_0B\left\{\frac{(\bar{Y}(0) - \eta(0))(t_0-r_0-s)}{t_0-r_0} + h(s)\right\}ds
\end{aligned}
\]

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

\[\bar{Q}_{t_0-r_0} := \int_0^{t_0-r_0} s(t_0-r_0-s)e^{sA_0}BB^*e^{sA_0^*}ds\]

is invertible on \(\mathbb{H}_1\). So, we may take

\[e = -\bar{Q}_{t_0-r_0}^{-1}\left\{\bar{\xi}(0) - \xi(0) + \int_0^{t_0-r_0} \frac{t_0-r_0-s}{t_0-r_0}e^{-sA_1}B(\bar{Y}(0) - \eta(0))ds\right\}
\]

in (4.4) to derive \(\bar{X}_{t_0} = X_{t_0}\). Therefore, \((X_{t_0}, Y_{t_0}) = (\bar{X}_{t_0}, \bar{Y}_{t_0})\).
Lemma 4.2. Assume \( \lambda \in (0, \lambda_1) \) and do assume \( \lambda'' = \lambda - \lambda e^{\lambda_1 r_0} > 0 \); otherwise in the sequel it suffices to replace \( \lambda_1 \) by \( \lambda_1' \in (0, \lambda_1) \) which attends the maximum in the definition of \( \lambda'' \).

Let
\[
\tilde{W}(t) = W(t) + \int_0^t \phi(s) \, ds, \quad t \in [0, t_0],
\]
where
\[
\phi(t) := \sigma^{-1} \left( b(X_t, Y_t) - b(X_t, \bar{Y}_t) + \frac{1}{t_0 - r_0} \tilde{\lambda}(t) e^{\lambda A_2}(\eta(0) - \bar{\eta}(0)) + e^{\lambda A_2} h'(t) \right).
\]

By (4.2) and (4.4), for some constant \( C > 0 \) we have
\[
||X_t - \bar{X}_t||^2_\infty + ||Y_t - \bar{Y}_t||^2_\infty \leq C(||\xi - \bar{\xi}||^2_\infty + ||\eta - \bar{\eta}||^2_\infty), \quad t \in [0, t_0].
\]

Thus, by the Girsanov theorem (see, e.g., [3] Theorem 10.14], \( \{\tilde{W}(s)\}_{s \in [0, T]} \) is a cylindrical Wiener process under the weighted probability measure \( d\tilde{Q} := Rd\tilde{P} \) with
\[
R := \exp \left( - \int_0^{t_0} \langle \phi(s), dW(s) \rangle - \frac{1}{2} \int_0^{t_0} |\phi(s)|^2 ds \right).
\]

Now, we reformulate the equation for \( (\bar{X}(t), \bar{Y}(t)) \) as
\[
\begin{align*}
\frac{d\bar{X}(t)}{dt} &= \{A_1 \bar{X}(t) + B\bar{Y}(t)\} dt, \\
\frac{d\bar{Y}(t)}{dt} &= \{A_2 \bar{Y}(t) + b(X_t, Y_t)\} dt + \sigma d\tilde{W}(t), \quad t \in [0, t_0].
\end{align*}
\]

Then, by the weak uniqueness of the equation and using \( (\bar{X}_{t_0}, \bar{Y}_{t_0}) = (X_{t_0}, Y_{t_0}) \), we derive
\[
\begin{align*}
(P_{t_0} f(\xi, \eta))^2 &= \left\{ \mathbb{E}_{\tilde{Q}} f(X_{t_0}, Y_{t_0}) \right\}^2 = \left\{ \mathbb{E} f(X_{t_0}, Y_{t_0}) \right\}^2 \\
&\leq (\mathbb{E} R^2) \mathbb{E} f^2(X_{t_0}, Y_{t_0}) = (\mathbb{E} R^2) P_{t_0} f^2(\xi, \eta).
\end{align*}
\]

Combining this with (4.5) and the definitions of \( R \) and \( \phi \), 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 \( \lambda_1 \) by \( \lambda_1' \in (0, \lambda_1) \) which attends the maximum in the definition of \( \lambda'' \).

Lemma 4.2. Assume (B1)-(B3) and let (1.9) hold. Then there exists \( c > 0 \) such that for \( \lambda'' := \sup_{s \in (0, \lambda_1)} (s - \lambda e^{r_0}) > 0 \),
\[
\begin{align*}
&\|X_t^{\xi, \eta} - X_t^{\bar{\xi}, \bar{\eta}}\|_\infty + \|Y_t^{\xi, \eta} - Y_t^{\bar{\xi}, \bar{\eta}}\|_\infty \\
&\leq c(||\xi - \bar{\xi}||_\infty + ||\eta - \bar{\eta}||_\infty) e^{-\lambda'' t}, \quad t \geq 0, (\xi, \eta), (\bar{\xi}, \bar{\eta}) \in \mathcal{C}.
\end{align*}
\]

Proof. By (B1)-(B3), we have
\[
\begin{align*}
e^{\lambda_1 t} |X_t^{\xi, \eta}(t) - X_t^{\bar{\xi}, \bar{\eta}}(t)| - |\xi(0) - \bar{\xi}(0)| \\
&\leq \int_0^t e^{\lambda_1 s} \{ \delta |X_s^{\xi, \eta}(s) - X_s^{\bar{\xi}, \bar{\eta}}(s)| + \|B\| \cdot |Y_s^{\xi, \eta}(s) - Y_s^{\bar{\xi}, \bar{\eta}}(s)| \} \, ds, \\
e^{\lambda_1 t} |Y_t^{\xi, \eta}(t) - Y_t^{\bar{\xi}, \bar{\eta}}(t)| - |\eta(0) - \bar{\eta}(0)| \\
&\leq \int_0^t e^{\lambda_1 s} \{ K_1 ||X_s^{\xi, \eta} - X_s^{\bar{\xi}, \bar{\eta}}||_\infty + K_2 ||Y_s^{\xi, \eta} - Y_s^{\bar{\xi}, \bar{\eta}}||_\infty \} \, ds.
\end{align*}
\]
Next, let

\[(4.8)\]
\[
\alpha = \frac{\delta - K_2 + \sqrt{(K_2 - \delta)^2 + 4K_1\|B\|}}{2\|B\|}.
\]

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

\[(4.9)\]
\[
\alpha\delta + K_1 = \lambda'\alpha, \quad \alpha\|B\| + K_2 = \lambda'.
\]

Combining \((4.8)\), \((4.9)\) and \((4.7)\), we derive

\[
e^{\lambda t}(\alpha\|X_{t}^{\xi,\eta} - X_{s}^{\xi,\eta}\|_{\infty} + \|Y_{t}^{\xi,\eta} - Y_{s}^{\xi,\eta}\|_{\infty})
\]
\[
\leq e^{\lambda r_{0}}\left\{\alpha\|\xi - \xi\|_{\infty} + \|\eta - \eta\|_{\infty}
\right.
\]
\[
+ \int_{0}^{t} e^{\lambda s}(\delta\alpha + K_1)\|X_{s}^{\xi,\eta} - X_{s}^{\xi,\eta}\|_{\infty} + (\alpha\|B\| + K_2)\|Y_{s}^{\xi,\eta} - Y_{s}^{\xi,\eta}\|_{\infty})ds\}
\]
\[
\leq e^{\lambda r_{0}}\left\{\alpha\|\xi - \xi\|_{\infty} + \|\eta - \eta\|_{\infty}
\right.
\]
\[
+ \lambda' \int_{0}^{t} e^{\lambda s}(\alpha\|X_{s}^{\xi,\eta} - X_{s}^{\xi,\eta}\|_{\infty} + \|Y_{s}^{\xi,\eta} - Y_{s}^{\xi,\eta}\|_{\infty})ds\}.
\]

Therefore, we complete the proof by using Gronwall’s inequality and \(\lambda'' = \lambda_1 - \lambda' e^{\lambda r_{0}}\) as assumed above. \(\square\)

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

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

\[
\sup_{t \geq 0} E e^{\varepsilon(\|X_{t}^{\xi,\eta}\|_{\infty} + \|Y_{t}^{\xi,\eta}\|_{\infty})} < \infty, \quad (\xi, \eta) \in \mathcal{C}.
\]

**Proof.** By Lemma 4.2, it suffices to prove for \((\xi, \eta) = (0,0)\). Simply denote \((X_{t}, Y_{t}) = (X_{t}^{0,0}, Y_{t}^{0,0})\). We have

\[
X(t) = \int_{0}^{t} e^{(A_{1} - \delta)(t-s)}(\delta X(s) + BY(s))ds, \quad t \geq 0.
\]

Then \((B3)\) yields

\[(4.10)\]
\[
e^{\lambda t}|X(t)| \leq \int_{0}^{t} e^{\lambda s}\{\|B\| \cdot |Y(s)| + \delta|X(s)|\}ds.
\]

Next, according to \((B1)\) and \((B2)\),

\[(4.11)\]
\[
e^{\lambda t}|Y(t)| \leq \int_{0}^{t} e^{\lambda s}\{c_0 + K_1\|X_s\|_{\infty} + K_2\|Y_s\|_{\infty}\}ds + e^{\lambda t}\left|\int_{0}^{t} e^{A_{2}(t-s)}\sigma W(s)\right|
\]

15
holds for $c_0 := |b(0, 0)|$. Obviously, using $(\mathbb{H}_2, A_2)$ to replace $(\mathbb{H}, A)$, we see that (3.10) holds for
\[ Z_t(\theta) := \int_0^{(t+\theta)^+} e^{A_2(t-s)} \sigma W(s), \quad \theta \in [-r_0, 0]. \]
Combining (4.10), (4.11) and (4.9), for the present $Z_t$ we have
\[
eq e^{\lambda r}(\sup_{-r_0 \leq \theta \leq 0} (e^{\lambda(t+\theta)}|X(t+\theta)|) + \sup_{-r_0 \leq \theta \leq 0} (e^{\lambda(t+\theta)}|Y(t+\theta)|))
\leq e^{\lambda r}\left(\int_0^t e^{\lambda s}\{c_0 + (\alpha d + K_1)\|X_s\|_\infty + (\alpha\|B\| + K_2)\|Y_s\|_\infty\}ds + e^{\lambda t}\|Z_t\|_\infty\right)
\leq c_1 e^{\lambda t}(1 + \|Z_t\|_\infty) + e^{\lambda r}\left(\int_0^t e^{\lambda s}(\alpha\|X_s\|_\infty + \|Y_s\|_\infty)ds\right)
\]
for some constant $c_1 > 0$. By Gronwall’s inequality and $\lambda'' = \lambda_1 - \lambda e^{\lambda r}$ as assumed above, this yields
\[
\alpha\|X_t\|_\infty + \|Y_t\|_\infty \leq c_1(1 + \|Z_t\|_\infty) + c_1\lambda e^{\lambda r}\left(\int_0^t (1 + \|Z_s\|_\infty) e^{-\lambda'(t-s)}ds\right)
\leq c_2\left(1 + \|Z_t\|_\infty + \int_0^t \|Z_s\|_\infty e^{-\lambda''(t-s)}ds\right)
\]
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 (1.3). Then $P_t$ has a unique invariant measure $\mu$. Moreover, $\mu(e^{\varepsilon I\|F\|_w}) < \infty$ 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
\[
\rho((\xi, \eta), (\bar{\xi}, \bar{\eta})) = 1 \wedge (||\xi - \bar{\xi}||_\infty + ||\eta - \bar{\eta}||_\infty).
\]

Due to Lemmas 3.2, 3.3 and the argument in the proof of Lemma 3.2, we only need to prove that $\{\mu_t^{\xi, \eta}\}_{t \geq 0}$ is $W$-Cauchy as $t \to \infty$.

For any $t_2 > t_1 > 0$, let $(\tilde{X}(t), \tilde{Y}(t))$ solve equation (1.8) for $t \in [t_2 - t_1, t_2]$ with $(\tilde{X}_{t_2-t_1}, \tilde{Y}_{t_2-t_1}) = (\xi, \eta)$. Then the laws of $(\tilde{X}_{t_2}, \tilde{Y}_{t_2})$ is $\mu_{t_2}^{\xi, \eta}$. So,
\[
W(\mu_{t_1}^{\xi, \eta}, \mu_{t_2}^{\xi, \eta}) \leq \mathbb{E}(\|X_{t_2}^{\xi, \eta} - \tilde{X}_{t_2}\|_\infty + \|Y_{t_2}^{\xi, \eta} - \tilde{Y}_{t_2}\|_\infty).
\]

Next, repeating the proof of Lemma 4.2 for $t \in [t_2 - t_1, t_2]$ and $(\tilde{X}_t, \tilde{Y}_t)$ in place of $(X_t^{\xi, \eta}, Y_t^{\xi, \eta})$, we obtain
\[
\|X_{t_2}^{\xi, \eta} - \tilde{X}_{t_2}\|_\infty + \|Y_{t_2}^{\xi, \eta} - \tilde{Y}_{t_2}\|_\infty \leq c(||\xi - X_{t_2-t_1}^{\xi, \eta, m_1}\|_\infty + ||\eta - Y_{t_2-t_1}^{\xi, \eta, m_1}\|_\infty)e^{-\lambda''t_1}
\]
for some constant $c > 0$ independent of $t_1$ and $t_2$. Combining this with (4.12) and using Lemma 4.3, we prove $\lim_{t_1, t_2 \to \infty} W(\mu_{t_1}^{\xi, \eta}, \mu_{t_2}^{\xi, \eta}) = 0$. \qed

16
References

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

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

[3] Bao, J., Wang, F.-Y., Yuan, C., Hypercontractivity for Functional Stochastic Differential Equations, arXiv:1406.7580v2.

[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.

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

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

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

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

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

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

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

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

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

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

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