Paper:
Wang, F. (2017). Hypercontractivity and applications for stochastic Hamiltonian systems. *Journal of Functional Analysis*
http://dx.doi.org/10.1016/j.jfa.2017.03.015

Hypercontractivity and Applications for Stochastic Hamiltonian Systems *

Feng-Yu Wang

Center of Applied Mathematics, Tianjin University, Tianjin 300072, China

Department of Mathematics, Swansea University, Singleton Park, SA2 8PP, UK

Email: wangfy@bnu.edu.cn; F.Y.Wang@swansea.ac.uk

Abstract

The hypercontractivity is proved for the Markov semigroup associated with a class of stochastic Hamiltonian systems on Hilbert spaces. Consequently, the Markov semigroup converges exponentially to the invariant probability measure in entropy and is compact for large time. These strengthen the hypocoercivity results derived in the literature. Since the log-Sobolev inequality is invalid, we introduce a new argument to prove the hypercontractivity using coupling and dimension-free Harnack inequality. The main results are illustrated by concrete examples of the kinetic Fokker-Planck equation and highly degenerate diffusion processes.

AMS subject Classification: 65G17, 65G60.

Keywords: Hypercontractivity, stochastic Hamiltonian system, Harnack inequality, exponential convergence, compactness.

1 Introduction

To motivate the present study, we first recall the famous hypocoercivity result of C. Villani [14]. Consider the following degenerate SDE (stochastic differential equation) for (X_t, Y_t) on $\mathbb{R}^d \times \mathbb{R}^d$:

$$
\begin{cases} \mathrm{d}X_t = Y_t \,\mathrm{d}t, \\ \mathrm{d}Y_t = \{\nabla V(X_t) - Y_t\}\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t, \end{cases} \tag{1.1}
$$

*Supported in part by NNSFC(11431014) and Start-Up Fund of Tianjin University.

where $V \in C^2(\mathbb{R}^d)$ such that

$$\mu(\mathrm{d}x, \mathrm{d}y) := \mathrm{e}^{V(x) - \frac{1}{2}|y|^2} \mathrm{d}x \mathrm{d}y$$

is a probability measure on $\mathbb{R}^d \times \mathbb{R}^d$, and W_t is the d-dimensional Brownian motion. This type degenerate SDE is known as "Stochastic Hamiltonian System (Abbrev. SHS)" in probability theory (see [22]), and the distribution density of the solution solves the kinetic Fokker-Planck equation (see [14]). Let P_t be the Markov semigroup for the solution of (1.1). According to [14, Theorem 35], if there exists a constant $C > 0$ such that

$$|\nabla^2 V| \le C(1 + |\nabla V|)$$

and the following Poincaré inequality holds for $\mu_1(\mathrm{d}x) := \mu(\mathrm{d}x \times \mathbb{R}^d)$:

$$\mu_1(f^2) \le C\mu_1(|\nabla f|^2), \quad f \in C_b^1(\mathbb{R}^d), \mu_1(f) = 0,$$

then for some constants $c, \lambda > 0$ one has

$$\mu(|\nabla P_t f|^2 + (P_t f)^2) \le c\mathrm{e}^{-\lambda t} \mu(|\nabla f|^2 + f^2), \quad f \in C_b^1(\mathbb{R}^{2d}), \mu(f) = 0, t \ge 0. \tag{1.2}$$

See [6, 7, 8, 10] and references within for L^2-exponential convergence of the same type degenerate diffusion semigroups. The methodology used in these papers relies heavily on the explicit formulation of the invariant probability measure μ. In this paper, we investigate the hypercontractivity, a stronger property than the L^2-exponential convergence, for more general degenerate diffusion processes with inexplicit invariant probability measures.

The model we investigate here is the following SHS on $\mathbb{H} := \mathbb{H}_1 \times \mathbb{H}_2$, where $\mathbb{H}_1$ and $\mathbb{H}_2$ are two separable Hilbert spaces:

$$\begin{cases} \mathrm{d}X_t = (AX_t + BY_t)\,\mathrm{d}t, \\ \mathrm{d}Y_t = Z(X_t, Y_t)\mathrm{d}t + \sigma \mathrm{d}W_t, \end{cases} \tag{1.3}$$

where

- A is a densely defined (possibly unbounded) linear operator on $\mathbb{H}_1$;
- B is a bounded linear operator from $\mathbb{H}_2$ to $\mathbb{H}_1$;
- Z is a densely defined map from $\mathbb{H}$ to $\mathbb{H}_2$;
- σ is a linear operator on $\mathbb{H}_2$;
- W_t is the cylindrical Brownian motion on $\mathbb{H}_2$, i.e.

$$W_t = \sum_{i \ge 1} B_t^i e_i$$

for independent one-dimensional Brownian motions $\{B_t^i\}_{i \ge 1}$ and orthonormal basis $\{e_i\}_{i \ge 1}$ of $\mathbb{H}_2$.

See [11, 20, 21] for results on the existence and uniqueness of (mild) solutions, as well as Harnack inequality and gradient estimate of the associated Markov semigroup P_t. We intend to find out explicit conditions ensuring the existence and uniqueness of the invariant probability measure μ (whose formulation is in general unknown) and, furthermore, the hypercontractivity of P_t.

According to Nelson [12], P_t is called hypercontractive if it has an invariant probability measure μ such that

$$\|P_t\|_{L^2(\mu)\to L^4(\mu)} := \sup\{\|P_t f\|_{L^4(\mu)} : \ \mu(f^2) \le 1\} = 1 \text{ for some } t > 0.$$

By the semigroup property and the interpolation theorem, the norm $\|\cdot\|_{L^2(\mu)\to L^4(\mu)}$ can be replaced by $\|\cdot\|_{L^p(\mu)\to L^q(\mu)}$ for any $(p,q) \in (1,\infty)$ with $q > p$. As applications of the hypercontractivity, we will prove the compactness of P_t for large $t > 0$ and the exponential convergence in entropy.

Due to L. Gross (see e.g. [9]), the hypercontractivity of P_t follows from the log-Sobolev inequality

$$\mu(f^2 \log f^2) - \mu(f^2) \log \mu(f^2) \le C\mathscr{E}(f,f), \quad f \in \mathscr{D}(\mathscr{E})$$

for some constant $C > 0$, where $(\mathscr{E}, \mathscr{D}(\mathscr{E}))$ is the associated energy form. Because of this result, the log-Sobolev inequality has been intensively investigated for forty years. However, since the energy form $\mathscr{E}$ associated with (1.3) satisfies

$$\mathscr{E}(f,f) = \mu(|\sigma^* \nabla_y f|^2) = 0$$

for $f \in C_b^1(\mathbb{H})$ with $f(x,y)$ depending only on x, the log-Sobolev inequality is invalid. So, to prove the hypercontractivity we need to develop a new argument.

The remainder of the paper is organized as follows. In Section 2, we introduce a general result on the hypercontractivity using coupling and dimension-free Harnack inequality initiated from [15]. This result is then applied in Sections 3 and 4 to finite- and infinite-dimensional SHS respectively. Finally, concrete examples are presented in Section 5 to illustrate our main results.

2 Hypercontractivity using Harnack inequality

In this section, we introduce a general result on the hypercontractivity using Harnack inequality. The basic idea of the study goes back to [15] for elliptic diffusion semigroups on manifolds, see also [2] for a recent study of functional SDEs.

For a probability space $(E, \mathscr{B}, \mu)$, let P_t be a Markov semigroup on $\mathscr{B}_b(E)$ such that μ is P_t-invariant, i.e. $\mu(P_t f) = \mu(f)$ for $f \in L^1(\mu)$ and $t \ge 0$. Recall that a process (X_t, Y_t) on $E \times E$ is called a coupling of the Markov process with semigroup P_t, if

$$(P_t f)(X_0) = \mathbb{E}\big(f(X_t)|X_0\big), \quad (P_t f)(Y_0) = E\big(f(Y_t)|Y_0\big), \quad f \in \mathscr{B}_b(E), t \ge 0.$$

Theorem 2.1. *Assume that the following three conditions hold for some measurable functions $\rho : E \times E \to (0,\infty)$ and $\phi : [0,\infty) \to (0,\infty)$ with $\lim_{t\to\infty} \phi(t) = 0$:*

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

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

(*ii*) *For any* $(X_0, Y_0) \in E \times E$, *there exists a coupling* (X_t, Y_t) *associated to* P_t *such that*

$$\rho(X_t, Y_t) \le \phi(t) \rho(X_0, Y_0), \quad t \ge 0;$$

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

Then μ *is the unique invariant probability measure and* P_t *is hypercontractive. Consequently,* P_t *is compact in* $L^2(\mu)$ *for large* $t > 0$, *and there exist constants* $c, \lambda > 0$ *such that*

$$\begin{aligned} &\mu((P_t f) \log P_t f) \le c \mathrm{e}^{-\lambda t} \mu(f \log f), \quad t \ge 0, f \ge 0, \mu(f) = 1; \\ &\|P_t f - \mu(f)\|_{L^2(\mu)} \le c \mathrm{e}^{-\lambda t} \|f - \mu(f)\|_{L^2(\mu)}, \quad f \in L^2(\mu), t \ge 0. \end{aligned} \tag{2.1}$$

To prove this result, we introduce two propositions on the hypercontractivity and applications for bounded linear operators. The first is generalized from [16] where symmetric Markov operators are considered.

Proposition 2.2. *Let* P *be a bounded linear operator on* $L^2(\mu)$ *such that* $P1 = 1$ *and* μ *is* P-*invariant, i.e.* $\mu(Pf) = \mu(f)$ *for* $f \in L^2(\mu)$. *If* $\|P\|^4_{L^2(\mu) \to L^4(\mu)} < 2$, *then*

(1) $\|P - \mu\|_{L^2(\mu)} := \sup\{\|Pf - \mu(f)\|_{L^2(\mu)} : \ \mu(f^2) \le 1\} < 1$;

(2) $\|P^n\|_{L^2(\mu) \to L^4(\mu)} = 1$ *for large enough* $n \in \mathbb{N}$.

Proof. (1) Let $\delta(P) := \|P\|^4_{L^2(\mu) \to L^4(\mu)} < 2$. For any $f \in L^2(\mu)$ with $\mu(f^2) = 1$ and $\mu(f) = 0$, we intend to prove

$$\mu((Pf)^2) \le \inf_{\varepsilon \in (0,1)} \frac{\sqrt{8\varepsilon^2 + \delta(P)} - 3\varepsilon}{1 - \varepsilon}. \tag{2.2}$$

Without loss of generality, we assume $\mu((Pf)^3) \ge 0$, otherwise it suffices to replace f by $-f$. For any $\varepsilon \in (0,1)$, let $g_\varepsilon = \sqrt{\varepsilon} + \sqrt{1-\varepsilon} f$. Then $\mu(g_\varepsilon^2) = 1$. Since $P1 = 1$, $\mu(Pf) = \mu(f) = 0, \mu((Pf)^3) \ge 0, \mu(g_\varepsilon^2) = 1$ and $\mu((Pf)^4) \ge \mu((Pf)^2)^2$, we have

$$\begin{aligned} &\delta(P) \ge \mu((Pg_\varepsilon)^4) \\ &= \varepsilon^2 + (1-\varepsilon)^2 \mu((Pf)^4) + 6\varepsilon(1-\varepsilon)\mu((Pf)^2) + 4\varepsilon^{\frac{3}{2}}\sqrt{1-\varepsilon}\mu(Pf) + 4\sqrt{\varepsilon}(1-\varepsilon)^{\frac{3}{2}}\mu((Pf)^3) \\ &\ge (1-\varepsilon)^2 \mu((Pf)^2)^2 + 6\varepsilon(1-\varepsilon)\mu((Pf)^2) + \varepsilon^2. \end{aligned}$$

This implies (2.2). According to the calculations in [16, pages 2632-2633], $\delta(P) < 2$ and (2.2) imply

$$\|P - \mu\|^2_{L^2(\mu)} \le \inf_{\varepsilon \in (0,1)} \frac{\sqrt{8\varepsilon^2 + \delta(P)} - 3\varepsilon}{1 - \varepsilon} < 1.$$

(2) For $f \in L^2(\mu)$ with $\mu(f^2) = 1$, let $\hat{f} = f - \mu(f)$. We have $\mu(P^m \hat{f}) = 0, m \geq 1$. Let $\theta := \|P - \mu\|_{L^2(\mu)}$. Then

$$\mu((P^m \hat{f})^2) \leq \theta^{2m} \mu(\hat{f}^2), \quad m \geq 1,$$

so that

$$\begin{aligned}
&\mu((P^{m+1} f)^4) = \mu(f)^4 + 4\mu(f)\mu((P^{m+1}\hat{f})^3) + 6\mu(f)^2\mu((P^{m+1}\hat{f})^2) + \mu((P^{m+1}\hat{f})^4) \\
&\leq \mu(f)^4 + 4\|P\|^3_{L^2(\mu)\to L^3(\mu)} |\mu(f)| \mu((P^m \hat{f})^2)^{\frac{3}{2}} \\
&\quad + 6\mu(f)^2 \mu((P^{m+1}\hat{f})^2) + \|P\|^4_{L^2(\mu)\to L^4(\mu)} \mu((P^m \hat{f})^2)^2 \\
&\leq \mu(f)^4 + 4\|P\|^3_{L^2(\mu)\to L^3(\mu)} \theta^{3m} |\mu(f)| \mu(\hat{f}^2)^{\frac{3}{2}} \\
&\quad + 6\theta^{2(m+1)} \mu(f)^2 \mu(\hat{f}^2) + \|P\|^4_{L^2(\mu)\to L^4(\mu)} \theta^{4m} \mu(\hat{f}^2)^2.
\end{aligned}$$

Since $\theta \in (0,1)$ due to (1), $\|P\|_{L^2(\mu)\to L^3(\mu)} \leq \|P\|_{L^2(\mu)\to L^4(\mu)} < \infty$, and

$$2|\mu(f)|\mu(\hat{f}^2)^{\frac{3}{2}} \leq \mu(f)^2\mu(\hat{f}^2) + \mu(\hat{f}^2)^2,$$

this implies that for large enough $m \geq 1$,

$$\mu((P^{m+1}f)^4) \leq \mu(f)^4 + 2\mu(f)^2\mu(\hat{f}^2) + \mu(\hat{f}^2)^2 = \mu(f^2)^2 = 1.$$

Therefore, $\|P^n\|_{L^2(\mu)\to L^4(\mu)} \leq 1$ holds for large enough $n \geq 1$. $\square$

Next, we present a result on exponential convergence implied by the hypercontractivity, which is well known in the literature of symmetric Markov semigroups.

Proposition 2.3. *Let P be a posivity-preserving linear operator on $L^1(\mu)$ such that μ is P-invariant and $\|P\|_{L^p(\mu)\to L^q(\mu)} \leq 1$ holds for some constants $q > p > 1$. Then*

$$\mu((Pf)\log Pf) \leq \frac{(p-1)q}{p(q-1)} \mu(f \log f), \quad f \geq 0, \mu(f) = 1. \tag{2.3}$$

Consequently,

$$\mu\big((Pf)^2\big) \leq \frac{(p-1)q}{p(q-1)} \mu(f^2), \quad f \in L^2(\mu), \mu(f) = 0. \tag{2.4}$$

Proof. Let $f \in L^2(\mu)$ with $\mu(f) = 0$. By applying (2.3) to $f_s := \frac{1+sf}{1+s\mu(f)}$, multiplying with s^{-2} and letting $s \to 0$, we prove (2.4). So, it suffices to prove (2.3). For any $\varepsilon \in (0, p-1)$, let

$$r = \frac{p-1-\varepsilon}{(1+\varepsilon)(p-1)}, \quad \delta(\varepsilon) = \frac{p(q-1)\varepsilon}{(p-1-\varepsilon)q + \varepsilon p}.$$

Then

$$\frac{1}{1+\varepsilon} = r + \frac{1-r}{p}, \quad \frac{1}{1+\delta(\varepsilon)} = r + \frac{1-r}{q}.$$

Since $\|P\|_{L^1(\mu)} = 1$ and $\|P\|_{L^p(\mu)\to L^q(\mu)} \le 1$, Riesz-Thorin's interpolation theorem implies $\|P\|_{L^{1+\varepsilon}(\mu)\to L^{1+\delta(\varepsilon)}(\mu)} \le 1$. So, for any $f \in \mathscr{B}_b^+(E)$ with $\mu(f) = 1$,

$$\int_E (Pf^{\frac{1}{1+\varepsilon}})^{1+\delta(\varepsilon)}\mathrm{d}\mu \le 1, \quad \varepsilon \in (0, p-1).$$

Since the equality holds for $\varepsilon = 0$, this implies

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0} \int_E (Pf^{\frac{1}{1+\varepsilon}})^{1+\delta(\varepsilon)}\mathrm{d}\mu \le 0,$$

which is equivalent to (2.3). □

Proof of Theorem 2.1. (a) According to [19, Proposition 3.1], (i) implies that μ is the unique invariant probability measure of P_{t_0}, and P_{t_0} has a density with respect to μ. So, by [22, Theorem 2.3], if $\|P_t\|_{L^2(\mu)\to L^4(\mu)} < \infty$ then P_{t_0+t} is compact in $L^2(\mu)$. Therefore, according to Propositions 2.2 and 2.3, it remains to prove $\|P_t\|^4_{L^2(\mu)\to L^4(\mu)} < 2$ for large enough $t > 0$.

(b) Let $f \in \mathscr{B}_b(E)$ with $\mu(f^2) \le 1$. By (i) and (ii) we have

$$\begin{aligned}(P_{t+t_0}f(\xi))^2 &\le \mathbb{E}(P_{t_0}f(X_t))^2 \le \mathbb{E}\Big[(P_{t_0}f^2(Y_t))\mathrm{e}^{c_0\rho(X_t,Y_t)^2}\Big]\\ &\le (P_{t_0+t}f^2(\eta))\mathrm{e}^{c_0\phi(t)^2\rho(\xi,\eta)^2}, \quad t \ge 0, (\xi,\eta) \in E\times E.\end{aligned}$$

Equivalently,

$$(P_{t_0+t}f(\xi))^2\mathrm{e}^{-c_0\phi(t)^2\rho(\xi,\eta)^2} \le P_{t_0+t}f^2(\eta), \quad t \ge 0, (\xi,\eta) \in E\times E.$$

Integrating with respect to $\mu(\mathrm{d}\eta)$ gives

$$(P_{t_0+t}f(\xi))^2 \int_E \mathrm{e}^{-c_0\phi(t)^2\rho(\xi,\eta)^2}\mu(\mathrm{d}\eta) \le \int_E P_{t_0+t}f^2(\eta)\mu(\mathrm{d}\eta) = \mu(f^2) \le 1, \quad t \ge 0, \xi \in E.$$

Thus,

$$(P_{t_0+t}f(\xi))^4 \le \frac{1}{\big(\int_E \exp[-c_0\phi(t)^2\rho(\xi,\eta)^2]\mu(\mathrm{d}\eta)\big)^2}, \quad \mu(f^2) \le 1, t \ge 0, \xi \in E.$$

Then by Jensen's inequality, for $t \ge 0$

$$\begin{aligned}\sup_{\mu(f^2)\le 1}\int_E (P_{t+t_0}f(\xi))^4\mu(\mathrm{d}\xi) &\le \int_E \frac{\mu(\mathrm{d}\xi)}{(\int_E \exp[-c_0\phi(t)^2\rho(\xi,\eta)^2]\mu(\mathrm{d}\eta))^2}\\ &\le \int_E\bigg(\int_E \mathrm{e}^{c_0\phi(t)^2\rho(\xi,\eta)^2}\mu(\mathrm{d}\eta)\bigg)^2\mu(\mathrm{d}\xi) \le \int_{E\times E}\mathrm{e}^{2c_0\phi(t)^2\rho(\xi,\eta)^2}\mu(\mathrm{d}\xi)\mu(\mathrm{d}\eta).\end{aligned} \tag{2.5}$$

Since $\lim_{t\to\infty}\phi(t) = 0$, it follows from (iii) that

$$\lim_{t\to\infty}\int_{E\times E}\mathrm{e}^{2c_0\phi(t)^2\rho(\xi,\eta)^2}\mu(\mathrm{d}\xi)\mu(\mathrm{d}\eta) = 1.$$

Combining this with (2.5) we prove $\|P_t\|^4_{2\to4} < 2$ for large enough $t > 0$. □

3 Hypercontractivity for finite-dimensional SHS

In this section, we consider the equation (1.3) with $\mathbb{H} = \mathbb{R}^{m+d}$ for some $m, d \geq 1$. Let $\|\cdot\|$ denote the operator norm. To verify conditions (i)-(iii) in Theorem 2.1, we make the following assumptions.

$(A1)$ σ is invertible and $\mathrm{Rank}[B, AB, \cdots, A^{m-1}B] = m$.

$(A2)$ $Z : \mathbb{R}^{m+d} \to \mathbb{R}^d$ is Lipschitz continuous.

$(A3)$ There exist constants $r, \theta > 0$ and $r_0 \in (-\|B\|^{-1}, \|B\|^{-1})$ such that

$$\begin{aligned}&\big\langle r^2(x-\bar{x}) + rr_0 B(y-\bar{y}), A(x-\bar{x}) + B(y-\bar{y})\big\rangle\\&+\big\langle Z(x,y) - Z(\bar{x},\bar{y}), y-\bar{y} + rr_0 B^*(x-\bar{x})\big\rangle\\&\leq -\theta(|x-\bar{x}|^2 + |y-\bar{y}|^2), \quad (x,y),(\bar{x},\bar{y}) \in \mathbb{R}^{m+d}.\end{aligned}$$

The rank condition in $(A1)$ is known as Kalman's condition, when σ is invertible it is equivalent to the Hörmander condition. We will prove the Harnack inequality in condition (i) using $(A1)$ and $(A2)$, and verify conditions (ii) and (iii) by Assumption $(A3)$.

Theorem 3.1. *Assume $(A1)$, $(A2)$ and $(A3)$. Let P_t be the Markov semigroup associated with (1.3). Then*

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

(2) *P_t is hypercontractive, i.e. $\|P_t\|_{2\to 4} = 1$ for large $t > 0$;*

(3) *P_t is compact in $L^2(\mu)$ for large $t > 0$, and there exist constants $c, \lambda > 0$ such that (2.1) holds.*

In a similar spirit of (1.2), under a generalized curvature condition [3] proved the following entropy-information inequality for some constants $c, \lambda > 0$:

$$\mu\big((P_t f)\log P_t f + (P_t f)|\nabla \log P_t f|^2\big) \leq c\mathrm{e}^{-\lambda t}\mu\big(f\log f + f|\nabla \log f|^2\big), \quad f \geq 0, \mu(f) = 1, t \geq 0.$$

This does not imply the entropy inequality in (2.1).

According to Theorem 2.1 and Proposition 2.3, Theorem 3.1 follows from the following three lemmas which correspond to conditions (i)-(iii) respectively. The first lemma provides the desired Harnack inequality. Although the Harnack inequality has been investigated in [11, 20] for SHS, the resulting results are not enough for our purpose: the inequality established in [11] (see Corollary 4.2 therein) contains a worse exponential term, while the assumption (H) in [20] does not hold if Z is not second order differentiable. So, we present below a new version of Harnack inequality for SHS using coupling by change of measures. See [18, Chapter 1] for more results on the coupling by change measures and applications.

Lemma 3.2. *Assume (A1) and (A2). For any $t_0 > 0$, there exists a constant $c_0 > 0$ such that*

$$(P_{t_0} f)^2(\xi) \le (P_{t_0} f^2(\eta)) \mathrm{e}^{c_0|\xi-\eta|^2}, \quad f \in \mathscr{B}_b(\mathbb{R}^{m+d}), \xi, \eta \in \mathbb{R}^{m+d}.$$

Proof. Let (X_t, Y_t) solve the equation (1.3) with $(X_0, Y_0) = \eta \in \mathbb{R}^{m+d}$, and let $(\bar{X}_t, \bar{Y}_t)$ solve the following equation with $(\bar{X}_0, \bar{Y}_0) = \xi \in \mathbb{R}^{m+d}$:

$$\begin{cases} \mathrm{d}\bar{X}_t = (A\bar{X}_t + B\bar{Y}_t)\,\mathrm{d}t, \\ \mathrm{d}\bar{Y}_t = \Big\{ Z(X_t, Y_t) + \dfrac{Y_0 - \bar{Y}_0}{t_0} + \dfrac{\mathrm{d}}{\mathrm{d}t}\big(t(t_0 - t)B^* \mathrm{e}^{(t_0-t)A^*} b\big) \Big\} \mathrm{d}t + \sigma \mathrm{d}W_t, \end{cases} \tag{3.1}$$

where $b \in \mathbb{R}^m$ is to be determined such that $(X_{t_0}, Y_{t_0}) = (\bar{X}_{t_0}, \bar{Y}_{t_0})$. It is easy to see that

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}(X_t - \bar{X}_t) = A(X_t - \bar{X}_t) + B(Y_t - \bar{Y}_t), \\ \frac{\mathrm{d}}{\mathrm{d}t}(Y_t - \bar{Y}_t) = \frac{1}{t_0}(\bar{Y}_0 - Y_0) - \frac{\mathrm{d}}{\mathrm{d}t}\big\{t(t_0 - t)B^* \mathrm{e}^{(t_0-t)A^*} b\big\}. \end{cases}$$

Then

$$Y_t - \bar{Y}_t = \frac{t_0 - t}{t_0}(Y_0 - \bar{Y}_0) - t(t_0 - t)B^* \mathrm{e}^{(t_0-t)A^*} b, \tag{3.2}$$

and

$$\begin{aligned} X_t - \bar{X}_t &= \mathrm{e}^{At}(X_0 - \bar{X}_0) + \int_0^t \mathrm{e}^{A(t-s)} B(Y_s - \bar{Y}_s)\mathrm{d}s \\ &= \mathrm{e}^{At}(X_0 - \bar{X}_0) + \bigg(\int_0^t \mathrm{e}^{A(t-s)} \frac{t_0 - s}{t_0} \mathrm{d}s\bigg) B(Y_0 - \bar{Y}_0) \\ &\quad - \bigg(\int_0^t s(t_0 - s)\mathrm{e}^{A(t-s)} BB^* \mathrm{e}^{(t_0-s)A^*} \mathrm{d}s\bigg) b. \end{aligned} \tag{3.3}$$

We now take

$$b = Q_{t_0}^{-1} \bigg\{ \mathrm{e}^{t_0 A}(X_0 - \bar{X}_0) + \bigg(\int_0^{t_0} \frac{t_0 - s}{t_0} \mathrm{e}^{A(t_0-s)} \mathrm{d}s\bigg) B(Y_0 - \bar{Y}_0) \bigg\}, \tag{3.4}$$

where, according to [13, §3], the rank condition in $(A1)$ ensures the invertibility of the $m \times m$-matrix

$$Q_{t_0} := \int_0^{t_0} s(t_0 - s)\mathrm{e}^{A(t_0-s)} BB^* \mathrm{e}^{(t_0-s)A^*} \mathrm{d}s,$$

see (1) in the proof of [20, Theorem 4.2] for details. Then (3.2)-(3.4) imply $(X_{t_0}, Y_{t_0}) = (\bar{X}_{t_0}, \bar{Y}_{t_0})$.

In order to establish the Harnack inequality using Girsanov's theorem, let

$$\psi_t = Z(X_t, Y_t) - Z(\bar{X}_t, \bar{Y}_t) + \frac{1}{t_0}(Y_0 - \bar{Y}_0) + \frac{\mathrm{d}}{\mathrm{d}t}\big\{t(t_0 - t)B^* \mathrm{e}^{(t_0-t)A^*} b\big\}, \quad t \in [0, t_0].$$

Since Z is Lipschitz continuous, (3.2), (3.3) and (3.4) imply

$$|\psi_t|^2 \le c_1(|X_0 - \bar{X}_0|^2 + |Y_0 - \bar{Y}_0|^2) = c_1|\xi - \eta|^2, \quad t \in [0, t_0] \tag{3.5}$$

for some constant $c_1 > 0$. Moreover, according to the definition of ψ, (3.1) can be reformulated as

$$\begin{cases} \mathrm{d}\bar{X}_t = (A\bar{X}_t + B\bar{Y}_t)\,\mathrm{d}t, \\ \mathrm{d}\bar{Y}_t = Z(\bar{X}_t, \bar{Y}_t)\mathrm{d}t + \sigma \mathrm{d}\bar{W}_t, \end{cases}$$

where

$$\bar{W}_t := W_t + \sigma^{-1}\int_0^t \psi_s \mathrm{d}s, \quad t \in [0, t_0].$$

Let

$$R := \exp\bigg[-\int_0^{t_0} \langle \sigma^{-1}\psi_t, \mathrm{d}W_t\rangle - \frac{1}{2}\int_0^{t_0} |\sigma^{-1}\psi_t|^2 \mathrm{d}t\bigg]. \tag{3.6}$$

By (3.5) and Girsanov's theorem, $\tilde{W}_t$ is a d-dimensional Brownian motion under the probability measure $\mathrm{d}\mathbb{Q} := R\mathrm{d}\mathbb{P}$. Therefore, by the weak uniqueness of the equation (1.3) and using $(X_{t_0}, Y_{t_0}) = (\bar{X}_{t_0}, \bar{Y}_{t_0})$, we obtain

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

Noting that (3.5) and (3.6) imply $\mathbb{E}R^2 \le \mathrm{e}^{c_0|\xi - \eta|^2}$ for some constant $c_0 > 0$, we finish the proof. □

Lemma 3.3. *If ($A3$) holds, then there exist two constants $c, \lambda > 0$ such that for any two solutions (X_t, Y_t) and $(\tilde{X}_t, \tilde{Y}_t)$ of (1.3),*

$$|X_t - \tilde{X}_t|^2 + |Y_t - \tilde{Y}_t|^2 \le c\mathrm{e}^{-\lambda t}(|X_0 - \tilde{X}_0|^2 + |Y_0 - \tilde{Y}_0|^2), \quad t \ge 0.$$

Proof. Obviously, $X_t - \tilde{X}_t$ solves the ODE

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}(X_t - \tilde{X}_t) = A(X_t - \tilde{X}_t) + B(Y_t - \tilde{Y}_t), \\ \frac{\mathrm{d}}{\mathrm{d}t}(Y_t - \tilde{Y}_t) = \big(Z(X_t, Y_t) - Z(\tilde{X}_t, \tilde{Y}_t)\big)\mathrm{d}t. \end{cases} \tag{3.7}$$

Since $r_0 \in (-\|B\|^{-1}, \|B\|^{-1})$, for any $r > 0$ there exists a constant $C > 1$ such that

$$\begin{aligned} &\frac{1}{C}(|X_t - \tilde{X}_t|^2 + |Y_t - \tilde{Y}_t|^2) \\ &\le \Phi_t := \frac{r^2}{2}|X_t - \tilde{X}_t|^2 + \frac{1}{2}|Y_t - \tilde{Y}_t|^2 + rr_0\langle X_t - \tilde{X}_t, B(Y_t - \tilde{Y}_t)\rangle \\ &\le C(|X_t - \tilde{X}_t|^2 + |Y_t - \tilde{Y}_t|^2), \quad t \ge 0. \end{aligned} \tag{3.8}$$

Combining this with (3.7) and ($A3$), we obtain

$$\mathrm{d}\Phi_t \le -\theta(|X_t - \tilde{X}_t|^2 + |Y_t - \tilde{Y}_t|^2) \le -\frac{\theta}{C}\Phi_t \mathrm{d}t.$$

Therefore, $\Phi_t \le \Phi_0 \mathrm{e}^{-\theta t/C}$. This together with (3.8) implies the desired estimate. □

Lemma 3.4. *If $(A3)$ holds, then P_t has an invariant probability measure μ such that $\mu(\mathrm{e}^{\varepsilon|\cdot|^2}) < \infty$ for some constant $\varepsilon > 0$.*

Proof. Let (X_t, Y_t) solve (1.3) with $(X_0, Y_0) = 0 \in \mathbb{R}^{m+d}$. By a standard tightness argument, it suffices to prove

$$\sup_{t\geq 0} \mathbb{E}\mathrm{e}^{\varepsilon(|X_t|^2+|Y_t|^2)} < \infty \tag{3.9}$$

for some constant $\varepsilon > 0$. Since $r_0 \in (-\|B\|^{-1}, \|B\|^{-1})$, for any $r > 0$ there exists a constant $C > 1$ such that

$$\begin{aligned}\frac{1}{C}(|X_t|^2 + |Y_t|^2) \leq \Psi_t &:= \frac{r^2}{2}|X_t|^2 + \frac{1}{2}|Y_t|^2 + rr_0\langle X_t, BY_t\rangle \\ &\leq C(|X_t|^2 + |Y_t|^2), \quad t \geq 0.\end{aligned} \tag{3.10}$$

Moreover, $(A3)$ with $(\bar{x}, \bar{y}) = 0$ implies

$$\langle r^2x + rr_0By, Ax + By\rangle + \langle Z(x, y) - Z(0, 0), y + rr_0B^*x\rangle \leq -\theta(|x|^2 + |y|^2), \quad (x, y) \in \mathbb{R}^{m+d}.$$

Then there exist constants $c_1, c_2 > 0$ such that

$$\begin{aligned}&\langle r^2x + rr_0By, Ax + By\rangle + \langle Z(x, y), y + rr_0B^*x\rangle \\ &\leq |Z(0, 0)| \cdot |y + rB^*x| - \theta(|x|^2 + |y|^2) \leq c_1 - c_2(|x|^2 + |y|^2), \quad (x, y) \in \mathbb{R}^{m+d}.\end{aligned}$$

Thus, by (1.3), Itô's formula and (3.10), we may find out two constants $c_3, c_4 > 0$ such that

$$\begin{aligned}\mathrm{d}\Psi_t &\leq \big(c_3 - c_2(|X_t|^2 + |Y_t|^2)\big)\mathrm{d}t + \langle Y_t + rB^*X_t, \sigma\mathrm{d}W_t\rangle \\ &\leq (c_3 - c_4\Psi_t)\mathrm{d}t + \langle Y_t + rB^*X_t, \sigma\mathrm{d}W_t\rangle.\end{aligned}$$

By Itô's formula, for any $\varepsilon > 0$ there exists a local martingale M_t such that

$$\mathrm{d}\mathrm{e}^{\varepsilon\Psi_t} \leq \varepsilon\mathrm{e}^{\varepsilon\Psi_t}\Big(c_3 - c_4\Psi_t + \frac{\varepsilon^2}{2}|\sigma^*(Y_t + rB^*X_t)|^2\Big)\mathrm{d}t + \mathrm{d}M_t.$$

Noting that (3.10) implies $|\sigma^*(Y_t + rB^*X_t)|^2 \leq c_5\Psi_t$ for some constant $c_5 > 0$, by taking $\varepsilon = \frac{c_4}{c_5}$ we obtain

$$\mathrm{d}\mathrm{e}^{\varepsilon\Psi_t} \leq \varepsilon\mathrm{e}^{\varepsilon\Psi_t}\Big(c_3 - \frac{1}{2}c_4\Psi_t\Big)\mathrm{d}t + \mathrm{d}M_t \leq (c_6 - \mathrm{e}^{\varepsilon\Psi_t})\mathrm{d}t + \mathrm{d}M_t$$

for some constant $c_6 \geq 1$. Since $\mathrm{e}^{\varepsilon\Psi_0} = 1$, it follows that

$$\mathbb{E}\mathrm{e}^{\varepsilon\Psi_t} \leq c_6, \quad t \geq 0.$$

Because of (3.10), this implies (3.9) for small $\varepsilon > 0$. $\square$

4 Hypercontractivity for infinite-dimensional SHS

When $\mathbb{H}_2$ is infinite-dimensional and σ is not Hilbert-Schmidt, σW_t is ill defined on $\mathbb{H}_2$, so that the usual strong solution of (1.3) does not make sense. Alternatively, we consider the mild solution. To this end, we reformulate (1.3) on $\mathbb{H} := \mathbb{H}_1 \times \mathbb{H}_2$ as follows:

$$\begin{cases} \mathrm{d}X_t = (AX_t + BY_t - L_1 X_t)\,\mathrm{d}t, \\ \mathrm{d}Y_t = \{Z(X_t, Y_t) - L_2 Y_t\}\mathrm{d}t + \sigma \mathrm{d}W_t, \end{cases} \tag{4.1}$$

where $A : \mathbb{H}_1 \to \mathbb{H}_1, B : \mathbb{H}_2 \to \mathbb{H}_1$ and $\sigma : \mathbb{H}_2 \to \mathbb{H}_2$ are bounded linear operators; $(L_i, \mathscr{D}(L_i))$ is a positive definite self-adjoint operator on $\mathbb{H}_i, i = 1, 2$; and $Z : \mathbb{H} \to \mathbb{H}_2$ is measurable. This equation reduces to (1.3) if we regard $A - L_1$ as one operator and combine $Z(x, y)$ with $-L_2 y$. The unbounded operator L_2 plays a crucial role in the study of mild solutions (see [5]), while L_1 is the counterpart of L_2 for the first component process X_t, and the bounded operator A stands for a perturbation of L_1, see $(B3)$ below.

Let $\langle \cdot, \cdot \rangle, |\cdot|$ and $\|\cdot\|$ denote, respectively, the inner product, the norm and the operator norm on a Hilbert space. Moreover, for a linear operator $(L, \mathscr{D}(L))$ on a Hilbert space, and for $\lambda \in \mathbb{R}$, we write $L \geq \lambda$ if $\langle f, Lf \rangle \geq \lambda |f|^2$ holds for all $f \in \mathscr{D}(L)$.

To prove the hypercontractivity using Theorem 2.1, we will need the following assumptions.

$(B1)$ σ is invertible, L_2 has discrete spectrum with eigenbasis $\{e_i\}_{i\geq 1}$ and corresponding eigenvalues $0 < \lambda_1 \leq \lambda_2 \leq \cdots$ including multiplicities satisfy $\sum_{i=1}^{\infty} \frac{1}{\lambda_i} < \infty$.

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

$$|Z(x, y) - Z(\bar{x}, \bar{y})| \leq K_1 |x - \bar{x}| + K_2 |y - \bar{y}|, \quad (x, y), (\bar{x}, \bar{y}) \in \mathbb{H}.$$

$(B3)$ $L_1 - A \geq \lambda_1 - \delta$ for some constant $\delta \geq 0$, $BL_2 = L_1 B, AL_1 = L_1 A$, and for any $t > 0$

$$Q_t := \int_0^t \mathrm{e}^{sA} BB^* \mathrm{e}^{sA^*} \mathrm{d}s$$

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

It is well known that $(B1)$ and $(B2)$ imply the existence and uniqueness of mild solutions for (4.1), see [5]. Let P_t be the associated Markov semigroup.

Theorem 4.1. *Assume* $(B1)$, $(B2)$ *and* $(B3)$. *If*

$$\lambda_1 > \lambda' := \frac{1}{2}\Big(\delta + K_2 + \sqrt{(K_2 - \delta)^2 + 4K_1 \|B\|}\Big), \tag{4.2}$$

then all assertions in Theorem 3.1 *hold.*

As shown in the proof of Theorem 3.1, we need to verify conditions (i)-(iii) in Theorem 2.1. Let (X_t, Y_t) be a mild solution to (4.1). We have

$$\begin{cases} X_t = \mathrm{e}^{-(L_1 - A + \delta)t} X_0 + \int_0^t \mathrm{e}^{-(L_1 - A + \delta)(t-s)} (\delta X_s + B Y_s) \mathrm{d}s, \\ Y_t = \mathrm{e}^{-L_2 t} Y_0 + \int_0^t \mathrm{e}^{-L_2(t-s)} Z(X_s, Y_s) \mathrm{d}s + \xi_t, \end{cases} \tag{4.3}$$

where

$$\xi_t := \int_0^t \mathrm{e}^{-L_2(t-s)} \sigma \mathrm{d}W_s, \quad t \geq 0.$$

Due to $(B1)$, for any $T > 0$, the process

$$M_t^T := \int_0^t \mathrm{e}^{-L_2(T-s)} \sigma \mathrm{d}W_s, \quad t \in [0, T]$$

is a square integrable martingale on $\mathbb{H}$ with quadratic variation process

$$\langle M^T \rangle_t = \int_0^t \|\mathrm{e}^{-L_2(T-s)} \sigma\|_{HS}^2 \mathrm{d}s \leq \|\sigma\|^2 \sum_{i=1}^\infty \frac{1}{2\lambda_i} =: \alpha_0 < \infty, \quad t \in [0, T],$$

where $\| \cdot \|_{HS}$ is the Hilbert-Schmidt norm. This implies

$$\mathbb{E} \exp\Big[\frac{|M_t^T|^2}{2 + \alpha_0}\Big] \leq C, \quad T > 0, t \in [0, T] \tag{4.4}$$

for some constant $C > 0$. Indeed, since

$$\mathrm{d}|M_t^T|^2 = 2\langle M_t^T, \mathrm{d}M_t^T \rangle + \mathrm{d}\langle M^T \rangle_t, \quad t \in [0, T],$$

by Itô's formula, for any $r > 0$ we have

$$\begin{aligned} \mathrm{d}\Big\{ \exp\Big[\frac{r|M_t^T|^2 + 1}{\langle M^T \rangle_t + 1}\Big] \Big\} &= \exp\Big[\frac{r|M_t^T|^2 + 1}{\langle M^T \rangle_t + 1}\Big] \frac{2r}{\langle M^T \rangle_t + 1} \langle M_t^T, \mathrm{d}M_t^T \rangle \\ &- \exp\Big[\frac{r|M_t^T|^2 + 1}{\langle M^T \rangle_t + 1}\Big] \Big\{ \frac{r|M_t^T|^2 + 1 - r\langle M^T \rangle_t - r - 2r^2 |M_t^T|^2}{(\langle M^T \rangle_t + 1)^2} \Big\} \mathrm{d}\langle M^T \rangle_t, \quad t \in [0, T]. \end{aligned}$$

Since $\langle M^T \rangle_t \leq \alpha_0$, when $r \in (0, \frac{1}{2+\alpha_0}]$ the process $\exp\big[\frac{r|M_t^T|^2+1}{\langle M^T \rangle_t + 1}\big]$ for $t \in [0, T]$ is a supmartingale. In particular, by taking $r = \frac{1}{2+\alpha_0}$ we prove (4.4).

Since $\xi_T = M_T^T$ for any $T > 0$, (4.4) implies

$$\sup_{t \geq 0} \mathbb{E} \exp\Big[\frac{|\xi_t|^2}{2 + \alpha_0}\Big] \leq C. \tag{4.5}$$

We are now ready to prove the following four lemmas which imply Theorem 4.1 according to Theorem 2.1.

Lemma 4.2. *Assume $(B1)$, $(B2)$ and $(B3)$. For any $t_0>0$, there exists a constant $c_0>0$ such that*

$$(P_{t_0}f)^2(\xi)\le (P_{t_0}f^2(\eta))\mathrm{e}^{c_0|\xi-\eta|^2},\quad f\in\mathscr{B}_b(\mathbb{H}),\xi,\eta\in\mathbb{H}:=\mathbb{H}_1\times\mathbb{H}_2.$$

Proof. Let (X_t,Y_t) solve (4.1) with $(X_0,Y_0)=\eta$, and let $(\bar X_t,\bar Y_t)$ solve the following equation for $(\bar X_0,\bar Y_0)=\xi$:

$$\begin{cases}\mathrm{d}\bar X_t=(A\bar X_t+B\bar Y_t-L_1\bar X_t)\mathrm{d}t,\\ \mathrm{d}\bar Y_t=\Big\{Z(X_t,Y_t)-L_2\bar Y_t+\frac{1}{t_0}\mathrm{e}^{-L_2t}(Y_0-\bar Y_0)+\mathrm{e}^{-L_2t}\frac{\mathrm{d}}{\mathrm{d}t}\big(t(t_0-t)B^*\mathrm{e}^{(t_0-t)A^*}b\big)\Big\}\mathrm{d}t+\sigma\mathrm{d}W_t,\end{cases}$$

where $b\in\mathbb{H}_1$ will be determined latter such that $(X_{t_0},Y_{t_0})=(\bar X_{t_0},\bar Y_{t_0})$. We have

$$\begin{cases}\mathrm{d}(X_t-\bar X_t)=\big\{A(X_t-\bar X_t)+B(Y_t-\bar Y_t)-L_1(X_t-\bar X_t)\big\}\mathrm{d}t,\\ \mathrm{d}(Y_t-\bar Y_t)=-\Big\{L_2(Y_t-\bar Y_t)+\frac{1}{t_0}\mathrm{e}^{-L_2t}(Y_0-\bar Y_0)+\mathrm{e}^{-L_2t}\frac{\mathrm{d}}{\mathrm{d}t}\big(t(t_0-t)B^*\mathrm{e}^{(t_0-t)A^*}b\big)\Big\}\mathrm{d}t.\end{cases}$$

Then

$$Y_t-\bar Y_t=\frac{t_0-t}{t_0}\mathrm{e}^{-L_2t}(Y_0-\bar Y_0)-t(t_0-t)\mathrm{e}^{-L_2t}B^*\mathrm{e}^{(t_0-t)A^*}b,\quad t\in[0,t_0],\tag{4.6}$$

and, since $BL_2=L_1B, AL_1=L_1A$,

$$\begin{aligned}X_t-\bar X_t=&\mathrm{e}^{(A-L_1)t}(X_0-\bar X_0)+\int_0^t\frac{t_0-s}{t_0}\mathrm{e}^{(A-L_1)(t-s)}B\mathrm{e}^{-L_2s}(Y_0-\bar Y_0)\mathrm{d}s\\ &-\int_0^t s(t_0-s)\mathrm{e}^{(A-L_1)(t-s)}B\mathrm{e}^{-L_2s}B^*\mathrm{e}^{A^*(t_0-s)}b\,\mathrm{d}s\\ =&\mathrm{e}^{-tL_1}\Big\{\mathrm{e}^{At}(X_0-\bar X_0)+\int_0^t\frac{t_0-s}{t_0}\mathrm{e}^{A(t-s)}B(Y_0-\bar Y_0)\mathrm{d}s\\ &-\int_0^t s(t_0-s)\mathrm{e}^{A(t-s)}BB^*\mathrm{e}^{A^*(t_0-s)}b\,\mathrm{d}s\Big\}.\end{aligned}\tag{4.7}$$

According to $(B3)$, the operator

$$\tilde Q_{t_0}:=\int_0^{t_0}s(t_0-s)\mathrm{e}^{A(t_0-s)}BB^*\mathrm{e}^{A^*(t_0-s)}\mathrm{d}s$$

is invertible on $\mathbb{H}_1$. So, letting

$$b=\tilde Q_{t_0}^{-1}\Big\{\mathrm{e}^{At_0}(X_0-\bar X_0)+\int_0^{t_0}\frac{t_0-s}{t_0}\mathrm{e}^{A(t_0-s)}B(Y_0-\bar Y_0)\mathrm{d}s\Big\},$$

we conclude from (4.6) and (4.7) that $(X_{t_0},Y_{t_0})=(\bar X_{t_0},\bar Y_{t_0})$. Moreover, there exists a constant $C_1>0$ such that

$$|X_t-\bar X_t|+|Y_t-\bar Y_t|\le C_1(|X_0-\bar X_0|+|Y_0-\bar Y_0|),\quad t\in[0,t_0].\tag{4.8}$$

Since A, B are bounded, σ is reversible, and Z is Lipschitz continuous, this implies that the process

$$\psi_t := \sigma^{-1}\Big\{Z(X_t, Y_t) - Z(\bar X_t, \bar Y_t) + \frac{1}{t_0}\mathrm{e}^{-L_2 t}(Y_0 - \bar Y_0) + \mathrm{e}^{-L_2 t}\frac{\mathrm{d}}{\mathrm{d}t}\big(t(t_0 - t)B^*\mathrm{e}^{(t_0-t)A^*}\big)b\Big\}$$

satisfies

$$|\psi_t|^2 \le C_2(|X_0 - \bar X_0|^2 + |Y_0 - \bar Y_0|^2), \quad t \in [0, t_0]$$

for some constant $C_2 > 0$. By the Girsanove theorem,

$$\tilde W_t := W_t + \int_0^t \psi_s \mathrm{d}s, \quad t \in [0, t_0]$$

is a cylindrical Brownian motion on $\mathbb{H}_2$ under the probability measure $\mathrm{d}\mathbb{Q} := R\,\mathrm{d}\mathbb{P}$, where

$$R := \exp\bigg[-\int_0^{t_0}\langle \psi_s, \mathrm{d}W_s\rangle - \frac{1}{2}\int_0^{t_0}|\psi_s|^2\mathrm{d}s\bigg].$$

Rewrite the equation for $(\bar X_t, \bar Y_t)$ as

$$\begin{cases} \mathrm{d}\bar X_t = (A\bar X_t + B\bar Y_t - L_1\bar X_t)\mathrm{d}t, \\ \mathrm{d}\bar Y_t = \big\{Z(\bar X_t, \bar Y_t) - L_2\bar Y_t\big\}\mathrm{d}t + \sigma\mathrm{d}\tilde W_t. \end{cases}$$

By the weak uniqueness of the mild solutions to (4.1) and $(X_{t_0}, Y_{t_0}) = (\bar X_{t_0}, \bar Y_{t_0})$, we obtain

$$(P_{t_0}f(\xi))^2 = (\mathbb{E}_{\mathbb{Q}}f(\bar X_{t_0}, \bar Y_{t_0}))^2 = (\mathbb{E}[Rf(X_{t_0}, Y_{t_0})])^2 \le (P_{t_0}f^2)(\eta)\mathbb{E}R^2 \le (P_{t_0}f^2)(\eta)\mathrm{e}^{c_0|\xi-\eta|^2}$$

for some constant $c_0 > 0$. □

Lemma 4.3. *Assume $(B1)$, $(B2)$ and $(B3)$. Let (X_t, Y_t) solve (4.3) for $X_0 = Y_0 = 0$. If $\lambda_1 > \lambda'$, then there exists a constant $\varepsilon > 0$ such that* $\sup_{t\ge0}\mathbb{E}\mathrm{e}^{\varepsilon(|X_t|^2+|Y_t|^2)} < \infty$.

Proof. By $(B2)$, there exists a constant $c > 0$ such that

$$|Z(x, y)| \le c + K_1|x| + K_2|y|, \quad x, y \in \mathbb{H}.$$

Combining this with (4.3), and noting that $(B1)$ and $(B3)$ imply $L_1 - A + \delta \ge \lambda_1$ and $L_2 \ge \lambda_1$, we obtain

$$\begin{aligned} |X_t| &\le \int_0^t \mathrm{e}^{-\lambda_1(t-s)}(\delta|X_s| + \|B\|\cdot|Y_s|)\mathrm{d}s, \\ |Y_t| &\le \int_0^t \mathrm{e}^{-\lambda_1(t-s)}(c + K_1|X_s| + K_2|Y_s|)\mathrm{d}s + |\xi_t|. \end{aligned} \tag{4.9}$$

By $(B2)$ and $(B3)$, we have

$$\alpha := \frac{1}{2\|B\|}\Big(\delta - K_2 + \sqrt{(K_2-\delta)^2 + 4K_1\|B\|}\Big) \in (0, \infty).$$

Obviously, the definitions of α and λ' in (4.2) imply

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

So,

$$(\alpha\delta + K_1)s + (\alpha\|B\| + K_2)t = \lambda'(\alpha s + t), \quad s,t \geq 0.$$

Combining this with (4.9), we obtain

$$\begin{aligned}\alpha|X_t| + |Y_t| &\leq \int_0^t \mathrm{e}^{-\lambda_1(t-s)}\Big\{c + (\alpha\delta + K_1)|X_s| + (\alpha\|B\| + K_2)|Y_s|\Big\}\mathrm{d}s + |\xi_t| \\ &\leq \lambda' \int_0^t \mathrm{e}^{-\lambda_1(t-s)}(\alpha|X_s| + |Y_s|)\mathrm{d}s + |\xi_t| + \frac{c}{\lambda_1}.\end{aligned}$$

By Gronwall's inequality, this implies

$$\begin{aligned}\alpha|X_t| + |Y_t| &\leq |\xi_t| + \frac{c}{\lambda_1} + \lambda' \int_0^t \mathrm{e}^{-\lambda(t-s)}\Big(|\xi_s| + \frac{c}{\lambda_1}\Big)\mathrm{d}s \\ &\leq |\xi_t| + c_1 + \lambda' \int_0^t \mathrm{e}^{-\lambda(t-s)}|\xi_s|\mathrm{d}s, \quad t \geq 0\end{aligned} \tag{4.11}$$

for some constant $c_1 > 0$ and $\lambda := \lambda_1 - \lambda' > 0$.

Finally, applying Jensen's inequality to the probability measure $\nu(\mathrm{d}s) := \lambda \mathrm{e}^{-\lambda(t-s)}\mathrm{d}s$ on $(-\infty, t]$, we obtain

$$\begin{aligned}&\exp\Big[\varepsilon\Big(\lambda'\int_0^t \mathrm{e}^{-\lambda(t-s)}|\xi_s|\mathrm{d}s\Big)^2\Big] = \exp\Big[\frac{\varepsilon}{\lambda^2}\Big(\lambda'\int_{-\infty}^t 1_{[0,t]}(s)|\xi_s|\nu(\mathrm{d}s)\Big)^2\Big] \\ &\leq \int_{-\infty}^t \exp\Big[\frac{\varepsilon(\lambda')^2}{\lambda^2}1_{[0,t]}(s)|\xi_s|^2\Big]\nu(\mathrm{d}s) \\ &\leq c_2 + c_2\int_0^t \mathrm{e}^{-\lambda(t-s)}\exp\big[c_2\varepsilon|\xi_s|^2\big]\mathrm{d}s, \quad t,\varepsilon \geq 0\end{aligned}$$

for some constant $c_2 > 0$. Combining this with (4.5) and (4.11), we finish the proof. □

Lemma 4.4. *Assume (B1), (B2) and (B3). If $\lambda_1 > \lambda'$, then P_t has a unique invariant probability measure μ, and $\mu(\mathrm{e}^{\varepsilon|\cdot|^2}) < \infty$ holds for some constant $\varepsilon > 0$.*

Proof. According to [19, Proposition 3.1], the Harnack inequality in Lemma 4.2 implies that P_t has at most one invariant probability measure. So, it suffices to prove the existence of μ with $\mu(\mathrm{e}^{\varepsilon|\cdot|^2}) < \infty$ for some constant $\varepsilon > 0$.

Let $(X_t, Y_t)_{t\geq 0}$ solve (4.1) for $X_0 = Y_0 = 0$. For every $t \geq 0$, let μ_t be the distribution of (X_t, Y_t), which is a probability measure on $\mathbb{H}$. By the Markov property, if μ_t converges weakly to a probability measure μ as $t \to \infty$, then μ is an invariant probability measure of P_t and, by Lemma 4.3 and Fatou's lemma, $\mu(\mathrm{e}^{\varepsilon|\cdot|^2}) < \infty$ holds for some constant $\varepsilon > 0$. Therefore, it remains to prove the weak convergence of μ_t as $t \to \infty$.

Consider the L^1-Wasserstein distance

$$W(\nu_1,\nu_2) := \inf_{\pi\in\mathscr{C}(\nu_1,\nu_2)} \int_{\mathbb{H}\times\mathbb{H}} |\cdot|\mathrm{d}\pi$$

for two probability measures ν_1 and ν_2 on $\mathbb{H}\times\mathbb{H}$, where $\mathscr{C}(\nu_1,\nu_2)$ is the set of all couplings of these two measures. If μ_t is a W-Cauchy family as $t\to\infty$, i.e.

$$\lim_{t_1,t_2\to\infty} W(\mu_{t_1},\mu_{t_2}) = 0, \tag{4.12}$$

then it converges weakly as $t\to\infty$, see e.g. [4, Theorem 5.4 and Theorem 5.6].

To prove (4.12), for any $t_2 > t_1 > 0$, let $(X_t,Y_t)_{t\geq 0}$ solve (4.1) for $X_0 = Y_0 = 0$, and let $(\tilde{X}_t,\tilde{Y}_t)_{t\geq t_2-t_1}$ solve the following equation with $\tilde{X}_{t_2-t_1} = \tilde{Y}_{t_2-t_1} = 0$:

$$\begin{cases} \mathrm{d}\tilde{X}_t = (A\tilde{X}_t + B\tilde{Y}_t - L_1\tilde{X}_t)\mathrm{d}t, \\ \mathrm{d}\tilde{Y}_t = \big\{Z(\tilde{X}_t,\tilde{Y}_t) - L_2\tilde{Y}_t\big\}\mathrm{d}t + \sigma\mathrm{d}W_t, \quad t\geq t_2-t_1. \end{cases} \tag{4.13}$$

Then the distribution of (X_{t_2},Y_{t_2}) is μ_{t_2} while that of $(\tilde{X}_{t_2},\tilde{Y}_{t_2})$ is μ_{t_1}. By the definition of W, we have

$$W(\mu_{t_1},\mu_{t_2}) \leq \mathbb{E}(|X_{t_2}-\tilde{X}_{t_2}| + |Y_{t_2}-\tilde{Y}_{t_2}|). \tag{4.14}$$

On the other hand, (4.1), (4.13), $(B2)$ and $(B3)$ imply that for any $t\geq t_2-t_1$,

$$\begin{aligned} |X_t-\tilde{X}_t| &\leq \mathrm{e}^{-\lambda_1(t-t_2+t_1)}|X_{t_2-t_1}| + \int_{t_2-t_1}^t \mathrm{e}^{-\lambda_1(t-s)}(\delta|X_s-\tilde{X}_s| + \|B\|\cdot|Y_s-\tilde{Y}_s|)\mathrm{d}s, \\ |Y_t-\tilde{Y}_t| &\leq \mathrm{e}^{-\lambda_1(t-t_2+t_1)}|Y_{t_2-t_1}| + \int_{t_2-t_1}^t \mathrm{e}^{-\lambda_1(t-s)}(K_1|X_s-\tilde{X}_s| + K_2|Y_s-\tilde{Y}_s|)\mathrm{d}s. \end{aligned}$$

Then by (4.10), for $t\geq t_2-t_1$

$$\begin{aligned} &\alpha|X_t-\tilde{X}_t| + |Y_t-\tilde{Y}_t| \\ &\leq \mathrm{e}^{-\lambda_1(t+t_1-t_2)}(\alpha|X_{t_1}| + |Y_{t_1}|) + \lambda'\int_{t_2-t_1}^t \mathrm{e}^{-\lambda_1(t-s)}(\alpha|X_s-\tilde{X}_s| + |Y_s-\tilde{Y}_s|)\mathrm{d}s. \end{aligned} \tag{4.15}$$

By Gronwall's inequality, we obtain

$$\begin{aligned} \alpha|X_{t_2}-\tilde{X}_{t_2}| + |Y_{t_2}-\tilde{Y}_{t_2}| &\leq (\alpha|X_{t_1}| + |Y_{t_1}|)\mathrm{e}^{-\lambda_1 t_1}\left(1 + \lambda'\int_{t_2-t_1}^{t_2} \mathrm{e}^{\lambda'(t_2-s)}\mathrm{d}s\right) \\ &\leq 2(\alpha|X_{t_1}| + |Y_{t_1}|)\mathrm{e}^{-(\lambda_1-\lambda')t_1}. \end{aligned}$$

Since $\sup_{t\geq 0}\mathbb{E}(|X_t| + |Y_t|) < \infty$ due to Lemma 4.3, this together with (4.14) implies (4.12). The proof is therefore finished. □

Lemma 4.5. *Assume (B1), (B2) and (B3). If $\lambda_1 > \lambda'$, then there exists a constant $C > 0$ such that for any mild solutions (X_t, Y_t) and $(\tilde{X}_t, \tilde{Y}_t)$ of the equation (4.1),*

$$|X_t - \tilde{X}_t| + |Y_t - \tilde{Y}_t| \le C(|X_0 - \tilde{X}_0| + |Y_0 - \tilde{Y}_0|)\mathrm{e}^{-(\lambda_1 - \lambda')t}, \quad t \ge 0.$$

Proof. Similarly to the proof of (4.15), we have

$$\begin{aligned} &\alpha|X_t - \tilde{X}_t| + |Y_t - \tilde{Y}_t| \\ &\le \mathrm{e}^{-\lambda_1 t}(\alpha|X_0 - \tilde{X}_0| + |Y_0 - \tilde{Y}_0|) + \lambda' \int_0^t \mathrm{e}^{-\lambda_1(t-s)}(\alpha|X_s - \tilde{X}_s| + |Y_s - \tilde{Y}_s|)\mathrm{d}s, \quad t \ge 0. \end{aligned}$$

By Gronwall's inequality,

$$\alpha|X_t - \tilde{X}_t| + |Y_t - \tilde{Y}_t| \le \mathrm{e}^{-(\lambda_1 - \lambda')t}(\alpha|X_0 - \tilde{X}_0| + |Y_0 - \tilde{Y}_0|), \quad t \ge 0.$$

This completes the proof. □

5 Some Examples

In this section, we present three examples to illustrate Theorems 3.1 and 4.1, where the first includes the kinetic Fokker-Planck equation discussed in [14] for $V(x) = -\frac{1}{2}|x|^2 + \nabla W$ with small $\|\nabla^2 W\|_\infty$, the second is highly degenerate in the sense that m can be much larger than d, and the last is an infinite-dimensional model.

Example 5.1. Let $d = m$ and σ be invertible, $A = 0, B = I$, and $Z(x, y) = \nabla W(x) - x - y$ for some $W \in C^2(\mathbb{R}^d)$. If $\|\nabla^2 W\|_\infty < 1$ is small enough such that

$$1 > \inf_{r_0 \in (0,1)} \Big\{ \frac{\|\nabla^2 W\|_\infty^2}{2r_0(1 - \|\nabla^2 W\|_\infty)(1 + \sqrt{1 + 4r_0})} + \frac{r_0}{2}\big(1 + \sqrt{1 + 4r_0}\big) \Big\}, \tag{5.1}$$

then all assertions in Theorem 3.1 hold. In particular, (5.1) holds if $\|\nabla^2 W\|_\infty \le \frac{1}{2}$.

Proof. It is trivial that (A1) and (A2) hold. To verify (A3), let $r > 0$ and $r_0 \in (0, 1) = (0, \|B\|^{-1})$. By $A = 0, B = I$ and the formulation of Z, we have

$$\begin{aligned} &\langle r^2(x - \bar{x}) + rr_0 B(y - \bar{y}), A(x - \bar{x}) + B(y - \bar{y})\rangle + \langle Z(x, y) - Z(\bar{x}, \bar{y}), y - \bar{y} + rr_0 B^*(x - \bar{x})\rangle \\ &= (r^2 - 1 - rr_0)\langle x - \bar{x}, y - \bar{y}\rangle + rr_0|y - \bar{y}|^2 + \langle \nabla W(x) - \nabla W(\bar{x}), y - \bar{y} + rr_0(x - \bar{x})\rangle \\ &\qquad - rr_0|x - \bar{x}|^2 - |y - \bar{y}|^2. \end{aligned}$$

Take

$$r = \frac{1}{2}\big(1 + \sqrt{1 + 4r_0}\big) \tag{5.2}$$

such that $r^2 - 1 - rr_0 = 0$, we obtain

$$\langle r^2(x-\bar x) + rr_0 B(y-\bar y), A(x-\bar x) + B(y-\bar y)\rangle + \langle Z(x,y) - Z(\bar x,\bar y), y - \bar y + rr_0 B^*(x-\bar x)\rangle$$
$$\le -\big(rr_0 - \|\nabla^2 W\|_\infty rr_0 - \gamma\big)|x-\bar x|^2 - \Big(1 - rr_0 - \frac{\|\nabla W\|_\infty^2}{4\gamma}\Big)|y-\bar y|^2, \quad \gamma > 0.$$

Therefore, (A_3) holds for some constants $r_0 \in (0,1)$ and $\theta > 0$ if

$$1 > \inf_{r_0\in(0,1)} \inf_{\gamma\in(0, rr_0 - \|\nabla^2 W\|_\infty rr_0)} \Big(rr_0 + \frac{\|\nabla W\|_\infty^2}{4\gamma}\Big),$$

which is equivalent to (5.1) due to (5.2). It remains to prove (5.1) for $\|\nabla^2 W\|_\infty \le \frac{1}{2}$. Since (5.1) is trivial for $\|\nabla^2 W\|_\infty = 0$, we assume that $\|\nabla^2 W\|_\infty \in (0, \frac{1}{2}]$. In this case we simply take $r_0 = \|\nabla^2 W\|_\infty$ such that

$$\begin{aligned}&\frac{\|\nabla^2 W\|_\infty^2}{2r_0(1-\|\nabla^2 W\|_\infty)(1+\sqrt{1+4r_0})} + \frac{r_0}{2}\big(1+\sqrt{1+4r_0}\big)\\ &< \|\nabla^2 W\|_\infty\Big(\frac{1}{2} + \frac{1}{2}\big(1+\sqrt{3}\big)\Big) \le \frac{1}{2}\Big(1 + \frac{1}{2}\sqrt{3}\Big) < 1.\end{aligned}$$

$\square$

Example 5.2. Let σ be invertible, $m = kd$ for some natural number $k \ge 2$, and

$$\begin{aligned}&By = (0,\cdots,0,y) \in \mathbb{R}^{kd}, \quad y \in \mathbb{R}^d,\\ &Z(x,y) = b(y) - x_k, \quad y \in \mathbb{R}^d, x = (x_1,x_2,\cdots,x_k) \in \mathbb{R}^{kd},\\ &A(x_1,x_2,\cdots,x_k) = (\gamma x_2 - x_1, \gamma x_3 - x_2, \cdots, \gamma x_k - x_{k-1}, 0), \quad x_1,\cdots,x_k \in \mathbb{R}^d,\end{aligned}$$

where $\gamma \ne 0$ is a constant, and $b : \mathbb{R}^d \to \mathbb{R}^d$ satisfies

$$|b(y) - b(\bar y)| \le K|y-\bar y|, \quad \langle b(y) - b(\bar y), y - \bar y\rangle \le -\beta|y-\bar y|^2, \quad y,\bar y \in \mathbb{R}^d \tag{5.3}$$

for some constants $K, \beta > 0$. If

$$0 < |\gamma| < 1 \wedge \frac{2\beta}{2+K^2}, \tag{5.4}$$

then assertions in Theorem 3.1 hold.

Proof. It is easy to see that when $\gamma \ne 0$, the rank condition in $(A1)$ holds. Since b is Lipchitz continuous and σ is invertible, by Theorem 3.1 it suffices to verify $(A3)$. We simply take

$r=1$. For any $r_0\in(0,1)=(0,\|B\|^{-1})$, we have

$$
\begin{aligned}
&\langle r^2(x-\bar{x})+rr_0B(y-\bar{y}), A(x-\bar{x})+B(y-\bar{y})\rangle+\langle Z(x,y)-Z(\bar{x},\bar{y}), y-\bar{y}+rr_0B^*(x-\bar{x})\rangle\\
&=r_0|y-\bar{y}|^2+\sum_{i=1}^{k-1}\Big\{\gamma\langle x_i-\bar{x}_i, x_{i+1}-\bar{x}_{i+1}\rangle-|x_i-\bar{x}_i|^2\Big\}\\
&\qquad+\big\langle b(y)-b(\bar{y}), y-\bar{y}+r_0(x_k-\bar{x}_k)\big\rangle-r_0|x_k-\bar{x}_k|^2\\
&\le-(\beta-r_0)|y-\bar{y}|^2-r_0|x_k-\bar{x}_k|^2+r_0K|y-\bar{y}|\cdot|x_k-\bar{x}_k|\\
&\qquad-\sum_{i=1}^{k-1}\Big\{|x_i-\bar{x}_i|^2-\frac{|\gamma|}{2}|x_i-\bar{x}_i|^2-\frac{|\gamma|}{2}|x_{i+1}-\bar{x}_{i+1}|^2\Big\}\\
&\le-\sum_{i=1}^{k-1}(1-|\gamma|)|x_i-\bar{x}_i|^2-\Big(r-\frac{|\gamma|}{2}-\frac{r_0K^2}{4\alpha}\Big)|x_k-\bar{x}_k|^2-(\beta-r_0-\alpha r_0)|y-\bar{y}|^2,\quad \alpha>0.
\end{aligned}
$$

So, $(A3)$ holds for some $\theta>0$ provided $|\gamma|<1$ and

$$\sup_{r_0\in(0,1\wedge\frac{\beta}{1+\alpha}),\alpha>0}\Big(r_0-\frac{|\gamma|}{2}-\frac{K^2r_0}{4\alpha}\Big)>0.$$

Letting $r_0\uparrow 1\wedge\frac{\beta}{1+\alpha}$, we conclude that $(A3)$ holds provided $|\gamma|<1$ and

$$\sup_{\alpha>0}\Big(1\wedge\frac{\beta}{1+\alpha}\Big)\Big(1-\frac{K^2}{4\alpha}\Big)>\frac{|\gamma|}{2}.$$

By taking $\alpha=\frac{1}{2}K^2$ we see that this inequality follows from (5.4). $\square$

Finally, we present an example for Theorem 4.1 in the spirit of Example 5.2 that $\mathbb{H}_2$ is a subspace of $\mathbb{H}_1$.

Example 5.3. Let $\{u_i\}_{i\ge1}$ be an orthonormal basis on $\mathbb{H}_1$, and let $\mathbb{H}_2=\overline{\mathrm{span}}\{u_{2i}:i\ge1\}$. Take $B=I_{\mathbb{H}_2}$ and

$$L_1u_{2i}=\lambda_iu_{2i},\ L_1u_{2i-1}=\lambda_iu_{2i-1},\quad i\ge1,$$

where $0<\lambda_i\uparrow\infty$ with $\sum_{i\ge1}\lambda_i^{-1}<\infty$. Moreover, let $L_2=L_1|_{\mathbb{H}_2}$ and

$$Ax=\gamma\lambda_1\sum_{i=1}^{\infty}\langle x,u_{2i}\rangle u_{2i-1},\quad x\in\mathbb{H}_1$$

for some constant $\gamma\in\mathbb{R}$. Finally, let Z satisfy

$$|Z(x,y)-Z(\bar{x},\bar{y})|\le\alpha\lambda_1|x-\bar{x}|+\beta\lambda_1|y-\bar{y}|$$

for some constants $\alpha,\beta\ge0$. Then all assertions in Theorem 3.1 hold provided

$$\sqrt{1+\gamma^2}+4\beta+\sqrt{(2\beta-1-\sqrt{1+\gamma^2})^2+8\alpha}<7. \tag{5.5}$$

Proof. It is easy to see that $BL_2 = L_1 B, AL_1 = L_1 A$. According to Theorem 4.1, it suffices to prove

(a) For some $\delta > 0$ such that $L_1 - A \geq \lambda_1 - \delta$ and the condition (4.2) hold.

(b) For any $t_0 > 0$, Q_{t_0} is invertible on $\mathbb{H}_1$.

Proof of (a) We have

$$\begin{aligned}
\langle (L_1 - A)x, x\rangle &= \langle L_2 \pi x, \pi x\rangle - \langle Ax, x\rangle \\
&\geq \lambda_1 \sum_{i\geq 1} \langle x, u_{2i}\rangle^2 - \gamma \sum_{i\geq 1} \langle x, u_{2i}\rangle \langle x, u_{2i-1}\rangle \\
&\geq (\lambda_1 - \delta) \sum_{i\geq 1} \langle x, u_{2i}\rangle^2 - \frac{\gamma^2}{4\delta} \sum_{i\geq 1} \langle x, u_{2i-1}\rangle^2, \quad x \in \mathbb{H}_1.
\end{aligned}$$

Taking

$$\delta = \frac{1 + \sqrt{1+\gamma^2}}{2} \lambda_1$$

such that $\frac{\gamma^2}{4\delta} = \delta - \lambda_1$, we have $L_1 - A \geq \lambda_1 - \delta$ as required, and the condition (5.5) is equivalent to (4.2).

Proof of (b) We may simply assume $\gamma\lambda_1 = 1$, so that

$$A^* x = \sum_{i=1}^{\infty} \langle x, u_{2i-1}\rangle u_{2i}, \quad x \in \mathbb{H}_1.$$

Since $A^2 = (A^*)^2 = 0$ and BB^* is the orthogonal projection onto $\mathbb{H}_2$, for any $x \in \mathbb{H}_1$ we have

$$\begin{aligned}
\mathrm{e}^{sA} BB^* \mathrm{e}^{sA^*} x &= (I + sA) BB^* \{x + sA^* x\} \\
&= \sum_{i=1}^{\infty} \big(\langle x, u_{2i}\rangle + s\langle x, u_{2i-1}\rangle\big) \big\{u_{2i} + s u_{2i-1}\big\}.
\end{aligned}$$

Then

$$\begin{aligned}
\langle Q_{t_0} x, x\rangle &= \sum_{i=1}^{\infty} \int_0^{t_0} \big\{\langle x, u_{2i}\rangle^2 + 2s\langle x, u_{2i-1}\rangle \langle x, u_{2i}\rangle + s^2 \langle x, u_{2i-1}\rangle^2\big\} \mathrm{d}s \\
&= t_0 \sum_{i=1}^{\infty} \Big\{\langle x, u_{2i}\rangle^2 + t_0 \langle x, u_{2i-1}\rangle \langle x, u_{2i}\rangle + \frac{t_0^2}{3} \langle x, u_{2i-1}\rangle^2\Big\} \\
&\geq t_0 \sum_{i=1}^{\infty} \Big\{(1-r)\langle x, u_{2i}\rangle^2 + \Big(\frac{1}{3} - \frac{1}{4r}\Big) t_0^2 \langle x, u_{2i-1}\rangle^2\Big\}, \quad r > 0.
\end{aligned}$$

Taking $r \in (0,1)$ but close enough to 1, we conclude that $\langle Q_{t_0} x, x\rangle \geq c|x|^2$ holds for some constant $c > 0$ and all $x \in \mathbb{H}_1$. Therefore, Q_{t_0} is invertible. □

Acknowledgement. The author would like to thank the referee for helpful comments.

References

[1] 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.

[2] J. Bao, F.-Y. Wang, C. Yuan, *Hypercontractivity for functional stochastic differential equations,* Stoch. Proc. Appl. 125(2015) 3636–3656.

[3] F. Baudoin, *Bakry-Emery meet Villani,* arXiv:1308.4938

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

[5] G.D. Prato, J. Zabczyk, *Stochastic Equations in Infinite Dimensions,* Cambridge University Press, 1992.

[6] J. Dolbeault, C. Mouhot, C. Schmeiser, *Hypocoercivity for kinetic equations with linear relaxation terms,* C. R. Math. Acad. Sci. Paris 347(2009), 511–516.

[7] R. Duan, *Hypocoercivity of linear degenerately dissipative kinetic equations,* Nonlinearity 24(2011), 2165–2189.

[8] S. Gadat, L. Miclo, *Spectral decompositions and L^2-operator norms of toy hypocoercive semi-groups,* Kinetic and related models 6(2013), 317–372.

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

[10] M. Grothaus, P. Stilgenbauer, *Hypocoercivity for kolmogorov backward evolution equations and applications,* J. Funct. Anal. 267(2014), 3515–3556.

[11] A. Guillin, F.-Y. Wang, *Degenerate Fokker-Planck equations : Bismut formula, gradient estimate and Harnack inequality,* J. Diff. Equat. 253(2012), 20–40.

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

[13] T. Seidman, *How violent are fast controls?,* Math. Control Signals Systems 1(1988), 89–95.

[14] C. Villani, *Hypocoercivity,* Mem. Amer. Math. Soc. 202(950)(2009).

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

[16] F.-Y. Wang, *Spectral gap for hyperbounded operators,* Proc. Amer. Math. Soc. 132(2004), 2629–2638.

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

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

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

[20] F.-Y. Wang, T. Zhang, *Gradient estimates for stochastic evolution equations with non-Lipschitz coefficients,* J. Math. Anal. Appl. 365(2010), 1–11.

[21] F.-Y. Wang, T. Zhang, *Degenerate SDEs in Hilbert spaces with rough drifts,* Infin. Dimens. Anal. Quant. Probab. Relat. Top. 18 (2015), no. 4, 1550026, 25 pp

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