Paper:
Wang, F. (2017). Distribution dependent SDEs for Landau type equations. *Stochastic Processes and their Applications*
http://dx.doi.org/10.1016/j.spa.2017.05.006

Distribution-Dependent SDEs for Landau Type Equations*

Feng-Yu Wang

Laboratory of Mathematical and Complex Systems, Beijing Normal University, Beijing 100875, China

Department of Mathematics, Swansea University, Singleton Park, SA2 8PP, United Kingdom

wangfy@bnu.edu.cn, F.-Y.Wang@swansea.ac.uk

Abstract

The distribution-dependent stochastic differential equations (DDSDEs) describe stochastic systems whose evolution is determined by both the microcosmic site and the macrocosmic distribution of the particle. The density function associated with a DDSDE solves a Landau type nonlinear PDE. Due to the distribution-dependence, some standard techniques developed for SDEs do not apply. By iterating in distributions, a strong solution is constructed using SDEs with control. By proving the uniqueness, the distribution of solutions is identified with a nonlinear semigroup P_t^* on the space of probability measures. The exponential contraction as well as Harnack inequalities and applications are investigated for the nonlinear semigroup P_t^* using coupling by change of measures. The main results are illustrated by homogeneous Landau equations.

AMS subject Classification: 60J75, 47G20, 60G52.
Keywords: Distribution-dependent SDEs, homogeneous Landau equation, Wasserstein distance, exponential convergence, Harnack inequality.

1 Introduction

A fundamental application of the Itô SDE is to solve Kolmogorov's problem [13] of determining Markov processes whose distribution density satisfies the Fokker-Planck-Kolmogorov equation. Let W_t be the d-dimensional Brownian motion on a complete probability space with nature filtration $(\Omega, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$, and let $b : \mathbb{R}^d \to \mathbb{R}^d; \sigma : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ be measurable. Then the distribution density of the solution to the SDE

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t \tag{1.1}$$

*Supported in part by NNSFC (11626245, 11431014).

satisfies the parabolic equation

$$\partial_t f_t = \frac{1}{2}\sum_{i,j=1}^d \partial_i\partial_j\big\{(\sigma\sigma^*)_{ij} f_t\big\} - \sum_{i=1}^d \partial_i\{b_i f_t\}, \tag{1.2}$$

which describes the time evolution of the probability density function of the velocity of a particle under the influence of drag forces and random forces. If b and σ are "almost" locally Lipchitzian, then the SDE (1.1) has a unique strong solution up to life time (c.f. [7]). When σ is invertible (i.e. the SDE is non-degenerate), this condition has been largely weakened as $|b| + |\nabla\sigma| \in L^p_{loc}(\mathrm{d}x)$ for some $p > d$, see [26] and references within.

However, in many cases the distribution density satisfies a nonlinear PDE, for instance, the Landau type equation

$$\partial_t f_t = \frac{1}{2}\mathrm{div}\bigg\{\int_{\mathbb{R}^d} a(\cdot - z)\big(f_t(z)\nabla f_t - f_t\nabla f_t(z)\big)\mathrm{d}z\bigg\}, \tag{1.3}$$

for some reference coefficient $a : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$. This includes the homogenous Landau equation where $d = 3$ and

$$a(x) = |x|^{2+\gamma}\Big(I - \frac{x\otimes x}{|x|^2}\Big), \quad x \in \mathbb{R}^3$$

for some constant $\gamma \in [-3, 1]$. When $\gamma \in [0, 1]$, the existence, uniqueness, regularity estimates, and exponential convergence have been investigated for good enough initial distributions, see [5, 6, 4] and references within. To describe the solution of (1.3) using stochastic processes, consider the following distribution-dependent SDE (DDSDE) for $b = \mathrm{div}a$ and σ such that $\sigma\sigma^* = a$:

$$\mathrm{d}X_t = (b * \mathscr{L}_{X_t})(X_t)\mathrm{d}t + (\sigma * \mathscr{L}_{X_t})(X_t)\mathrm{d}W_t, \tag{1.4}$$

where $\mathscr{L}_\xi$ denotes the distribution of a random variable ξ, and

$$(f * \mu)(x) := \int_{\mathbb{R}^d} f(x - z)\mu(\mathrm{d}z)$$

for a function f and a probability measure μ. By Itô's formula and the integration by parts formula, the distribution density of X_t is a weak solution to (1.3). For the homogenous Landau equation with $\gamma \in [0, 1]$ and initial distribution density f_0 satisfying

$$\int_{\mathbb{R}^3} f_0(x)\big(f_0(x) + \mathrm{e}^{|x|^\alpha}\big)\mathrm{d}x < \infty \ \text{ for some } \alpha > \gamma, \tag{1.5}$$

the existence and uniqueness of weak solutions to (1.4) has been proved in [8] by an approximation argument using particle systems. This approximation is known as *propagation of chaos* according to Kac [12], see also [9, 10] and references within.

In this paper, we aim to investigate the (pathwise) strong solutions of (1.4) and characterize their distribution properties.

In general, for measurable maps

$$b : [0, \infty) \times \mathbb{R}^d \times \mathscr{P}(\mathbb{R}^d) \to \mathbb{R}^d; \quad \sigma : [0, \infty) \times \mathbb{R}^d \times \mathscr{P}(\mathbb{R}^d) \to \mathbb{R}^d \otimes \mathbb{R}^d,$$

we consider the following DDSDE on $\mathbb{R}^d$:

$$\mathrm{d}X_t = b_t(X_t, \mathscr{L}_{X_t})\,\mathrm{d}t + \sigma_t(X_t, \mathscr{L}_{X_t})\,\mathrm{d}W_t. \tag{1.6}$$

When more than one probability measures on Ω are concerned, we use $\mathscr{L}_{X_t}|_{\mathbb{P}}$ instead of $\mathscr{L}_{X_t}$ to emphasize the distribution under probability $\mathbb{P}$. Due to technical reasons, we will restrict ourselves to the following P_t^*-invariant subspace of $\mathscr{P}$ for some $\theta \in [1, \infty)$:

$$\mathscr{P}_\theta := \left\{ \nu \in \mathscr{P} : \ \nu(|\cdot|^\theta) := \int_{\mathbb{R}^d} |x|^\theta \nu(\mathrm{d}x) < \infty \right\},$$

which is a polish space under the L^θ-Wasserstein distance

$$\mathbb{W}_\theta(\mu_1, \mu_2) := \inf_{\pi \in \mathscr{C}(\mu_1, \mu_2)} \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^\theta \pi(\mathrm{d}x, \mathrm{d}y) \right)^{\frac{1}{\theta}}, \quad \mu_1, \mu_2 \in \mathscr{P}_\theta,$$

where $\mathscr{C}(\mu_1, \mu_2)$ is the set of all couplings for μ_1 and μ_2.

The following definition is standard in the literature of SDEs.

Definition 1.1. (1) For any $s \geq 0$, a continuous adapted process $(X_t)_{t \geq s}$ on $\mathbb{R}^d$ is called a (strong) solution of (1.6) from time s, if

$$\int_s^t \mathbb{E}\{|b_r(X_r, \mathscr{L}_{X_r})| + \|\sigma_r(X_r, \mathscr{L}_{X_r})\|^2\}\mathrm{d}r < \infty, \quad t > s,$$

and $\mathbb{P}$-a.s.,

$$X_t = X_s + \int_s^t b_r(X_r, \mathscr{L}_{X_r})\mathrm{d}r + \int_s^t \sigma_r(X_r, \mathscr{L}_{X_r})\mathrm{d}W_r, \quad t \geq s.$$

We say that (1.6) has (strong or pathwise) existence and uniqueness, if for any $s \geq 0$ and $\mathscr{F}_s$-measurable random variable $X_{s,s}$ with $\mathbb{E}|X_{s,s}|^2 < \infty$, the equation from time s has a unique solution $(X_{s,t})_{t \geq s}$. We simply denote $X_{0,t} = X_t$.

(2) A couple $(\tilde{X}_t, \tilde{W}_t)_{t \geq s}$ is called a weak solution to (1.6) from time s, if $\tilde{W}_t$ is the d-dimensional Brownian motion with respect to a complete filtration probability space $(\tilde{\Omega}, \{\tilde{\mathscr{F}}_t\}_{t \geq 0}, \tilde{\mathbb{P}})$, and $\tilde{X}_t$ solves the DDSDE

$$\mathrm{d}\tilde{X}_t = b_t(\tilde{X}_t, \mathscr{L}_{\tilde{X}_t}|_{\tilde{\mathbb{P}}})\mathrm{d}t + \sigma_t(\tilde{X}_t, \mathscr{L}_{\tilde{X}_t}|_{\tilde{\mathbb{P}}})\mathrm{d}\tilde{W}_t, \quad t \geq s. \tag{1.7}$$

(3) (1.6) is said to have weak uniqueness in $\mathscr{P}_\theta$, if for any $s \geq 0$, any two weak solutions of the equation from time s with common initial distribution in $\mathscr{P}_\theta$ are equal in law. Precisely, if $s \geq 0$ and $(\bar{X}_{s,t}, \bar{W}_t)_{t \geq s}$ with respect to $(\bar{\Omega}, \{\bar{\mathscr{F}}_t\}_{t \geq 0}, \bar{\mathbb{P}})$ and $(\tilde{X}_{s,t}, \tilde{W}_t)_{t \geq s}$ with respect to $(\tilde{\Omega}, \{\tilde{\mathscr{F}}_t\}_{t \geq 0}, \tilde{\mathbb{P}})$ are weak solutions of (1.6), then $\mathscr{L}_{\bar{X}_{s,s}}|_{\bar{\mathbb{P}}} = \mathscr{L}_{\tilde{X}_{s,s}}|_{\tilde{\mathbb{P}}}$ implies $\mathscr{L}_{\bar{X}_{s,\cdot}}|_{\bar{\mathbb{P}}} = \mathscr{L}_{\tilde{X}_{s,\cdot}}|_{\tilde{\mathbb{P}}}$.

When (1.6) has strong existence and uniqueness, the solution $(X_t)_{t\ge 0}$ is a Markov process in the sense that for any $s \ge 0$, $(X_t)_{t\ge s}$ is determined by solving the equation from time s with initial state X_s. More precisely, letting $\{X_{s,t}(\xi)\}_{t\ge s}$ denote the solution of the equation from time s with initial state $X_{s,s} = \xi$, the existence and uniqueness imply

$$X_{s,t}(\xi) = X_{r,t}(X_{s,r}(\xi)), \quad t \ge r \ge s \ge 0, \xi \text{ is } \mathscr{F}_s\text{-measurable with } \mathbb{E}|\xi|^\theta < \infty. \tag{1.8}$$

However, in general the solution is not strong Markovian since we do not have $\mathscr{L}_{X_\tau} = \mathscr{L}_{X_t}$ on the set $\{\tau = t\}$ for a stopping time τ and $t > 0$. Moreover, the associated Makov operators $(P_{s,t})_{t\ge s}$ given by

$$P_{s,t}f(x) := \mathbb{E}f(X_{s,t}(x)), \quad x \in \mathbb{R}^d, f \in \mathscr{B}_b(\mathbb{R}^d)$$

is not a semigroup, see (1.10) below.

When the DDSDE has $\mathscr{P}_\theta$-weak uniqueness (in the classical case it follows from the pathwise uniqueness according to Yamada-Watanabe), we may define a semigroup $(P_t^*)_{t\ge s}$ on $\mathscr{P}_\theta$ by letting $P_{s,t}^*\mu = \mathscr{L}_{X_{s,t}}$ for $\mathscr{L}_{X_{s,s}} = \mu \in \mathscr{P}_\theta$. Indeed, by (1.8) we have

$$P_{s,t}^* = P_{r,t}^* P_{s,r}^*, \quad t \ge r \ge s \ge 0. \tag{1.9}$$

To see that $(P_{s,t})_{t\ge s}$ is not a semigroup, we write

$$(P_{s,t}f)(\mu) = (P_{s,t}^*\mu)(f) := \int_{\mathbb{R}^d} f\mathrm{d}(P_{s,t}^*\mu), \quad f \in \mathscr{B}_b(\mathbb{R}^d), t \ge 0, \mu \in \mathscr{P}_\theta.$$

Then $P_{s,t}f(x) = (P_{s,t}f)(\delta_x)$, where δ_x is the Dirac measure at point x. Since $(\mathscr{L}_{X_{s,t}})_{t\ge s}$ solves a nonlinear equation as indicated in the beginning, the semigroup $P_{s,t}^*$ is nonlinear; i.e.

$$P_{s,t}^*\mu \ne \int_{\mathbb{R}^d} (P_{s,t}^*\delta_x)\mu(\mathrm{d}x), \quad t > s \ge 0$$

for a non-trivial distribution μ. In other words, in general

$$(P_{s,t}f)(\mu) \ne \mu(P_{s,t}f) := \int_{\mathbb{R}^d} P_{s,t}f\mathrm{d}\mu, \quad t > s \ge 0,$$

so that

$$\begin{aligned}(P_{s,t}f)(\mu) &:= \int_{\mathbb{R}^d} f\mathrm{d}(P_{s,t}^*\mu) = \int_{\mathbb{R}^d} f\mathrm{d}(P_{r,t}^*P_{s,r}^*\mu) = (P_{r,t}f)(P_{s,r}^*\mu) \\ &\ne \int_{\mathbb{R}^d} (P_{r,t}f)\mathrm{d}(P_{s,r}^*\mu) = (P_{s,r}(P_{r,t}f))(\mu), \quad t > s \ge 0.\end{aligned} \tag{1.10}$$

Although the semigroup $P_{s,t}^*$ is nonlinear, we may also investigate the ergodicity in the time homogeneous case when σ_t and b_t do not depend on t. In this case we have $P_{s,t}^* = P_{t-s}^*$ for $t \ge s \ge 0$. We call $\mu \in \mathscr{P}_\theta$ an invariant probability measure of P_t^* if $P_t^*\mu = \mu$ for all $t \ge 0$, and we call the solution ergodic if there exists $\mu \in \mathscr{P}_\theta$ such that $\lim_{t\to\infty} P_t^*\nu = \mu$ weakly for any $\nu \in \mathscr{P}_\theta$. Obviously, the ergodicity implies that P_t^* has a uniqueness invariant probability measure.

When b and σ are bounded and Lipschitz continuous in $(x,\nu) \in \mathbb{R}^d \times \mathbb{W}_2$, the weak solution of (1.6) has been constructed in [16] by using propagation of chaos. In this paper, we investigate the existence, uniqueness and distribution properties of the strong solutions. To explain the difficulty of the study, let us recall some standard techniques developed for (1.1) with locally bounded coefficients. Firstly, by a truncation argument one reduces an SDE with locally bounded coefficients to that with bounded coefficients, so that when σ is invertible the existence of weak solutions is ensured by the Girsanov transform and the uniqueness follows from Zvonkin type argument, see e.g. [26] and references within. Then the SDE has a unique strong solution according to Yamada-Watanabe's principle [25]. However, these techniques do not apply to DDSDEs: since the coefficients depend on the distribution which is not pathwisely determined, the truncation argument and Yamada-Watanabe's principle do not work; since the distribution of solution depends on the reference probability measure, the Girsanov transform method is invalid for the construction of weak solutions. Moreover, due to the lack of strong Markovian property, one can not let the marginal processes move together after the coupling time, so that the classical coupling argument does not apply. To overcome the difficulty caused by distribution-dependence, we will approximate the DDSDE (1.6) using classical ones by iterating in distributions, see Lemma 2.3 below. This enables us to construct the strong solution. However, since the approximating SDEs depend on the initial distributions, this method does not provide other properties from existing results for classical SDEs. Fortunately, we are able to develop coupling argument to investigate the $\mathbb{W}_2$-exponential convergence, Harnack inequality and applications for the associated nonlinear semigroup.

In Section 2, we investigate the existence, uniqueness and time-space continuity of solutions. In Section 3, we study the $\mathbb{W}_2$-exponential contraction of P_t^*, which implies the exponential ergodicity in the time-homogenous case. In Sections 4 and 5, we use coupling by change of measures to establish Harnack and shift Harnack inequalities and make applications. Finally, in Section 6, we apply the main results to specific models including the homogeneous Landau equation. These results provide pointwise estimates on the distributions, which are essentially different from existing results on L^p-estimates and Sobolev regularities derived in [5, 6, 4] for the homogeneous Landau equation.

2 Existence, uniqueness and time-space continuity

As already explained in Introduction that the distribution dependence of coefficients may cause trouble in the study of DDSDEs. To get rid of the distribution dependence, we will iterate (1.6) in distributions. To prove the convergence of solutions to iterating SDEs, we make the following assumptions on the continuity, monotonicity and growth of coefficients.

$(H1)$ (Continuity) For every $t \geq 0$, b_t is continuous on $\mathbb{R}^d \times \mathscr{P}_\theta$. Moreover, there exist increasing $K_{\sigma,1}, K_{\sigma,2} \in C([0,\infty);(0,\infty))$ such that

$$\|\sigma_t(x,\mu) - \sigma_t(y,\nu)\|^2 \leq K_{\sigma,1}(t)|x-y|^2 + K_{\sigma,2}\mathbb{W}_\theta(\mu,\nu)^2, \quad t \geq 0, x, y \in \mathbb{R}^d, \mu, \nu \in \mathscr{P}_\theta.$$

$(H2)$ (Monotonicity) There exists increasing $K_{b,1}, K_{b,2} \in C([0,\infty);(0,\infty))$ such that

$$2\big\langle b_t(x,\mu) - b_t(y,\nu), x-y\big\rangle \le K_{b,1}(t)|x-y|^2 + K_{b,2}(t)\mathbb{W}_\theta(\mu,\nu)|x-y|,$$
$$t\ge 0; x,y\in\mathbb{R}^d; \mu,\nu\in\mathscr{P}_\theta.$$

$(H3)$ (Growth) b is bounded on bounded sets in $[0,\infty)\times\mathbb{R}^d\times\mathscr{P}_\theta$, and there exists increasing $K_{b,3}\in C([0,\infty);(0,\infty))$ such that

$$|b_t(0,\mu)|^\theta \le K_{b,3}(t)\big\{1++\mu(|\cdot|^\theta)\big\},\quad x\in\mathbb{R}^d, t\ge 0, \mu\in\mathscr{P}_\theta.$$

2.1 Main results

We first consider the existence, uniqueness and $\mathbb{W}_\theta$-Lipschitz continuous in initial distributions.

Theorem 2.1. *Assume $(H1)$-$(H3)$ for some $\theta\in[1,\infty)$ such that $K_{\sigma,2}=0$ when $\theta<2$.*

(1) *The DDSDE (1.6) has strong/weak existence and uniqueness with initial distributions in $\mathscr{P}_\theta$. Moreover, for any $p\ge\theta$ and $s\ge 0$, $\mathbb{E}|X_{s,s}|^p<\infty$ implies*

$$\mathbb{E}\sup_{t\in[s,T]}|X_{s,t}|^p<\infty,\quad T\ge t\ge s\ge 0. \tag{2.1}$$

(2) *There exists increasing $\psi:[0,\infty)\to[0,\infty)$ such that for any two solutions $X_{s,t}$ and $Y_{s,t}$ of (1.6) with $\mathscr{L}_{X_{s,s}},\mathscr{L}_{Y_{s,s}}\in\mathscr{P}_\theta$,*

$$\mathbb{E}|X_{s,t}-Y_{s,t}|^\theta \le \big(\mathbb{E}|X_{s,s}-Y_{s,s}|^\theta\big)\mathrm{e}^{\int_s^t\psi(r)\mathrm{d}r},\quad t\ge s\ge 0. \tag{2.2}$$

Consequently,

$$\lim_{\mathbb{E}|X_{s,s}-Y_{s,s}|^\theta\to 0}\mathbb{P}\Big(\sup_{r\in[s,t]}|X_{s,r}-Y_{s,r}|\ge\varepsilon\Big)=0,\quad t>s\ge 0,\varepsilon>0; \tag{2.3}$$

and

$$\mathbb{W}_\theta(P_{s,t}^*\mu_0, P_{s,t}^*\nu_0)^\theta \le \mathbb{W}_2(\mu_0,\nu_0)^\theta\mathrm{e}^{\int_s^t\psi(r)\mathrm{d}r},\quad t\ge s\ge 0. \tag{2.4}$$

Next, we consider the continuity of $X_t(x)$ in $t,x)\in[0,\infty)\times\mathbb{R}^d$, where $X_t(x)$ for $X_0=x$. Since assumptions $(H1)$-$(H3)$ are weaker for larger θ, and $\delta_x\in\mathscr{P}_\theta$ for any $\theta\ge 1$, by Theorem 2.1 the DDSDE (1.6) has a unique solution $X_t(x)$ for $X_0=x$. The next result says that X is continuous in $(t,x)\in[0,\infty)\times\mathbb{R}^d$ provided b has a polynomial growth. Because the coefficients depend on the distribution of solution, it seems hard to prove the flow property, for instance to prove that $\mathbb{P}$-a.s. for all t the map $X_t(\cdot):\mathbb{R}^d\to\mathbb{R}^d$ is a diffeormorphism, by using techniques developed in the classical setting. So, we leave the study to the future.

Theorem 2.2. *Assume* ($H1$)-($H3$) *for some* $\theta \geq 1$. *If there exists* $p \geq 1$ *such that*

$$|b_t(x,\mu)| \leq K(t)\big\{1+|x|^p+\mu(|\cdot|^p)\big\},\quad t\geq 0, x\in\mathbb{R}^d \tag{2.5}$$

holds for some increasing function $K:[0,\infty)\to(0,\infty)$, *then* $\mathbb{P}$*-a.s. the map*

$$[0,\infty)\times\mathbb{R}^d \ni (t,x)\mapsto X_t(x)\in\mathbb{R}^d$$

is continuous.

To prove these results, we first approximate (1.6) using classical SDEs by iterating in distributions.

2.2 An approximation argument using classical SDEs

We fixed $s\geq 0$ and $\mathscr{F}_s$-measurable random variable $X_{s,s}$ on $\mathbb{R}^d$ with $\mathbb{E}|X_{s,s}|^\theta<\infty$. Let

$$X_{s,t}^{(0)} = X_{s,s},\quad \mu_{s,t}^{(0)} = \mathscr{L}_{X_{s,s}},\quad t\geq s.$$

For any $n\geq 1$, let $(X_{s,t}^{(n)})_{t\geq s}$ solve the classical SDE

$$\mathrm{d}X_{s,t}^{(n)} = b_t(X_{s,t}^{(n)},\mu_{s,t}^{(n-1)})\mathrm{d}t + \sigma_t(X_{s,t}^{(n)},\mu_{s,t}^{(n-1)})\,\mathrm{d}W_t,\quad X_{s,s}^{(n)} = X_{s,s}, t\geq s, \tag{2.6}$$

where $\mu_{s,t}^{(n-1)} = \mathscr{L}_{X_{s,t}^{(n-1)}}$.

Lemma 2.3. *Assume* ($H1$)-($H3$) *for some* $\theta\in[1,\infty)$.

(1) *For every* $n\geq 1$, *the SDE* (2.6) *has a unique strong solution and*

$$\mathbb{E}\sup_{t\in[s,T]}|X_{s,t}^{(n)}|^\theta<\infty,\quad T>s, n\geq 1. \tag{2.7}$$

(2) *If either* $\theta\geq 2$ *or* $\sigma(x,\mu)$ *does not depend on* μ, *then for any* $T>0$ *there exists* $t_0>0$ *which is independent on* $s\in[0,T]$ *and* $X_{s,s}$, *such that*

$$\mathbb{E}\sup_{t\in[s,s+t_0]}|X_{s,t}^{(n+1)}-X_{s,t}^{(n)}|^\theta \leq 2^\theta \mathrm{e}^{-n}\mathbb{E}\sup_{t\in[s,s+t_0]}|X_{s,t}^{(1)}|^2,\quad s\in[0,T], n\geq 1. \tag{2.8}$$

Proof. Without loss of generality, we only prove for $s=0$.

(1) We first prove that the SDE (2.6) has a unique strong solution and (2.7) holds. By ($H1$), $b_t(x,\mu_t^{(0)})$ and $\sigma_t(x,\mu_t^{(0)})$ are continuous in x. Then the SDE (2.6) for $n=1$ has a weak solution up to life time (see [16, Theorem 6.1.6] and [11, p.155-163]). Next, by Itô's formula it is easy to see that ($H2$) implies the pathwise uniqueness. According to the Yamada-Watanabe principle [25], the SDE has a unique solution up to life time. It remains to prove (2.7). By ($H3$) and Itô's formula we have

$$\begin{aligned} \mathrm{d}|X_t^{(1)}|^2 &= 2\langle\sigma_t(X_t^{(1)},\mu_t^{(0)})\mathrm{d}W_t, X_t^{(1)}\rangle \\ &\quad + \big\{2\big\langle b_t(X_t^{(1)},\mu_t^{(0)}), X_t^{(1)}\big\rangle + \|\sigma_t(X_t^{(1)},\mu_t^{(0)})\|_{HS}^2\big\}\mathrm{d}t. \end{aligned} \tag{2.9}$$

By $(H1)$ with $y=0, \nu=\delta_0$, we have

$$\|\sigma_t(x,\mu)\|_{HS}^2 \le K(t)\big\{1+|x|^2+\mu(|\cdot|^\theta)^{\frac{2}{\theta}}\big\} \tag{2.10}$$

for some increasing $K:[0,\infty)\to[0,\infty)$. Combining this with $(H2)$ and $(H3)$, we may find increasing $H:[0,\infty)\to(0,\infty)$ such that

$$\begin{aligned}
&\max\big\{2\big\langle b_t(x,\mu_t^{(0)}),x\big\rangle,\ \|\sigma_t(x,\mu_t^{(0)})\|_{HS}^2\big\}\\
&\le \max\{2\big\langle b_t(x,\mu_t^{(0)})-b_t(0,\mu_t^{(0)}),x\big\rangle+2|b_t(0,\mu_t^{(0)})|\cdot|x|,\ \|\sigma_t(x,\mu_t^{(0)})\|_{HS}^2\big\}\\
&\le H(t)\big\{1+|x|^2+\mu_t^{(0)}(|\cdot|^\theta)^{\frac{2}{\theta}}\big\},\quad t\ge 0, x\in\mathbb{R}^d.
\end{aligned}$$

Thus, by (2.9), (2.10) and using Itô's formula, there exists a constant $c_1(\theta)>0$ such that

$$\begin{aligned}
\mathrm{d}(1+|X_t^{(1)}|^2)^{\frac{\theta}{2}} \le&\theta(1+|X_t^{(1)}|^2)^{\frac{\theta-2}{2}}\langle\sigma_t(X_t^{(1)},\mu_t^{(0)})\mathrm{d}W_t, X_t^{(1)}\rangle\\
&+c_1(\theta)H(t)\Big\{(1+|X_t^{(1)}|^2)^{\frac{\theta}{2}}+\mu_t^{(0)}(|\cdot|^\theta)^{\frac{2}{\theta}\vee 1}\Big\}\mathrm{d}t.
\end{aligned}$$

Letting $\tau_N=\inf\{t\ge 0: |X_t^{(1)}|\ge N\}$, we conclude from this, (2.10) and the BDG inequality yield that for some increasing $\Psi:[0,\infty)\to(0,\infty)$,

$$\begin{aligned}
&\mathbb{E}\sup_{s\in[0,t\wedge\tau_N]}\big(1+|X_s^{(1)}|^2\big)^{\frac{\theta}{2}} \le c_1(\theta)H(t)\mathbb{E}\int_0^{t\wedge\tau_N}\Big\{\big(1+|X_s^{(1)}|^2\big)^{\frac{\theta}{2}}+\mu_t^{(0)}(|\cdot|^\theta)^{\frac{2}{\theta}\vee 1}\Big\}\mathrm{d}s\\
&\quad+c_2(\theta)K(t)\mathbb{E}\bigg(\int_0^{t\wedge\tau_N}\Big\{\big(1+|X_s^{(1)}|^2\big)^{\theta}+\big(1+|X_s^{(1)}|^2\big)^{\theta-\frac{1}{2}}\mu_s^{(0)}(|\cdot|^\theta)^{\frac{2}{\theta}\vee 1}\Big\}\mathrm{d}s\bigg)^{\frac{1}{2}}\\
&\le \Psi(t)\mathbb{E}\int_0^{t\wedge\tau_N}\Big\{\big(1+|X_s^{(1)}|^2\big)^{\frac{\theta}{2}}+\mu_s^{(0)}(|\cdot|^\theta)^{\frac{2}{\theta}\vee 1}\Big\}\mathrm{d}s+\frac{1}{2}\mathbb{E}\sup_{s\in[0,t\wedge\tau_N]}\big(1+|X_s^{(1)}|^2\big)^{\frac{\theta}{2}}.
\end{aligned}$$

Therefore,

$$\mathbb{E}\sup_{s\in[0,t\wedge\tau_N]}\big(1+|X_s^{(1)}|^2\big)^{\frac{\theta}{2}} \le 2\Psi(t)\int_0^t\Big\{\big(1+|X_{s\wedge\tau_N}^{(1)}|^2\big)^{\frac{\theta}{2}}+\mu_s^{(0)}(|\cdot|^\theta)^{\frac{2}{\theta}\vee 1}\Big\}\mathrm{d}s.$$

By Gronwall's lemma and letting $N\to\infty$, we arrive at

$$\mathbb{E}\sup_{s\in[0,t]}\big(1+|X_s^{(1)}|^2\big)^{\frac{\theta}{2}} \le \Big(1+2t\Psi(t)\sup_{s\in[0,t]}\big(\mathbb{E}|X_s^{(0)}|^\theta\big)^{\frac{2}{\theta}\vee 1}\Big)\exp\big[2t\Psi(t)\big]<\infty.$$

Therefore, (2.7) holds for $n=1$.

Now, assume that the assertion holds for $n=k$ for some $k\ge 1$, we intend to prove it for $n=k+1$. This can be done in the same way by using $(X_\cdot^{(k+1)},\mu_\cdot^{(k)},X_\cdot^{(k)})$ in place of $(X_\cdot^{(1)},\mu_\cdot^{(0)},X_\cdot^{(0)})$. So, we omit the proof to save space.

(2) To prove (2.8), for $n\ge 1$ we simply denote

$$\xi_t^{(n)}=X_t^{(n+1)}-X_t^{(n)},$$

$$\Lambda_t^{(n)} = \sigma_t(X_t^{(n+1)}, \mu_t^{(n)}) - \sigma_t(X_t^{(n)}, \mu_t^{(n-1)}),$$
$$B_t^{(n)} = b_t(X_t^{(n+1)}, \mu_t^{(n)}) - b_t(X_t^{(n)}, \mu_t^{(n-1)}).$$

Below we prove for 1) $\theta \geq 2$ and 2) $\theta < 2$ but $K_{\sigma,2} = 0$ respectively.

Let $\theta \geq 2$. By $(H1)$, $(H2)$ and Itô's formula, there exists increasing $K_0 : [0,\infty) \to [0,\infty)$ such that

$$\text{(2.11)} \qquad \mathrm{d}|\xi_t^{(n)}|^2 \leq 2\langle \Lambda_t^{(n)} \mathrm{d}W_t, \xi_t^{(n)}\rangle + K_0(t)\big\{|\xi_t^{(n)}|^2 + \mathbb{W}_\theta(\mu_t^{(n)}, \mu_t^{(n-1)})^2\big\}\mathrm{d}t.$$

Combining this with $(H1)$ and using the BDG inequality, we may find out increasing functions $K_1, K_2 : [0,\infty) \to (0,\infty)$ such that

$$\begin{aligned}
&\mathbb{E}\sup_{s\in[0,t]} |\xi_s^{(n)}|^\theta \leq 2^{\frac{\theta}{2}-1}\bigg|\sup_{s\in[0,t]}\int_0^s 2\langle \Lambda_t^{(n)}\mathrm{d}W_t, \xi_t^{(n)}\rangle\bigg|^{\frac{\theta}{2}} \\
&\quad 2^{\frac{\theta}{2}-1}K_0(t)\bigg(\int_0^t \big\{|\xi_s^{(n)}|^2 + \mathbb{W}_\theta(\mu_s^{(n)}, \mu_s^{(n-1)})^2\big\}\mathrm{d}s\bigg)^{\frac{\theta}{2}} \\
&\leq K_1(t)\mathbb{E}\bigg(\int_0^t \big\{|\xi_s^{(n)}|^2\|\Lambda_s^{(n)}\|^2\big\}\mathrm{d}s\bigg)^{\frac{\theta}{2}} + K_1(t)\int_0^t \Big\{\mathbb{E}|\xi_s^{(n)}|^\theta + \mathbb{W}_\theta(\mu_s^{(n)}, \mu_s^{(n-1)})^\theta\Big\}\mathrm{d}s \\
&\leq \frac{1}{2}\mathbb{E}\sup_{s\in[0,t]}|\xi_s^{(n)}|^\theta + K_2(t)\int_0^t \Big\{\mathbb{E}|\xi_s^{(n)}|^\theta + \mathbb{W}_\theta(\mu_s^{(n)}, \mu_s^{(n-1)})^\theta\Big\}\mathrm{d}s.
\end{aligned}$$

Then

$$\mathbb{E}\sup_{s\in[0,t]}|\xi_s^{(n)}|^2 \leq 2K_2(t)\int_0^t \Big\{\mathbb{E}|\xi_s^{(n)}|^\theta + \mathbb{W}_\theta(\mu_s^{(n)}, \mu_s^{(n-1)})^\theta\Big\}\mathrm{d}s, \quad t \geq 0.$$

By Gronwall's lemma, we obtain

$$\text{(2.12)} \qquad \begin{aligned}
\mathbb{E}\sup_{s\in[0,t]}|\xi_s^{(n)}|^\theta &\leq 2tK_2(t)\mathrm{e}^{2tK_2(t)}\sup_{s\in[0,t]}\mathbb{W}_\theta(\mu_s^{(n)}, \mu_s^{(n-1)})^\theta \\
&\leq 2tK_2(t)\mathrm{e}^{2tK_2(t)}\mathbb{E}\sup_{s\in[0,t]}|\xi_s^{(n-1)}|^\theta, \quad t \geq 0.
\end{aligned}$$

Taking $t_0 > 0$ such that $2t_0K_2(t_0)\mathrm{e}^{2t_0K_2(t_0)} \leq \mathrm{e}^{-1}$, we arrive at

$$\mathbb{E}\sup_{s\in[0,t_0]}|\xi_s^{(n)}|^\theta \leq \mathrm{e}^{-1}\mathbb{E}\sup_{s\in[0,t_0]}|\xi_s^{(n-1)}|^\theta, \quad n \geq 1.$$

Since

$$\mathbb{E}\sup_{s\in[0,t_0]}|\xi_s^{(0)}|^\theta \leq 2^{\theta-1}\mathbb{E}\Big\{|X_0|^\theta + \sup_{s\in[0,t_0]}|X_s^{(1)}|^\theta\Big\} \leq 2^\theta\mathbb{E}\sup_{s\in[0,t_0]}|X_s^{(1)}|^2,$$

we prove (2.8).

Let $\theta \in [1,2)$ but $K_{\sigma,2} = 0$. Then instead of (2.11) we have

$$\mathrm{d}|\xi_t^{(n)}|^2 \leq 2\langle \Lambda_t^{(n)}\mathrm{d}W_t, \xi_t^{(n)}\rangle + K_0(t)|\xi_t^{(n)}|\big\{|\xi_t^{(n)}| + \mathbb{W}_\theta(\mu_t^{(n)}, \mu_t^{(n-1)})\big\}\mathrm{d}t.$$

Since $\theta \le 2$, for any $\varepsilon > 0$, by Itô's formula we obtain

$$\mathrm{d}(\varepsilon+|X_t^{(n)}|^2)^{\frac{\theta}{2}} \le \theta(\varepsilon+|X_t^{(n)}|^2)^{\frac{\theta-2}{2}}\Big\{\langle\Lambda_t^{(n)}\mathrm{d}W_t, \xi_t^{(n)}\rangle + \frac{K_0(t)}{2}|\xi_t^{(n)}|\big\{|\xi_t^{(n)}|+\mathbb{W}_\theta(\mu_t^{(n)}, \mu_t^{(n-1)})\big\}\mathrm{d}t\Big\}.$$

Since $(H1)$ with $K_{\sigma,2} = 0$ implies $\|\Lambda_t^{(n)}\|^2 \le K_{\sigma,1}(t)|\xi_t^{(n)}|^2$, this and the BDG inequality yield

$$\begin{aligned}
\mathbb{E}\sup_{s\in[0,t]}(\varepsilon+|X_s^{(n)}|^2)^{\frac{\theta}{2}} &\le K_1(t)\mathbb{E}\bigg(\int_0^t(\varepsilon+|X_s^{(n)}|^2)^\theta \mathrm{d}s\bigg)^{\frac{1}{2}} \\
&\quad + K_1(t)\mathbb{E}\int_0^t\Big\{(\varepsilon+|X_s^{(n)}|^2)^{\frac{\theta}{2}} + (\varepsilon+|X_s^{(n)}|^2)^{\frac{\theta-1}{2}}W_\theta(\mu_s^{(n)}, \mu_s^{(n-1)})\Big\}\mathrm{d}s \\
&\le \frac{1}{2}\mathbb{E}\sup_{s\in[0,t]}(\varepsilon+|X_s^{(n)}|^2)^{\frac{\theta}{2}} + K_2(t)\int_0^t\Big\{(\varepsilon+|X_s^{(n)}|^2)^{\frac{\theta}{2}} + W_\theta(\mu_s^{(n)}, \mu_s^{(n-1)})^\theta\Big\}\mathrm{d}s
\end{aligned}$$

for some increasing $K_1, K_2 : [0,\infty) \to [0,\infty)$. Letting $\varepsilon \to 0$ and using Gronwall's inequality, we prove (2.12), which implies the desired estimate (2.8) as explained above. □

2.3 Proofs of Theorem 2.1 and Theorem 2.2

Proof of Theorem 2.1. Without loss of generality, we only consider the DDSDE (1.6) from time $s = 0$.

(1) Since the uniqueness follows from (2.3) which will be proved in the next step, in this step we only prove the existence and the estimate (2.1).

By Lemma 2.3, there exists an adapted continuous process $(X_t)_{t\in[0,t_0]}$ such that

$$\lim_{n\to\infty}\sup_{t\in[0,t_0]}\mathbb{W}_\theta(\mu_t^{(n)}, \mu_t)^\theta \le \lim_{n\to\infty}\mathbb{E}\sup_{t\in[0,t_0]}|X_t^{(n)} - X_t|^\theta = 0, \tag{2.13}$$

where μ_t is the distribution of X_t. Noting that due to (2.6)

$$X_t^{(n)} = X_0 + \int_0^t b_s(X_s^{(n)}, \mu_s^{(n-1)})\mathrm{d}s + \int_0^t \sigma_s(X_s^{(n)}, \mu_s^{(n-1)})\mathrm{d}W_s,$$

it follows from (2.13), $(H1)$ and $(H3)$ that $\mathbb{P}$-a.s.

$$X_t = X_0 + \int_0^t b_s(X_s, \mu_s)\mathrm{d}s + \int_0^t \sigma_s(X_s, \mu_s)\mathrm{d}W_s, \quad t \in [0, t_0].$$

Therefore, $(X_t)_{t\in[0,t_0]}$ is a solution to (1.6), and (2.13) implies $\mathbb{E}\sup_{s\in[0,t_0]}|X_s|^\theta < \infty$. Since $t_0 > 0$ is independent of X_0, we conclude that (1.6) has a unique solution $(X_t)_{t\ge0}$ with

$$\mathbb{E}\sup_{s\in[0,t]}|X_s|^\theta < \infty, \quad t \in (0,\infty). \tag{2.14}$$

It remains to prove (2.1) for $\mathbb{E}|X_0|^p < \infty$. As in the proof of (2.7) above, by $(H1)$-$(H3)$ and Itô's formula we have

$$\mathrm{d}|X_t|^2 \le 2\langle\sigma_t(X_t, \mathscr{L}_{X_t})\mathrm{d}W_t, X_t\rangle + H(t)\big\{1+|X_t|^2+(\mathbb{E}|X_t|^\theta)^{\frac{2}{\theta}}\big\}\mathrm{d}t$$

for some increasing function $H : [0,\infty) \to (0,\infty)$. Then applying Itô's formula to $(1+|X_t|^2)^{\frac{p}{2}}$ and repeating step (1) in the proof of Lemma 2.3, we prove (2.1).

(2) By Itô's formula, $(H2)$ and $(H1)$ with $K_{\sigma,2} = 0$ if $\theta < 2$, we have

$$\begin{aligned} \text{d}|X_t - Y_t|^2 \le\ & 2\big\langle X_t - Y_t, \{\sigma_t(X_t, \mathscr{L}_{X_t}) - \sigma_t(Y_t, \mathscr{L}_{Y_t})\}\text{d}W_t\big\rangle \\ & + K_1(t)\big\{|X_t - Y_t|^2 + 1_{\{\theta\ge 2\}}\mathbb{W}_\theta(\mathscr{L}_{X_t}, \mathscr{L}_{Y_t})^2 + |X_t - Y_t|W_\theta(\mathscr{L}_{X_t}, \mathscr{L}_{Y_t})\big\}\text{d}t. \end{aligned} \tag{2.15}$$

By Itô's formula we obtain

$$\begin{aligned} \text{d}|X_t - Y_t|^\theta \le\ & \theta|X_t - Y_t|^{\theta-2}\big\langle X_t - Y_t, \{\sigma_t(X_t, \mathscr{L}_{X_t}) - \sigma_t(Y_t, \mathscr{L}_{Y_t})\}\text{d}W_t\big\rangle \\ & + K_2(t)\big\{|X_t - Y_t|^\theta + \mathbb{W}_\theta(\mathscr{L}_{X_t}, \mathscr{L}_{Y_t})^\theta\big\}\text{d}t \end{aligned}$$

for some increasing $K_2 : [0,\infty) \to [0,\infty)$. Noting that $\mathbb{W}_\theta(\mathscr{L}_{X_t}, \mathscr{L}_{Y_t})^\theta \le \mathbb{E}|X_t - Y_t|^\theta < \infty$, this implies

$$\mathbb{E}|X_t - Y_t|^\theta \le \mathbb{E}|X_0 - Y_0|^\theta + 2\int_0^t K_2(s)\mathbb{E}|X_s - Y_s|^2\text{d}s.$$

By Gronwall's lemma, we prove (2.2).

To prove (2.3), let $\tau_\varepsilon := \inf\{t \ge 0 : |X_t - Y_t| \ge \varepsilon\}$ for $\varepsilon \in (0,1)$. By (2.2) and (2.15), there exists increasing $K : [0,\infty) \to [0,\infty)$ such that

$$\begin{aligned} & \mathbb{E}|X_{t\wedge\tau_\varepsilon} - Y_{t\wedge\tau_\varepsilon}|^\theta \\ & \le \mathbb{E}|X_0 - Y_0|^\theta + \mathbb{E}\int_0^{t\wedge\tau_\varepsilon} K(s)\big(\mathbb{E}|X_s - Y_s|^\theta + |X_s - Y_s|^\theta\big)\text{d}s \\ & \le \big\{1 + t\text{e}^{t\psi(t)}\big\}\mathbb{E}|X_0 - Y_0|^2 + \int_0^t K(s)\big\{\mathbb{E}|X_{s\wedge\tau_\varepsilon} - Y_{s\wedge\tau_\varepsilon}|^\theta\big\}\text{d}s, \quad t \ge 0. \end{aligned}$$

By Gronwall's lemma, there exists positive $\phi \in C([0,\infty))$ such that

$$\varepsilon^\theta\mathbb{P}(\tau_\varepsilon \le t) \le \mathbb{E}|X_{t\wedge\tau_\varepsilon} - Y_{t\wedge\tau_\varepsilon}|^\theta \le \phi(t)\mathbb{E}|X_0 - Y_0|^\theta, \quad t \ge 0.$$

Therefore,

$$\mathbb{P}\Big(\sup_{s\in[0,t]}|X_s - Y_s| \ge \varepsilon\Big) = \mathbb{P}(\tau_\varepsilon \le t) \le \varepsilon^{-\theta}\psi(t)\mathbb{E}|X_0 - Y_0|^\theta, \quad t, \varepsilon > 0.$$

Hence, (2.3) holds.

(3) Let (X_t, W_t) and $(\tilde X_t, \tilde W_t)$ with respect to $(\Omega, \{\mathscr{F}_t\}_{t\ge0}, \mathbb{P})$ and $(\tilde\Omega, \{\tilde{\mathscr{F}}_t\}_{t\ge0}, \tilde{\mathbb{P}})$ respectively be two weak solutions such that $\mathscr{L}_{X_0}|_{\mathbb{P}} = \mathscr{L}_{\tilde X_0}|_{\tilde{\mathbb{P}}}$. Then X_t solves (1.6) while $\tilde X_t$ solves

$$\text{d}\tilde X_t = b_t(\tilde X_t, \mathscr{L}_{\tilde X_t}|_{\tilde{\mathbb{P}}})\text{d}t + \sigma_t(\tilde X_t, \mathscr{L}_{\tilde X_t}|_{\tilde{\mathbb{P}}})\text{d}\tilde W_t. \tag{2.16}$$

To prove that $\mathscr{L}_X|_{\mathbb{P}} = \mathscr{L}_{\tilde X}|_{\tilde{\mathbb{P}}}$, let $\mu_t = \mathscr{L}_{X_t}|_{\mathbb{P}}$ and

$$\bar b_t(x) = b_t(x, \mu_t), \quad \bar\sigma_t(x) = \sigma(x, \mu_t), \quad t \ge 0, x \in \mathbb{R}^d.$$

By $(H1)$-$(H3)$, the SDE

$$\mathrm{d}\bar{X}_t = \bar{b}_t(\bar{X}_t)\mathrm{d}t + \bar{\sigma}_t(\bar{X}_t)\mathrm{d}\tilde{W}_t, \quad \bar{X}_0 = \tilde{X}_0 \tag{2.17}$$

has a unique solution for any initial points. According to Yamada-Watanabe, it also has weak uniqueness. Noting that

$$\mathrm{d}X_t = \bar{b}_t(X_t)\mathrm{d}t + \bar{\sigma}_t(X_t)\mathrm{d}W_t, \quad \mathscr{L}_{X_0}|_{\mathbb{P}} = \mathscr{L}_{\tilde{X}_0}|_{\tilde{\mathbb{P}}},$$

the weak uniqueness of (2.17) implies

$$\mathscr{L}_{\bar{X}}|_{\tilde{\mathbb{P}}} = \mathscr{L}_X|_{\mathbb{P}}. \tag{2.18}$$

So, (2.17) reduces to

$$\mathrm{d}\bar{X}_t = b_t(\bar{X}_t, \mathscr{L}_{\bar{X}_t}|_{\tilde{\mathbb{P}}})\mathrm{d}t + \sigma_t(\bar{X}_t, \mathscr{L}_{\bar{X}_t}|_{\tilde{\mathbb{P}}})\mathrm{d}\tilde{W}_t, \quad \bar{X}_0 = \tilde{X}_0.$$

Since by (1) the DDSDE (2.16) has a unique solution, we obtain $\bar{X} = \tilde{X}$. Therefore, the weak uniqueness follows from (2.18).

Finally, for any $\mu_0, \nu_0 \in \mathscr{P}_\theta$, take $\mathscr{F}_0$-measurable random variables X_0, Y_0 such that $\mathscr{L}_{X_0} = \mu_0, \mathscr{L}_{Y_0} = \nu_0$ and $\mathbb{W}_\theta(\mu_0, \nu_0)^\theta = \mathbb{E}|X_0 - Y_0|^\theta$. Since $\mathbb{W}_\theta(P_t^*\mu_0, P_t^*\nu_0)^\theta \le \mathbb{E}|X_t - Y_t|^\theta$, (2.2) implies (2.4).

□

Proof of Theorem 2.2. Since the assumptions are weaker for larger θ, we may and do assume that $\theta \ge 2$. By Kolmogorov's modification theorem, it suffices to prove

$$\mathbb{E}|X_t(x) - X_s(y)|^m \le \Phi(s,t;x,y)(|x-y| + |s-t|)^q, \quad |t-s| + |x-y| \le 1 \tag{2.19}$$

for some constants $m > 0, q > 1$ and locally bounded function Φ on $[0,\infty)^2 \times \mathbb{R}^{2d}$. Firstly, by (2.1) and (2.2), we may find out an increasing function $\psi : [0,\infty) \to (0,\infty)$ such that

$$\begin{aligned} \mathbb{E}|X_t(x) - X_t(y)|^{2\theta} &\le \big\{\mathbb{E}|X_t(x) - X_t(y)|^\theta\big\}^{\frac{2}{3}} \times \big\{\mathbb{E}|X_t(x) - X_t(y)|^{4\theta}\big\}^{\frac{1}{3}} \\ &\le \psi(t)(1 + |x| + |y|)^{\frac{4\theta}{3}}|x-y|^{\frac{2\theta}{3}}, \quad t \ge 0, x, y \in \mathbb{R}^d. \end{aligned} \tag{2.20}$$

Next, by $(H3)$ and (2.5), there exist a constant $C > 0$ and an increasing function $\phi : [0,\infty) \to (0,\infty)$ such that

$$\begin{aligned} \mathbb{E}|X_t(x) - X_s(x)|^{2\theta} &\le C\mathbb{E}\bigg(\int_s^t K(r)(1 + |X_r(x)|^p + \mathbb{E}|X_r(x)|^p)\mathrm{d}r\bigg)^{2\theta} \\ &\quad + C\mathbb{E}\bigg(\int_s^t K_2(r)(1 + |X_r(x)|^2 + \mathbb{E}|X_r(x)|^2)\mathrm{d}r\bigg)^\theta \\ &\le \phi(t)(1 + |x|^{2\theta p})(t-s)^\theta, \quad |t-s| \le 1, x \in \mathbb{R}^d. \end{aligned}$$

This together with (2.20) implies the desired (2.19) with $p = 4, q = \frac{2\theta}{3} > 1$. □

3 $\mathbb{W}_2$-Exponential contraction of $P^*_{s,t}$

We intend to estimate the Wasserstein distance of solutions with different initial distributions and investigate the exponential ergodicity. For simplicity, we only consider the $\mathbb{W}_2$-distance. To this end, we use the following condition to replace $(H2)$:

$(H2')$ There exist positive functions $C_2, C_2 \in L^1_{loc}(\mathrm{d}t)$ such that

$$\begin{aligned}&2\langle b_t(x,\mu) - b_t(y,\nu), x-y\rangle + \|\sigma_t(x,\mu) - \sigma_t(y,\nu)\|^2_{HS}\\ &\le C_1(t)\mathbb{W}_2(\mu,\nu)^2 - C_2(t)|x-y|^2, \quad t\ge 0; x,y\in\mathbb{R}^d; \mu,\nu\in\mathscr{P}_2.\end{aligned}$$

Theorem 3.1. *Assume* $(H1)$, $(H2')$ *and* $(H3)$.

(1) *For any* $\mu_0,\nu_0\in\mathscr{P}_2$,

$$\mathbb{W}_2(P^*_{s,t}\mu_0, P^*_{s,t}\nu_0)^2 \le \mathbb{W}_2(\mu_0,\nu_0)^2 \mathrm{e}^{\int_s^t \{C_1(r)-C_2(r)\}\mathrm{d}r}, \quad t\ge 0.$$

(2) *Let* $b_t = b$ *and* $\sigma_t=\sigma$ *do not depend on time* t *such that* $(H2')$ *holds for some constants* C_1 *and* C_2. *If* $C_2 > C_1$ *then* P_t *has a unique invariant probability measure* $\mu\in\mathscr{P}_2$ *such that*

$$\mathbb{W}_2(P^*_t\nu_0,\mu)^2 \le \mathbb{W}_2(\nu_0,\mu)^2\mathrm{e}^{-(C_2-C_1)t}, \quad t\ge 0, \nu_0\in\mathscr{P}_2.$$

Proof. (1) Without loss of generality, we only prove for $s=0$. Let X_t and Y_t be two solutions to (1.6) such that $\mathscr{L}_{X_0}=\mu_0, \mathscr{L}_{Y_0}=\nu_0$ and

$$\mathbb{W}_2(\mu_0,\nu_0)^2 = \mathbb{E}|X_0-Y_0|^2. \tag{3.1}$$

Simply denote $\mu_t = \mathscr{L}_{X_t}, \nu_t = \mathscr{L}_{Y_t}, t\ge 0$. By $(H2')$ and Itô's formula we have

$$\begin{aligned}\mathrm{d}|X_t-Y_t|^2 &\le 2\big\langle X_t - Y_t, \{\sigma_t(X_t,\mu_t) - \sigma_t(Y_t,\nu_t)\}\mathrm{d}W_t\big\rangle\\ &\quad + \big\{C_1(t)\mathbb{W}_2(\mu_t,\nu_t)^2 - C_2(t)|X_t-Y_t|^2\big\}\mathrm{d}t.\end{aligned}$$

Noting that $\mathbb{W}_2(\mu_s,\nu_s)^2\le\mathbb{E}|X_s-Y_s|^2$, combining this with (3.1) we obtain

$$\mathbb{E}|X_t-Y_t|^2 \le \mathbb{W}_2(\mu_0,\nu_0)^2 + \int_0^t \big\{[C_1(s)-C_2(s)]\mathbb{E}|X_s-Y_s|^2\big\}\mathrm{d}s.$$

This implies the first assertion by Gronwall's lemma.

(2) Let δ_0 be the Dirac measure at point $0\in\mathbb{R}^d$. Then $P^*_t\delta_0 = \mathscr{L}_{X_t(0)}$. We first prove

$$\lim_{t\to\infty}\mathbb{W}_2(P^*_t\delta_0,\mu) = 0 \tag{3.2}$$

for some $\mu\in\mathscr{P}_2$. To this end, it suffices to show that $\{P^*_t\delta_0\}_{t\ge0}$ is a $\mathbb{W}_2$-Cauchy family when $t\to\infty$; that is,

$$\lim_{t\to\infty}\sup_{s\ge0}\mathbb{W}_2(P^*_t\delta_0, P^*_{t+s}\delta_0) = 0. \tag{3.3}$$

We will prove this using the shift-coupling and the weak uniqueness according to Theorem 2.1(3). More precisely, for any $s \geq 0$, it is easy to see that $(\bar{X}_t := X_{t+s}(0))_{t\geq 0}$ solves the DDSDE

$$\mathrm{d}\bar{X}_t = b(\bar{X}_t, \mathscr{L}_{\bar{X}_t})\mathrm{d}t + \sigma(\bar{X}_t, \mathscr{L}_{\bar{X}_t})\mathrm{d}\bar{W}_t, \quad \bar{X}_0 = X_s(0)$$

for the d-dimensional Brownian motion $\bar{W}_t := W_{t+s} - W_s$. So, by the weak uniqueness we have

$$P_t^*(P_s^*\delta_0) = \mathscr{L}_{\bar{X}_t} = \mathscr{L}_{X_{t+s}(0)} = P_{t+s}^*\delta_0, \quad s,t \geq 0. \tag{3.4}$$

Combining this with Theorem 3.1(1) and letting $X_t(P_s^*\delta_0)$ solve (1.6) with $\mathscr{L}_{X_0} = P_s^*\delta_0$, we obtain

$$\begin{aligned}
&\mathbb{W}_2(P_{t+s}^*\delta_0, P_t^*\delta_0)^2 = \mathbb{W}_2(\mathscr{L}_{X_t(P_s^*\delta_0)}, \mathscr{L}_{X_t(0)})^2 \\
&\leq \mathbb{W}_2(P_s^*\delta_0, \delta_0)^2 \mathrm{e}^{-(C_2-C_1)t} = \mathrm{e}^{-(C_2-C_1)t}\mathbb{E}|X_s(0)|^2, \quad s,t \geq 0.
\end{aligned}$$

This implies (3.3) provided

$$\sup_{s\geq 0} E|X_s(0)|^2 < \infty. \tag{3.5}$$

By $(H2')$ and $(H3)$ for constant $C_1 < C_2$ and K_2, it is easy to see that

$$2\langle b(x,\mu), x\rangle + \|\sigma(x,\mu)\|_{HS}^2 \leq C_0 - (C_2 - \varepsilon)|x|^2 + (C_1 + \varepsilon)\mu(|\cdot|^2)$$

holds for some constant $C_0 > 0$ and $\varepsilon := \frac{C_2 - C_1}{4} > 0$. By Itô's formula and Gronwall's lemma, this implies

$$\mathbb{E}|X_t(0)|^2 \leq C_0 \mathrm{e}^{-(C_2-C_1-2\varepsilon)t}, \quad t \geq 0.$$

Therefore, (3.5) holds.

Moreover, by (2.4) and (3.2) we have

$$\lim_{t\to\infty} \mathbb{W}_2(P_s^*\mu, P_s^*P_t^*\delta_0) = 0, \quad s \geq 0.$$

Combining this with (3.2) and (3.4), we obtain

$$\mathbb{W}_2(P_s^*\mu, \mu) \leq \lim_{t\to\infty} \mathbb{W}_2(P_s^*P_t^*\delta_0, P_t^*\delta_0) = \lim_{t\to\infty} \mathbb{W}_2(P_{t+s}^*\delta_0, P_t^*\delta_0) = 0.$$

Then μ is an invariant probability measure. Therefore, by Theorem 3.1(1) with $C_2 > C_1$, for any $\nu_0 \in \mathscr{P}_2$ we have

$$\mathbb{W}_2(P_t^*\nu_0, \mu)^2 = \mathbb{W}_2(P_t^*\nu_0, P_t^*\mu)^2 \leq \mathrm{e}^{-(C_2-C_1)t}\mathbb{W}_2(\nu_0, \mu)^2, \quad t \geq 0,$$

so that the proof is finished. □

4 Harnack inequality and applications

In this section, we investigate the dimension-free Harnack inequality in the sense of [20] and the log-Harnack inequality introduced in [15, 22] for the DDSDE (1.6), see [21] and references within for general results on these type Harnack inequalities and applications. To establish Harnack inequalities for DDSDEs using coupling by change of measures, we need to assume that the noise part is distribution-free; that is, we consider the following special version of (1.6):

$$\mathrm{d}X_t = b_t(X_t, \mathscr{L}_{X_t})\mathrm{d}t + \sigma_t(X_t)\mathrm{d}W_t. \tag{4.1}$$

Then

$$(P_t f)(\mu_0) = \int_{\mathbb{R}^d} f \mathrm{d}(P_t^* \mu_0) = \mathbb{E}f(X_t(\mu_0)), \quad f \in \mathscr{B}_b(\mathbb{R}^d), t \geq 0, \mu_0 \in \mathscr{P}_2,$$

where $X_t(\mu_0)$ solves (4.1) with initial distribution μ_0.

To make the study easy to follow, we first introduce the main steps in establishing Harnack inequalities using coupling by change of measures summarized in [21, §1.1].

(S1) Let $(X_t)_{t\geq 0}$ solve (4.1) with $\mathscr{L}_{X_0} = \mu_0$. By the uniqueness we have $\mu_t := \mathbb{P}_t^* \mu_0 = \mathscr{L}_{X_t}$, and the equation (4.1) reduces to

$$\mathrm{d}X_t = b_t(X_t, \mu_t)\mathrm{d}t + \sigma_t(X_t)\mathrm{d}W_t. \tag{4.2}$$

(S2) Construct a process $(Y_t)_{t\in[0,T]}$ such that for a weighted probability measure $\mathbb{Q} := R_T \mathbb{P}$,

$$X_T = Y_T \ \mathbb{Q}\text{-a.s.}, \quad \text{and } \mathscr{L}_{Y_T}|_{\mathbb{Q}} = P_T^* \nu_0 =: \nu_T. \tag{4.3}$$

Obviously, (S1) and (S2) implies

$$(P_T f)(\mu_0) = \mathbb{E}[f(X_T)] \text{ and } (P_T f)(\nu_0) = \mathbb{E}_{\mathbb{Q}}[f(Y_T)] = \mathbb{E}[R_T f(X_T)], \quad f \in \mathscr{B}_b(\mathbb{R}^d). \tag{4.4}$$

Combining this with (4.4) Young's inequality (see [1, Lemma 2.4]), we obtain the log-Harnack inequality:

$$\begin{aligned}(P_T \log f)(\nu_0) &\leq \mathbb{E}[R_T \log R_T] + \log \mathbb{E}[f(X_T)] \\ &= \log(P_T f)(\mu_0) + \mathbb{E}[R_T \log R_T], \quad f \in \mathscr{B}_b^+(\mathbb{R}^d);\end{aligned} \tag{4.5}$$

while using Hölder's inequality we prove the Harnack inequality with power $p > 1$:

$$(P_T f(\nu_0))^p = (\mathbb{E}[R_T f(X_T)])^p \leq (\mathbb{E}R_T^{\frac{p}{p-1}})^{p-1}(P_T f^p)(\mu_0), \quad f \in \mathscr{B}_b^+(\mathbb{R}^d). \tag{4.6}$$

To construct Y_t in (S2), we will need the following assumption.

(A) $\sigma_t(x)$ is invertible and locally Lipschitzian in x which is locally uniformly in $t \geq 0$, and there exist increasing functions $\kappa_0, \kappa_1, \kappa_2, \lambda : [0,\infty) \to (0,\infty)$ such that for any $t \in [0,T], x, y \in \mathbb{R}^d$ and $\mu, \nu \in \mathscr{P}_2$, we have

$$\|\sigma_t^{-1}\|_\infty \leq \lambda(t), \quad |b_t(0,\mu)|^2 + \|\sigma_t(x)\|^2 \leq \kappa_0(t)(1 + |x|^2 + \mu(|\cdot|^2)), \tag{4.7}$$

$$\begin{aligned}&2\langle b_t(x,\mu) - b_t(y,\nu), x - y\rangle + \|\sigma_t(x) - \sigma_t(y)\|_{HS}^2 \\ &\leq \kappa_1(t)|x-y|^2 + \kappa_2(t)|x-y|\mathbb{W}_2(\mu,\nu).\end{aligned} \tag{4.8}$$

Obviously, **(A)** implies assumptions $(H1)$-$(H3)$ in Theorem 2.1.

4.1 Main results

For any $\mu_0 \in \mathscr{P}_2$ and $r \geq 0$, let $B(\mu_0, r) = \{\nu \in \mathscr{P}_2 : \mathbb{W}_2(\mu_0, \nu) \leq r\}$. Let

$$\phi(s,t) = \lambda(t)^2 \bigg(\frac{\kappa_1(t)}{1 - \mathrm{e}^{-\kappa_1(t)(t-s)}} + \frac{t\kappa_2(t)^2 \exp[2(t-s)(\kappa_1(t)+\kappa_2(t))]}{2} \bigg), \quad 0 \leq s < t.$$

Under assumption **(A)**, we have the following result for the log-Harnack inequality and regularity estimates on P_t.

Theorem 4.1. *Assume* **(A)** *and let* $t > s \geq 0$.

(1) *For any* $\mu_0, \nu_0 \in \mathscr{P}_2$,

$$(P_{s,t} \log f)(\nu_0) \leq \log (P_{s,t} f)(\mu_0) + \phi(s,t) \mathbb{W}_2(\mu_0, \nu_0)^2, \quad f \in \mathscr{B}_b^+(\mathbb{R}^d). \tag{4.9}$$

Consequently,

$$|\nabla P_{s,t} f|^2 \leq 2\phi(s,t) \big\{ P_{s,t} f^2 - (P_{s,t} f)^2 \big\}, \quad f \in \mathscr{B}_b(\mathbb{R}^d). \tag{4.10}$$

(2) *For any different* $\mu_0, \nu_0 \in \mathscr{P}_2$, *and any* $f \in \mathscr{B}_b(\mathbb{R}^d)$,

$$\begin{aligned} &\frac{|(P_{s,t} f)(\mu_0) - (P_{s,t} f)(\nu_0)|^2}{\mathbb{W}_2(\mu_0, \nu_0)^2} \\ &\leq 2\phi(s,t) \sup_{\nu \in B(\mu_0, \mathbb{W}_2(\mu_0,\nu_0))} \big\{ (P_{s,t} f^2)(\nu) - (P_{s,t} f)^2(\nu) \big\}. \end{aligned} \tag{4.11}$$

Consequently,

$$\begin{aligned} \|P_{s,t}^* \mu_0 - P_{s,t}^* \nu_0\|_{var} &:= 2 \sup_{A \in \mathscr{B}_b(\mathbb{R}^d)} |(P_{s,t}^* \mu_0)(A) - (P_{s,t}^* \nu_0)(A)| \\ &\leq \sqrt{2\phi(s,t)} \mathbb{W}_2(\mu_0, \nu_0). \end{aligned} \tag{4.12}$$

Next, when $\|\sigma_t\|_\infty$ is locally bounded in t, we have the following result on Harnack inequality with power $p > 1$ and applications.

Theorem 4.2. *Assume* **(A)** *and that for some increasing* $\gamma : [0, \infty) \to (0, \infty)$,

$$|\{\sigma_t(x) - \sigma_t(y)\}^*(x-y)| \leq \gamma(t)|x-y|, \quad t \geq 0. \tag{4.13}$$

Let

$$p(t) = (1 + 4\lambda(t)\gamma(t))^2, \quad \Gamma(t) = \kappa_2(t)^2 \lambda(t)^2 T \mathrm{e}^{2\kappa_1(t) + 2\kappa_2(t)}.$$

Then for any $\mu_0, \nu_0 \in \mathscr{P}_2$ *and* $\mathscr{F}_0$*-measurable random variables* X_0, Y_0 *with* $\mathscr{L}_{X_0} = \mu_0, \mathscr{L}_{Y_0} = \nu_0$,

$$\begin{aligned} (P_{s,t} f)^p(\nu_0) \leq &(P_{s,t} f^p)(\mu_0) \exp\bigg[\frac{\sqrt{p}\,\Gamma(t) \mathbb{W}_2(\mu_0,\nu_0)^2}{(\sqrt{p}+1)[2(\sqrt{p}-1)^2 - 16\lambda(t)^2\gamma(t)^2]} \bigg] \\ &\times \bigg(\mathbb{E} \exp\bigg[\frac{2\lambda(t)^2 \kappa_1(t) |X_0 - Y_0|^2}{(\sqrt{p}-1)^2 (1 - \mathrm{e}^{-\kappa_1(t)(t-s)})} \bigg] \bigg)^{\frac{\sqrt{p}\,(\sqrt{p}-1)^2}{(\sqrt{p}+1)[2(\sqrt{p}-1)^2 - 16\lambda(t)^2\gamma(t)^2]}}, \\ &t > s \geq 0, p \geq p(t), f \in \mathscr{B}_b^+(\mathbb{R}^d). \end{aligned} \tag{4.14}$$

In particular, for any $x,y\in\mathbb{R}^d$, $t>s\geq 0$ and $p\geq p(t)$,

$$(P_{s,t}f)^p(x)\leq (P_{s,t}f^p)(y)\exp\bigg[\frac{\sqrt{p}|x-y|^2\big(\Gamma(t)+\frac{2\kappa_1(t)\lambda(t)^2}{1-\exp[-\kappa_1(t)(t-s)]}\big)}{(\sqrt{p}+1)[2(\sqrt{p}-1)^2-16\lambda(t)^2\gamma(t)^2]}\bigg]. \tag{4.15}$$

Below we present some consequences of the above Harnack inequalities.

Corollary 4.3. *Assume* **(A)** *and let $t>s\geq 0$.*

(1) *For any $\mu_0,\nu_0\in\mathscr{P}_2$, $P_{s,t}^*\mu_0$ and $P_{s,t}^*\nu_0$ are equivalent and the Radon-Nykodim derivative satisfies the entropy estimate*

$$\int_{\mathbb{R}^d}\Big\{\log\frac{\mathrm{d}P_{s,t}^*\nu_0}{\mathrm{d}P_{s,t}^*\mu_0}\Big\}\mathrm{d}P_{s,t}^*\nu_0\leq\phi(s,t)\mathbb{W}_2(\mu_0,\nu_0)^2. \tag{4.16}$$

Consequently, in the situation of Theorem 3.1(2),

$$\int_{\mathbb{R}^d}\Big\{\log\frac{\mathrm{d}P_t^*\nu_0}{\mathrm{d}P_t^*\mu}\Big\}\mathrm{d}P_t^*\nu_0\leq\phi(0,1)\mathrm{e}^{-(C_2-C_1)(t-1)}\mathbb{W}_2(\mu,\nu_0)^2,\quad t\geq 1.$$

(2) *If* (4.13) *holds, then for any $t>s\geq 0$ and $p\geq p(t)$,*

$$\begin{aligned}&\int_{\mathbb{R}^d}\Big\{\frac{\mathrm{d}P_{s,t}^*\nu_0}{\mathrm{d}P_{s,t}^*\mu_0}\Big\}^{\frac{1}{p}}\mathrm{d}(P_{s,t}^*\nu_0)\\ &\leq\exp\bigg[\frac{\Gamma(t)\mathbb{W}_2(\mu_0,\nu_0)^2}{(1+p^{-\frac12})[2(\sqrt{p}-1)^2-16\lambda(t)^2\gamma(t)^2]}\bigg]\\ &\quad\times\bigg(\mathbb{E}\exp\bigg[\frac{2\lambda(t)^2\kappa_1(t)|X_0-Y_0|^2}{(\sqrt{p}-1)^2(1-\mathrm{e}^{-\kappa_1(t)(t-s)})}\bigg]\bigg)^{\frac{(\sqrt{p}-1)^2}{(1+p^{-\frac12})[2(\sqrt{p}-1)^2-16\lambda(t)^2\gamma(t)^2]}}\end{aligned} \tag{4.17}$$

for $\mathscr{F}_0$-measurable random variables X_0,Y_0 with $\mathscr{L}_{X_0}=\mu_0,\mathscr{L}_{Y_0}=\nu_0$.

Proof. According to the proof of [21, Theorem 1.4.1], when μ_0 and ν_0 are Dirac measures, these results follow from (4.9) and (4.15) respectively. In general, the proof is completely similar. Precisely, for a $(P_{s,t}^*\mu_0)$-null set A and $n\geq 1$, we apply (4.9) to $f:=n1_A+1$, so that

$$(P_{s,t}^*\nu_0)(A)\log(n+1)=(P_{s,t}^*\log f)(\nu_0)\leq\phi(s,t)\mathbb{W}_2(\mu_0,\nu_0),\quad n\geq 1.$$

Letting $n\to\infty$ we obtain $(P_{s,t}^*\nu_0)(A)=0$, so that $P_{s,t}^*\nu_0$ is absolutely continuous with respect to $P_{s,t}^*\mu_0$. By the symmetry, $P_{s,t}^*\mu_0$ is also absolutely continuous with respect to $P_{s,t}^*\nu_0$. Moreover, (4.16) follows from (4.9) by taking $f=\frac{\mathrm{d}P_{s,t}^*\nu_0}{\mathrm{d}P_{s,t}^*\mu_0}$, while (4.17) follows from (4.14) by taking $f=(\frac{\mathrm{d}P_{s,t}^*\nu_0}{\mathrm{d}P_{s,t}^*\mu_0})^{\frac1p}$. □

4.2 Proof of Theorem 4.1

Without loss of generality, we only prove for $s = 0$. As in [23, §2], for fixed $T > 0$, let

$$\xi_t = \frac{1}{\kappa_1(T)}\Big(1 - \mathrm{e}^{\kappa_1(T)(t-T)}\Big), \quad t \in [0, T]. \tag{4.18}$$

Let $\nu_t = P_t^*\nu_0$ and let Y_0 be $\mathscr{F}_0$-measurable with $\mathscr{L}_{Y_0} = \nu_0$. Consider the SDE

$$\mathrm{d}Y_t = \Big\{b_t(Y_t, \nu_t) + \frac{1}{\xi_t}\sigma_t(Y_t)\sigma_t(X_t)^{-1}(X_t - Y_t)\Big\}\mathrm{d}t + \sigma_t(Y_t)\mathrm{d}W_t. \tag{4.19}$$

By **(A)** and $\sup_{t\in[0,T]} \nu_t(|\cdot|^2) < \infty$ due to Theorem 2.1, this SDE has a unique solution $(Y_t)_{t\in[0,T)}$. Let

$$\tau_n := T \wedge \inf\{t \in [0, T) : |X_t| + |Y_t| \geq n\}, \quad n \geq 1.$$

We have $\tau_n \uparrow T$ as $n \uparrow \infty$. To verify step (S2), we first prove that

$$R_s := \exp\bigg[\int_0^s \frac{1}{\xi_t}\big\langle\sigma_t(X_t)^{-1}(Y_t - X_t), \mathrm{d}W_t\big\rangle - \frac{1}{2}\int_0^s \frac{|\sigma_t(X_t)^{-1}(Y_t - X_t)|^2}{\xi_t^2}\mathrm{d}t\bigg] \tag{4.20}$$

is a uniformly integrable martingale for $s \in [0, T]$.

Lemma 4.4. *Assume* **(A)**. *Let X_0, Y_0 be two $\mathscr{F}_0$-measurable random variables such that $\mathscr{L}_{X_0} = \mu_0, \mathscr{L}_{Y_0} = \nu_0$ and*

$$\mathbb{E}|X_0 - Y_0|^2 = \mathbb{W}_2(\mu_0, \nu_0)^2. \tag{4.21}$$

Then $(R_s)_{s\in[0,T]}$ is a uniformly integrable martingale with

$$\sup_{t\in[0,T]} \mathbb{E}[R_t \log R_t] \leq \phi(0, T)\mathbb{W}_2(\mu_0, \nu_0)^2. \tag{4.22}$$

Proof. By **(A)**, for any $n \geq 1$ the process $(R_{s\wedge\tau_n})_{s\in[0,T)}$ is a uniformly integrable continuous martingale. Since $\tau_n \uparrow T$ as $n \uparrow \infty$, by the martingale convergence theorem, it suffices to prove

$$\sup_{t\in[0,T],n\geq 1} \mathbb{E}[R_{t\wedge\tau_n} \log R_{t\wedge\tau_n}] \leq \phi(0, T)\mathbb{W}_2(\mu_0, \nu_0)^2. \tag{4.23}$$

We fix $t \in (0, T)$ and $n \geq 1$. By Girsnaov's theorem,

$$\tilde{W}_s := W_s - \frac{1}{\xi_s}\sigma_s(X_s)^{-1}(Y_s - X_s), \quad s \in [0, t \wedge \tau_n]$$

is a d-dimensional Brownian motion under the weighted probability $\mathbb{Q}_{t,n} := R_{t\wedge\tau_n}\mathbb{P}$. Reformulating (4.2) and (4.19) as

$$\begin{aligned}
\mathrm{d}X_s &= \Big\{b_s(X_s, \mu_s) - \frac{X_s - Y_s}{\xi_s}\Big\}\mathrm{d}s + \sigma_s(X_s)\mathrm{d}\tilde{W}_s,\\
\mathrm{d}Y_s &= b_s(Y_s, \nu_s)\mathrm{d}s + \sigma_s(Y_s)\mathrm{d}\tilde{W}_s, \quad s \in [0, t \wedge \tau_n],
\end{aligned}$$

by **(A)** and Itô's formula under probability $\mathbb{Q}_{t,n}$, we obtain

$$\mathrm{d}|X_s - Y_s|^2 \le \Big\{\kappa_1(s)|X_s - Y_s|^2 + \kappa_2(s)|X_s - Y_s|\mathbb{W}_2(\mu_s, \nu_s) - \frac{2|X_s - Y_s|^2}{\xi_s}\Big\}\mathrm{d}s + \mathrm{d}M_s$$

for $s \in [0, t \wedge \tau_n]$ and some $\mathbb{Q}_{t,n}$-martingale M_s. Then

$$\begin{aligned} \mathrm{d}\frac{|X_s - Y_s|^2}{\xi_s} \le & \frac{\mathrm{d}M_s}{\xi_s} + \frac{\kappa_2(s)^2\mathbb{W}_2(\mu_s, \nu_s)^2}{2}\mathrm{d}s \\ & - \frac{|X_s - Y_s|^2}{\xi_s^2}\Big\{2 - \kappa_1(s)\xi_s + \xi_s' - \frac{1}{2}\Big\}\mathrm{d}s, \quad s \in [0, t \wedge \tau_n]. \end{aligned} \tag{4.24}$$

By (4.18) and the monotonicity of κ_1, we have

$$2 - \kappa_1(s)\xi_s + \xi_s' - \frac{1}{2} \ge 2 - \kappa_1(T)\xi_s + \xi_s' - \frac{1}{2} = \frac{1}{2}.$$

Moreover, since (4.8) implies $(H2)$ for $K_1 = \frac{\kappa_1 + \sqrt{\kappa_1^2 + 4\kappa_2^2}}{2} \le \kappa_1 + \kappa_2$, it follows from Theorem 2.1 that

$$\mathbb{W}_2(\mu_s, \nu_s) \le W_2(\mu_0, \nu_0)\mathrm{e}^{s\{\kappa_1(T) + \kappa_2(T)\}}, \quad s \in [0, T].$$

Substituting these into (4.24) and using (4.21), we arrive at

$$\begin{aligned} & \mathbb{E}_{\mathbb{Q}_{t,n}} \int_0^{t\wedge\tau_n} \frac{|X_s - Y_s|^2}{\xi_s^2}\mathrm{d}s \\ & \le \frac{2}{\xi_0} + \mathbb{E}_{\mathbb{Q}_{t,n}} \int_0^{t\wedge\tau_n} \kappa_2(s)^2\mathbb{W}_2(\mu_s, \nu_s)^2\mathrm{d}s \\ & \le \bigg[\frac{2}{\xi_0} + T\kappa_2(T)^2 \exp[2T(\kappa_1(T) + \kappa_2(T))]\bigg]\mathbb{W}_2(\mu_0, \nu_0)^2. \end{aligned} \tag{4.25}$$

Writing

$$\log R_{t\wedge\tau_n} = \int_0^{t\wedge\tau_n} \frac{1}{\xi_s}\big\langle \sigma_s(X_s)^{-1}(Y_s - X_s), \mathrm{d}\tilde{W}_s\big\rangle + \frac{1}{2}\int_0^{t\wedge\tau_n} \frac{|\sigma_s(X_s)^{-1}(Y_s - X_s)|^2}{\xi_s^2}\mathrm{d}s,$$

by $\|\sigma_t^{-1}\| \le \lambda(t)$ due to (4.7), $\xi_0 = \frac{1}{\kappa_1(T)}(1 - \mathrm{e}^{-\kappa_1(T)T})$ due to (4.18), and using (4.25), we arrive at

$$\mathbb{E}[R_{t\wedge\tau_n} \log R_{t\wedge\tau_n}] = \frac{1}{2}\mathbb{E}_{\mathbb{Q}_{t,n}} \int_0^{t\wedge\tau_n} \frac{|\sigma_s(X_s)^{-1}(Y_s - X_s)|^2}{\xi_s^2}\mathrm{d}s \le \phi(0, T)\mathbb{W}_2(\mu_0, \nu_0)^2.$$

Therefore, (4.23) holds since $t \in (0, T)$ and $n \ge 1$ are arbitrary. □

Proof of Theorem 4.1. (1) By Lemma 4.4 and the Girsanov theorem, $\mathrm{d}\mathbb{Q} := R_T\mathrm{d}\mathbb{P}$ is a probability measure such that

$$\tilde{W}_s := W_s - \int_0^s \frac{\sigma_t(X_t)^{-1}(Y_t - X_t)}{\xi_t}\mathrm{d}t, \quad s \in [0, T] \tag{4.26}$$

is a d-dimensional Brownian motion. Then (4.19) reduces to

$$\mathrm{d}Y_t = b_t(Y_t, \nu_t) + \sigma_t(Y_t)\mathrm{d}\tilde{W}_t. \tag{4.27}$$

Consider the DDSDE

$$\mathrm{d}\tilde{X}_t = b_t(\tilde{X}_t, \mathscr{L}_{\tilde{X}_t}|_{\tilde{\mathbb{P}}})\mathrm{d}t + \sigma_t(\tilde{X}_t)\mathrm{d}\tilde{W}_t, \quad \tilde{X}_0 = Y_0.$$

By the weak uniqueness we have $\mathscr{L}_{\tilde{X}_t}|_{\tilde{\mathbb{P}}} = P_t^*\nu_0 = \nu_t$ for $t \in [0,T]$. Combining this with (4.27) and the strong uniqueness, we conclude that $\tilde{X}_t = Y_t$ for $t \in [0,T]$. In particular, $\mathscr{L}_{Y_T} = \nu_T$ as required in (S2). Therefore, (4.5) and Lemma 4.4 lead to

$$(P_T \log f)(\nu_0) \le \log(P_T f)(\mu_0) + \phi(0,T)\mathbb{W}_2(\mu_0,\nu_0)^2.$$

In particular,

$$P_T \log f(x) \le (\log P_T f)(y) + \phi(0,T)|x-y|^2, \quad x,y \in \mathbb{R}^d, f \in \mathscr{B}_b(\mathbb{R}^d).$$

According to [2, Proposition 2.3], this implies (4.10).

(2) Let $\mathbb{W}_2(\mu_0,\nu_0) > 0$. We first assume that μ_0 is absolutely continuous with respect to the Lebesgue measure. In this case, by [18, Theorem 10.4.1] (see [14] when ν_0 is also absolutely continuous), there exists a measurable map $F : \mathbb{R}^d \to \mathbb{R}^d$ such that $\Xi(x) := x + F(x)$ maps μ_0 into ν_0; that is, $\nu_0 = \mu_0 \circ \Xi^{-1}$. Let

$$\Xi_s(x) = x + sF(x), \quad \mu_s = \mu_0 \circ \Xi_s^{-1}, \quad s \in [0,1].$$

Then it is easy to see that

$$\mathbb{W}_2(\mu_s,\mu_t) = |t-s|\mathbb{W}_2(\mu_0,\nu_0), \quad s,t \in [0,1].$$

Now, for any $n \ge 1$ and $0 \le i \le n-1$, we have

$$\mathbb{W}_2(\mu_{i/n}, \mu_{(i+1)/n}) = \varepsilon_n := \frac{1}{n}\mathbb{W}_2(\mu_0.\nu_0).$$

For any $f \in \mathscr{B}_b(\mathbb{R}^d)$ and $c > 0$, when n is large enough such that $c\varepsilon_n f + 1 > 0$, the log-Harnack inequality implies

$$P_T \log(c\varepsilon_n f + 1)(\mu_{i/n}) \le \log(c\varepsilon_n P_T f + 1)(\mu_{(i+1)/n}) + \varepsilon_n^2 \phi(0,T), \quad 0 \le i \le n-1. \tag{4.28}$$

By Taylor's expansion, there exists a constant $c(f) > 0$ depending on $\|f\|_\infty$ such that

$$\begin{aligned}
&\Big|P_T \log(c\varepsilon_n f + 1)(\mu_{i/n}) - c\varepsilon_n (P_T f)(\mu_{i/n}) + \frac{(c\varepsilon_n)^2}{2}(P_T f^2)(\mu_{i/n})\Big| \le \frac{c(f)}{n^3},\\
&\Big|\log(c\varepsilon_n P_T f + 1)(\mu_{(i+1)/n}) - c\varepsilon_n (P_T f)(\mu_{(i+1)/n}) + \frac{(c\varepsilon_n)^2}{2}(P_T f)^2(\mu_{(i+1)/n})\Big| \le \frac{c(f)}{n^3}.
\end{aligned}$$

Substituting these into (4.28), we obtain

$$\big|(P_T f)(\mu_{(i+1)/n}) - (P_T f)(\mu_{i/n})\big|$$

$$\le \frac{c\varepsilon_n}{2}\big|(P_Tf^2)(\mu_{i/n}) - (P_Tf)^2(\mu_{(i+1)/n})\big| + \frac{\phi(0,T)\varepsilon_n}{c} + \frac{2c(f)}{cn^2\mathbb{W}_2(\mu_0,\nu_0)}, \quad 0\le i\le n-1.$$

Therefore,

$$\begin{aligned}|(P_Tf)(\mu_0) - (P_Tf)(\nu_0)| &\le \sum_{i=0}^{n-1}\big|(P_Tf)(\mu_{(i+1)/n}) - (P_Tf)(\mu_{i/n})\big| \\ &\le \frac{c}{2}\sum_{i=0}^{n-1}\varepsilon_n\big|(P_Tf^2)(\mu_{i/n}) - (P_Tf)^2(\mu_{(i+1)/n})\big| + \frac{\phi(0,T)\mathbb{W}_2(\mu_0,\nu_0)}{c} + \mathrm{O}(n^{-1}).\end{aligned}$$

Noting that $\varepsilon_n = \frac{1}{n}\mathbb{W}_2(\mu_0,\nu_0)$, by letting $n\to\infty$, we obtain

$$\frac{|(P_Tf)(\mu_0) - (P_Tf)(\nu_0)|}{\mathbb{W}_2(\mu_0,\nu_0)} \le \frac{c}{2}\sup_{\nu\in B(\mu_0,\mathbb{W}_2(\mu_0,\nu_0))}\big|(P_Tf^2)(\nu) - (P_Tf)^2(\nu)\big| + \frac{\phi(0,T)}{c}.$$

Minimizing the upper bound in $c>0$, we prove (4.11). Since

$$(P_T^*\nu)(A) - \{(P_T^*\nu)(A)\}^2 \le \frac{1}{4}, \quad A\in\mathscr{B}(\mathbb{R}^d), \nu\in\mathscr{P},$$

(4.12) follows from (4.11) with $f = 1_A$.

In general, for any $\mu_0\in\mathscr{P}_2$, we take a sequence $\{\mu_0^{(n)}\}_{n\ge1}\subset\mathscr{P}_2$ converging to μ_0 in $\mathbb{W}_2$ and having densities with respect to the Lebesgue measure. Then $\mu_0^{(n)}$ converges to μ_0 weakly. By (4.10), this implies

$$\begin{aligned}&\lim_{n\to0}(P_{s,t}f(\mu_0^{(n)}) = (P_{s,t}f(\mu_0),\\ &\lim_{n\to\infty}\sup_{\nu\in B(\mu_0^{(n)},\mathbb{W}_2(\mu_0^{(n)},\nu_0))}\big\{(P_{s,t}f^2)(\nu) - (P_{s,t}f)^2(\nu)\big\}\\ &= \sup_{\nu\in B(\mu_0,\mathbb{W}_2(\mu_0,\nu_0))}\big\{(P_{s,t}f^2)(\nu) - (P_{s,t}f)^2(\nu)\big\}.\end{aligned}$$

Therefore, by (4.11) with $\mu_0^{(n)}$ replacing μ_0 which we just proved, and letting $n\to\infty$, we finish the proof. □

4.3 Proof of Theorem 4.2

Again, we only prove for $s=0$. By (4.13), for any $r>0$ we have

$$\begin{aligned}&\exp\bigg[-\frac{r\Gamma(T)\mathbb{W}_2(\mu_0,\nu_0)^2}{\lambda(T)^2}\bigg]\mathbb{E}_{\mathbb{Q}}\Big[\mathrm{e}^{r\int_0^{t\wedge\tau_n}\frac{|X_s-Y_s|^2}{\xi_s^2}\mathrm{d}s}\Big]\\ &\le \mathbb{E}_{\mathbb{Q}}\Big[\mathrm{e}^{\frac{2r|X_0-Y_0|^2}{\xi_0}+4r\int_0^{t\wedge\tau_n}\frac{\langle X_s-Y_s,\{\sigma_s(X_s)-\sigma_s(Y_s)\}\mathrm{d}\tilde W_s\rangle}{\xi_s}}\Big]\\ &\le \mathbb{E}_{\mathbb{Q}}\Big[\mathrm{e}^{\frac{2r|X_0-Y_0|^2}{\xi_0}}\mathbb{E}_{\mathbb{Q}}\Big(\mathrm{e}^{4r\int_0^{t\wedge\tau_n}\frac{\langle X_s-Y_s,\{\sigma_s(X_s)-\sigma_s(Y_s)\}\mathrm{d}\tilde W_s\rangle}{\xi_s}}\Big|\mathscr{F}_0\Big)\Big] \qquad (4.29)\\ &\le \mathbb{E}_{\mathbb{Q}}\Big[\mathrm{e}^{\frac{2r|X_0-Y_0|^2}{\xi_0}}\sqrt{\mathbb{E}_{\mathbb{Q}}\Big(\mathrm{e}^{32r^2\int_0^{t\wedge\tau_n}\frac{|\{\sigma_s(X_s)-\sigma_s(Y_s)\}^*(X_s-Y_s)|^2}{\xi_s^2}\mathrm{d}s}\Big|\mathscr{F}_0\Big)}\Big]\\ &\le \sqrt{\mathbb{E}_{\mathbb{Q}}\mathrm{e}^{\frac{4r|X_0-Y_0|^2}{\xi_0}}}\sqrt{\mathbb{E}_{\mathbb{Q}}\mathrm{e}^{32r^2\gamma(T)^2\int_0^{t\wedge\tau_n}\frac{|X_s-Y_s|^2}{\xi_s^2}\mathrm{d}s}},\end{aligned}$$

where we have used the inequality $\mathbb{E}_{\mathbb{Q}}(\mathrm{e}^{M_t}|\mathscr{F}_0) \le \sqrt{\mathbb{E}_{\mathbb{Q}}(\mathrm{e}^{2\langle M\rangle_t}|\mathscr{F}_0)}$ for a continuous $\mathbb{Q}$-martingale M_t. When $r \le \frac{1}{32\gamma(T)^2}$, by Jensen's inequality

$$\sqrt{\mathbb{E}_{\mathbb{Q}}\mathrm{e}^{32r^2\gamma(T)^2\int_0^{t\wedge\tau_n}\frac{|X_s-Y_s|^2}{\xi_s^2}\mathrm{d}s}} \le \left(\mathbb{E}_{\mathbb{Q}}\left[\mathrm{e}^{r\int_0^{t\wedge\tau_n}\frac{|X_s-Y_s|^2}{\xi_s^2}\mathrm{d}s}\right]\right)^{16r\gamma(T)^2},$$

so that (4.29) implies

$$\mathbb{E}_{\mathbb{Q}}\left[\mathrm{e}^{r\int_0^{t\wedge\tau_n}\frac{|X_s-Y_s|^2}{\xi_s^2}\mathrm{d}s}\right] \le \mathrm{e}^{\frac{r\Gamma(T)\mathbb{W}_2(\mu_0,\nu_0)^2}{\lambda(T)^2(1-16r\gamma(T)^2)}}\left(\mathbb{E}_{\mathbb{Q}}\left[\mathrm{e}^{\frac{4r\kappa_1(T)|X_0-Y_0|^2}{1-\mathrm{e}^{-\kappa_1(T)T}}}\right]\right)^{\frac{1}{2-32r\gamma(T)^2}}.$$

Letting $n\uparrow\infty$ and $t\uparrow T$, and noting that $\mathbb{Q}|_{\mathscr{F}_0}=\mathbb{P}|_{\mathscr{F}_0}$ since $R_0=1$, we obtain

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}\left[\mathrm{e}^{r\int_0^{T}\frac{|X_s-Y_s|^2}{\xi_s^2}\mathrm{d}s}\right] \le &\mathrm{e}^{\frac{r\Gamma(T)\mathbb{W}_2(\mu_0,\nu_0)^2}{\lambda(T)^2(1-16r\gamma(T)^2)}}\left(\mathbb{E}\left[\mathrm{e}^{\frac{4r\kappa_1(T)|X_0-Y_0|^2}{1-\mathrm{e}^{-\kappa_1(T)T}}}\right]\right)^{\frac{1}{2-32r\gamma(T)^2}},\\ &\text{if } 0\le r\le\frac{1}{32\gamma(T)^2}.\end{aligned}\tag{4.30}$$

On the other hand, letting $M_t=\int_0^t\frac{1}{\xi_s}\langle\sigma_s(X_s)^{-1}(Y_s-X_s),\mathrm{d}\tilde{W}_s\rangle,\ t\in[0,T]$, by (4.20) we have $R_T=\mathrm{e}^{M_T+\frac{1}{2}\langle M\rangle_T}$. So,

$$\begin{aligned}&\mathbb{E}R_T^{\frac{p}{p-1}}=\mathbb{E}_{\mathbb{Q}}R_T^{\frac{1}{p-1}}=\mathbb{E}_{\mathbb{Q}}\mathrm{e}^{\frac{M_T}{p-1}+\frac{\langle M\rangle_T}{2(p-1)}}\\ &\le\left(\mathbb{E}_{\mathbb{Q}}\exp\left[\frac{1}{\sqrt{p}-1}M_T-\frac{1}{2(\sqrt{p}-1)^2}\langle M\rangle_T\right]\right)^{\frac{1}{1+\sqrt{p}}}\left(\mathbb{E}_{\mathbb{Q}}\exp\left[\frac{\langle M\rangle_T}{2(\sqrt{p}-1)^2}\right]\right)^{\frac{\sqrt{p}}{\sqrt{p}+1}}\\ &\le\left(\mathbb{E}_{\mathbb{Q}}\exp\left[\frac{\lambda(T)^2}{2(\sqrt{p}-1)^2}\int_0^T\frac{|X_s-Y_s|^2}{\xi_s^2}\mathrm{d}s\right]\right)^{\frac{\sqrt{p}}{\sqrt{p}+1}}.\end{aligned}\tag{4.31}$$

Since $p\ge p(T)$ implies $\frac{\lambda(T)^2}{2(\sqrt{p}-1)^2}\le\frac{1}{32\gamma(T)^2}$, this and (4.30) with $r=\frac{\lambda(T)^2}{2(\sqrt{p}-1)^2}$ yield

$$\mathbb{E}R_T^{\frac{p}{p-1}}\le\mathrm{e}^{\frac{\sqrt{p}\,\Gamma(T)\mathbb{W}_2(\mu_0,\nu_0)^2}{(\sqrt{p}+1)\{2(\sqrt{p}-1)^2-16\lambda(T)^2\gamma(T)^2\}}}\left(\mathbb{E}_{\mathbb{Q}}\mathrm{e}^{\frac{2\lambda(T)^2\kappa_1(T)|X_0-Y_0|^2}{(\sqrt{p}-1)^2(1-\mathrm{e}^{-\kappa_1(T)T})}}\right)^{\frac{\sqrt{p}\,(\sqrt{p}-1)^2}{(\sqrt{p}+1)\{2(\sqrt{p}-1)^2-16\lambda(T)^2\gamma(T)^2\}}}.$$

Substituting into (4.6), we finish the proof.

5 Shift Harnack inequality and applications

In this section we establish the shift Harnack inequality and integration by parts formula introduced in [24]. Since the study for the multiplicative noise case is very complicated, here we only consider the additive noise for which the DDSDE (1.6) reduces to

$$\mathrm{d}X_t=b_t(X_t,\mathscr{L}_{X_t})\mathrm{d}t+\sigma_t\mathrm{d}W_t.\tag{5.1}$$

Theorem 5.1. *Let $\sigma : [0,\infty) \to \mathbb{R}^d \otimes \mathbb{R}^d$ and $b : [0,\infty) \times \mathbb{R}^d \times \mathscr{P}_2 \to \mathbb{R}^d$ are measurable such that σ_t is invertible with $\|\sigma_t\| + \|\sigma_t^{-1}\|$ locally bounded in $t \geq 0$, and $b_t(\cdot, \mu_t)$ is differentiable with*

$$\int_0^T \|\nabla b_t(\cdot, \mu_t)\|_\infty^2 \mathrm{d}t < \infty, \quad T > 0, \mu_\cdot \in C([0,T]; \mathscr{P}_2).$$

(1) *For any $p > 1, t > s \geq 0, \mu_0 \in \mathscr{P}_2, v \in \mathbb{R}^d$ and $f \in \mathscr{B}_b^+(\mathbb{R}^d)$,*

$$\begin{aligned}(P_{s,t}f)^p(\mu_0) \leq &(P_{s,t}f^p(v + \cdot))(\mu_0)\\ &\times \exp\left[\frac{p\sqrt{p}\,|v|^2 \int_s^t \|\sigma_r^{-1}\|^2\big\{1 + (r-s)\|\nabla b_r(\cdot, P_{s,r}^*\mu_0)\|_\infty\big\}^2 \mathrm{d}r}{2(p-1)(\sqrt{p}+1)(t-s)^2}\right].\end{aligned}$$

Moreover, for any $f \in \mathscr{B}_b^+(\mathbb{R}^d)$,

$$(P_{s,t}\log f)(\mu_0) \leq \log(P_{s,t}f(v+\cdot))(\mu_0) + \frac{|v|^2}{2(t-s)^2}\int_s^t \|\sigma_r^{-1}\|^2\Big(1+(r-s)\|\nabla b_r(\cdot,\mu_r)\|_\infty\Big)^2 \mathrm{d}r.$$

(2) *For any $t > s \geq 0, f \in C^1(\mathbb{R}^d)$ and $\mathscr{F}_s$-measurable random variable $X_{s,s}$ with $\mu_0 := \mathscr{L}_{X_{s,s}} \in \mathscr{P}_2$,*

$$\mathbb{E}(\nabla_v f)(X_{s,t}) = \mathbb{E}\bigg[\frac{f(X_{s,t})}{t-s}\int_s^t (r-s)\big\langle \sigma_r^{-1}\nabla_v b_r(\cdot, P_{s,r}^*\mu_0)(X_{s,r}),\ \mathrm{d}W_r\big\rangle\bigg], \quad v \in \mathbb{R}^d.$$

Proof. Without loss of generality, we only prove for $s = 0$ and $t = T$ for some fixed time $T > 0$. Denote $\mu_t = P_t^*\mu_0 = \mathscr{L}_{X_t}, t \geq 0$. Then (5.1) becomes

$$\mathrm{d}X_t = b_t(X_t, \mu_t)\mathrm{d}t + \sigma_t \mathrm{d}W_t, \quad \mathscr{L}_{X_0} = \mu_0. \tag{5.2}$$

Let $Y_t = X_t + \frac{tv}{T},\ t \in [0,T]$. Then

$$\mathrm{d}Y_t = b_t(Y_t, \mu_t)\mathrm{d}t + \sigma_t \mathrm{d}\tilde{W}_t, \quad \mathscr{L}_{Y_0} = \mu_0, t \in [0,T],$$

where

$$\begin{aligned}\tilde{W}_t &:= W_t + \int_0^t \xi_s \mathrm{d}s,\\ \xi_t &:= \sigma_t^{-1}\Big\{\frac{v}{T} + b_t(X_t,\mu_t) - b_t\Big(X_t + \frac{tv}{T}, \mu_t\Big)\Big\}.\end{aligned}$$

Let $R_T = \exp[-\int_0^T \langle \xi_t, \mathrm{d}W_t\rangle - \frac{1}{2}\int_0^T |\xi_s|^2 \mathrm{d}s]$. By the Girsanov theorem we obtain

$$(P_T f)(\mu_0) = \mathbb{E}[R_T f(Y_T)] = \mathbb{E}[R_T f(X_T + v)] \leq (P_T f^p(v+\cdot))^{\frac{1}{p}}(\mu_0)\big(\mathbb{E}R_T^{\frac{p}{p-1}}\big)^{\frac{p-1}{p}}.$$

This proves (1) since similarly to (4.31), we have

$$\mathbb{E}R_T^{\frac{p}{p-1}} = \mathbb{E}_{\mathbb{Q}}R_T^{\frac{1}{p-1}} \leq \Big(\mathbb{E}_{\mathbb{Q}}\mathrm{e}^{\frac{p}{2(p-1)^2}\int_0^T |\xi_s|^2 \mathrm{d}s}\Big)^{\frac{\sqrt{p}}{\sqrt{p}+1}}$$

$$\leq \exp\left[\frac{p\sqrt{p}\,|v|^2\int_0^T \|\sigma_t^{-1}\|^2\big\{1+t\|\nabla b_t(\cdot,P_t^*\mu_0)\|_\infty\big\}^2\mathrm{d}t}{2(p-1)^2(\sqrt{p}+1)T^2}\right].$$

To prove (2), we let $X_t^\varepsilon = X_t + \frac{\varepsilon t v}{T}$ for $\varepsilon\in(0,1)$ and $t\in[0,T]$. Using εv replace v, the above argument implies

$$(P_Tf)(\mu_0) = \mathbb{E}[R_T^\varepsilon f(X_T+\varepsilon v)],\quad \varepsilon\in(0,1),$$

where

$$R_T^\varepsilon := \exp\left[-\int_0^T\langle\xi_t^\varepsilon,\mathrm{d}W_t\rangle - \frac{1}{2}\int_0^T|\xi_s^\varepsilon|^2\mathrm{d}s\right],$$
$$\xi_t^\varepsilon := \sigma_t^{-1}\left\{\frac{\varepsilon v}{T} + b_t(X_t,\mu_t) - b_t\Big(X_t+\frac{\varepsilon t v}{T},\mu_t\Big)\right\}.$$

Therefore,

$$0 = \lim_{\varepsilon\to0}\frac{1}{\varepsilon}\mathbb{E}[R_T^\varepsilon f(X_T+\varepsilon v) - f(X_T)]$$
$$= \mathbb{E}[(\nabla_v f)(X_T)] - \mathbb{E}\left[\frac{f(X_T)}{T}\int_0^T r\big\langle\sigma_r^{-1}\nabla_v b_r(\cdot,P_{0,r}^*\mu_0)(X_r),\ \mathrm{d}W_r\big\rangle\right].$$

Then the proof is finished. □

As applications of Theorem 5.1, we have the following estimates on the density of $P_{s,t}^*$.

Corollary 5.2. *In the situation of Theorem* 5.1*, for any $t>s\geq0$ and $\mu_0\in\mathscr{P}_2$, $(P_{s,t}^*\mu_0)(\mathrm{d}x)=\rho_{s,t}^{\mu_0}(x)\mathrm{d}x$ for some density function $\rho_{s,t}^{\mu_0}$ satisfying the following estimates:*

$$\begin{aligned}&\int_{\mathbb{R}^d}|\nabla\log\rho_{s,t}^{\mu_0}(x)|^p\rho_{s,t}^{\mu_0}(x)\mathrm{d}x\\ &\leq\left\{\frac{(1\vee\frac{p(p-1)}{2})}{(t-s)^2}\int_s^t(r-s)^2\|\sigma_r^{-1}\|^2\|\nabla b_r(\cdot,P_{s,r}^*\mu_0)\|_\infty^2\mathrm{d}r\right\}^{\frac{p}{2}(1\wedge\frac{1}{p-1})},\quad p>1;\end{aligned}\tag{5.3}$$

$$\begin{aligned}&\int_{\mathbb{R}^d}\big\{\rho_{s,t}^{\mu_0}(x)\big\}^{\frac{p}{p-1}}\mathrm{d}x\\ &\leq\left(\frac{p\sqrt{p}\int_s^t\|\sigma_r^{-1}\|^2\{1+(r-s)\|\nabla b_2(\cdot,P_{s,r}^*\mu_0)\|_\infty\}^2\mathrm{d}r}{4\pi(p-1)(\sqrt{p}+1)(t-s)^2}\right)^{\frac{d}{2(p-1)}},\quad p>1;\end{aligned}\tag{5.4}$$

$$\begin{aligned}&\int_{\mathbb{R}^d}\rho_{s,t}^{\mu_0}(x)\log\rho_{s,t}^{\mu_0}(x)\mathrm{d}x\\ &\leq\frac{d}{2}\log\left(\frac{\int_s^t\|\sigma_r^{-1}\|^2\{1+(r-s)\|\nabla b_2(\cdot,P_{s,r}^*\mu_0)\|_\infty\}^2\mathrm{d}r}{4\pi(\sqrt{p}+1)(t-s)^2}\right).\end{aligned}\tag{5.5}$$

Proof. According to [24, Theorem 2.4, Theorem 2.5], the desired assertions follow from Theorem 5.1 . Indeed, by [24, Theorem 2.4], the integration by parts formula in Theorem 5.1(2) implies the existence of density $\rho_{s,t}^{\mu_0}$ and

$$\int_{\mathbb{R}^d} |\nabla \log \rho_{s,t}^{\mu_0}(x)|^p \rho_{s,t}^{\mu_0}(x)\mathrm{d}x = \mathbb{E}|\nabla \log \rho_{s,t}^{\mu_0}|^p(X_{s,t}) = \mathbb{E}\big|\mathbb{E}(N|X_{s,t})\big|^p \le \mathbb{E}|N|^p,$$

where

$$N := \frac{1}{t-s}\int_s^t (r-s)\{\sigma_r^{-1}\nabla b_r(\cdot, P_{s,r}^*\mu_0)(X_{s,r})\}^*\mathrm{d}W_r.$$

Then estimate (5.3) follows. Moreover, it is easy to see that the proof of [24, Theorem 2.5] also applies to $P_{s,t}^*\mu_0$ in place of $P(x,\cdot)$, so that estimates (5.4) and (5.5) follow from Theorem 5.1(1).

□

6 DDSDEs for homogeneous Landau equation

We consider the homogeneous Landau equation with $r \in [0,1]$ on $\mathbb{R}^3$ (see e.g. [19]):

$$\partial_t f_t = \frac{1}{2}\mathrm{div}\bigg(\int_{\mathbb{R}^3} |\cdot - y|^{2+\gamma}\Big(I - \frac{(\cdot - y)\otimes(\cdot - y)}{|\cdot - y|^2}\Big)\Big\{f_t(y)\nabla f_t - f_t \nabla f_t(y)\Big\}\mathrm{d}y\bigg). \tag{6.1}$$

Let $a(x) = |x|^\gamma(|x|^2 I - x\otimes x)$ and

$$b_0(x) := \mathrm{div} a(x) = -2|x|^\gamma x, \quad \sigma_0(x) := |x|^{\frac{\gamma}{2}}\begin{pmatrix} x_2 & 0 & x_3 \\ -x_1 & x_3 & 0 \\ 0 & -x_2 & -x_1 \end{pmatrix}. \tag{6.2}$$

Then $\sigma_0\sigma_0^* = a$. Take

$$\begin{aligned} b_t(x,\mu) = b(x,\mu) &:= -2\int_{\mathbb{R}^d} |x-z|^\gamma (x-z)\mu(\mathrm{d}z), \\ \sigma_t(x,\mu) = \sigma(x,\mu) &:= \int_{\mathbb{R}^d} \sigma_0(x-z)\mu(\mathrm{d}z). \end{aligned} \tag{6.3}$$

Then the density of $\mathscr{L}_{X_t}$ for the DDSDE (1.6) is a weak solution to (6.1). In this section we consider (1.6) for this specific choice of b and σ.

6.1 The case with Maxwell molecules: $\gamma = 0$

When $\gamma = 0$, both b_0 and σ_0 in (6.2) are Lipschitz continuous. Below we consider a more general model. For two Lipschitz continuous maps

$$b_0 : \mathbb{R}^d \to \mathbb{R}^d, \quad \sigma_0 : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d,$$

let

$$b^\alpha(x,\mu) := \int_{\mathbb{R}^d} b_0(x-\alpha z)\mu(\mathrm{d}z),\ \ \sigma^\alpha(x,\mu) := \int_{\mathbb{R}^d} \sigma_0(x-\alpha z)\mu(\mathrm{d}z), \quad \alpha\in\mathbb{R}, x\in\mathbb{R}^d, \mu\in\mathscr{P}_2.$$

For fixed $\alpha,\beta\in\mathbb{R}$, consider the DDSDE

$$\mathrm{d}X_t = b^\alpha(X_t,\mathscr{L}_{X_t})\mathrm{d}t + \sigma^\beta(X_t,\mathscr{L}_{X_t})\mathrm{d}W_t. \tag{6.4}$$

Theorem 6.1. *Let $\alpha,\beta\in\mathbb{R}$, $B_0 := \|\nabla b_0\|_\infty < \infty$ and $C_0 := \sup_{|v|=1,x\in\mathbb{R}^d}\|\nabla_v\sigma_0(x)\|_{HS}^2 < \infty$. Moreover, let $K_0\in\mathbb{R}$ such that*

$$\langle b_0(x)-b_0(y), x-y\rangle \le K_0|x-y|^2,\quad x,y\in\mathbb{R}^d.$$

(1) *For any $\mathscr{F}_0$-measurable X_0 with $\mathbb{E}|X_0|^2<\infty$, the equation* (6.4) *has a unique solution and $\sup_{t\in[0,T]}\mathbb{E}|X_t|^2<\infty$ for all $T>0$. Moreover, $X_t(x)$ is jointly continuous in $(t,x)\in[0,\infty)\times\mathbb{R}^d$.*

(2) *For any $\mu_0,\nu_0\in\mathscr{P}_2$,*

$$\mathbb{W}_2(P_t^*\mu_0, P_t^*\nu_0)^2 \le \mathbb{W}_2(\mu_0,\nu_0)^2\mathrm{e}^{(2K_0+C_0(1+|\beta|)^2+2|\alpha|B_0)t},\quad t\ge 0.$$

If, in particular, $2K_0+C_0(1+|\beta|)^2+2|\alpha|<0$, then P_t^* has a unique invariant probability measure.

(3) *If $\beta=0$ and σ_0 is invertible with $\lambda:=\|\sigma_0^{-1}\|_\infty<\infty$, then assertions in Theorem* 4.1*, Theorem* 4.2 *and Corollary* 4.3 *hold for*

$$\phi(s,t) = \lambda^2\bigg(\frac{2K_0+C_0}{1-\mathrm{e}^{-(2K_0+C_0)(t-s)}} + (t-s)|\alpha|\mathrm{e}^{2(t-s)(2K_0+C_0+2|\alpha|)}\bigg).$$

(4) *If σ_0 is constant and invertible, then assertions in Theorem* 5.1 *and Corollary* 5.2 *hold for $\sigma_r\equiv\sigma_0$.*

Proof. Since b_0 and σ_0 are Lipschitz continuous, it is easy to see that $(H1)$-$(H3)$ and (2.5) hold for $(b_t,\sigma_t)\equiv(b^\alpha,\sigma^\beta)$ for all $t\ge0$. Then the first assertion follows from Theorems 2.1 and 2.2.

To prove the second assertion using Theorem 3.1(2), we observe that for any $\pi\in\mathscr{C}(\mu,\nu)$.

$$\begin{aligned}&\langle b^\alpha(x,\mu)-b^\alpha(y,\nu), x-y\rangle\\ &= \int_{\mathbb{R}^d}\Big(\langle b_0(x-\alpha z)-b_0(y-\alpha z), x-y\rangle + \langle b_0(y-\alpha z)-b_0(y-\alpha z'), x-y\rangle\Big)\pi(\mathrm{d}z,\mathrm{d}z')\\ &\le K_0|x-y|^2 + B_0|\alpha|\cdot|x-y|\int_{\mathbb{R}^d}|z-z'|\pi(\mathrm{d}z,\mathrm{d}z').\end{aligned}$$

Then

$$\begin{aligned}2\langle b^\alpha(x,\mu)-b^\alpha(y,\nu), x-y\rangle &\le 2K_0|x-y|^2 + 2|\alpha|B_0\mathbb{W}_1(\mu,\nu)|x-y|\\ &\le (2K_0+|\alpha|B_0)|x-y|^2 + |\alpha|B_0\mathbb{W}_2(\mu,\nu)^2.\end{aligned}\tag{6.5}$$

Similarly,

$$\begin{aligned}\|\sigma^\beta(x,\mu)-\sigma^\beta(y,\nu)\|_{HS}^2 &\le C_0\big\{|x-y|+|\beta|\mathbb{W}_1(\mu,\nu)\big\}^2\\ &\le C_0(1+|\beta|)|x-y|^2+C_0(|\beta|+\beta^2)\mathbb{W}_2(\mu,\nu)^2.\end{aligned}$$

Combining this with (6.5) we obtain

$$\begin{aligned}&2\langle b^\alpha(x,\mu)-b^\alpha(y,\nu),x-y\rangle+\|\sigma^\beta(x,\mu)-\sigma^\beta(y,\nu)\|_{HS}^2\\ &\le\{2K_0+|\alpha|B_0+C_0(1+|\beta|)\}|x-y|^2+\{|\alpha|B_0+C_0|\beta|(1+|\beta|)\}\mathbb{W}_2(\mu,\nu)^2.\end{aligned}$$

Then the second assertion follows from Theorem 3.1(2).

Finally, by (6.5) and $\|\sigma_0(x)-\sigma_0(y)\|_{HS}^2\le C_0|x-y|^2$ we have

$$2\langle b^\alpha(x,\mu)-b^\alpha(y,\nu),x-y\rangle+\|\sigma_0(x)-\sigma_0(y)\|_{HS}^2\le(2K_0+C_0)|x-y|^2+2|\alpha||x-y|\mathbb{W}_2(\mu,\nu).$$

Then assumption **(A)** holds for $\lambda(t)=\lambda,\kappa_1(t)=2K_0+C_0$ and $\kappa_2(t)=2|\alpha|$. Therefore, assertions (3) and (4) follow from Theorem 4.1, Theorem 4.2, Corollary 4.3, Theorem 5.1 and Corollary 5.2. □

Coming back to the DDSDE for the homogeneous Landau equation with Maxwell molecules, i.e. (6.4) for b_0 and σ_0 in (6.2), Theorem 6.1 applies with $B_0=C_0=2$ and $K_0=-2$, so that we have the following result.

Corollary 6.2. *Let b_0 and σ_0 be in* (6.2) *and let $\alpha,\beta\in\mathbb{R}$. For any $\mathscr{F}_0$-measurable X_0 with $\mathbb{E}|X_0|^2<\infty$, the equation* (6.4) *has a unique solution and $\sup_{t\in[0,T]}\mathbb{E}|X_t|^2<\infty$ for all $T>0$. Moreover, $X_t(x)$ is jointly continuous in $(t,x)\in[0,\infty)\times\mathbb{R}^d$. Moreover, for any $\mu_0,\nu_0\in\mathscr{P}_2$,*

$$\mathbb{W}_2(P_t^*\mu_0,P_t^*\nu_0)^2\le\mathbb{W}_2(\mu_0,\nu_0)^2\mathrm{e}^{\{4(|\alpha|+|\beta|)+2\beta^2-2\}t},\quad t\ge0,\mu_0,\nu_0\in\mathscr{P}_2.$$

When $2(|\alpha|+|\beta|)+\beta^2<1$, P_t^* has a unique invariant probability measure μ and

$$\mathbb{W}_2(P_t^*\nu_0,\mu)^2\le\mathrm{e}^{-2(1-2|\alpha|-2|\beta|-\beta^2)t}\mathbb{W}_2(\nu_0,\mu)^2,\quad t\ge0,\nu_0\in\mathscr{P}_2.$$

When $\beta=\alpha=1$ which corresponds to the homogeneous Landau equation with Maxwell molecules,

$$\mathbb{W}_2(P_t^*\mu_0,P_t^*\nu_0)^2\le\mathbb{W}_2(\mu_0,\nu_0)^2\mathrm{e}^{8t},\quad t\ge0,\mu_0,\nu_0\in\mathscr{P}_2. \tag{6.6}$$

Remark 6.1. Let $N(z,A)$ denote the normal distribution on $\mathbb{R}^d$ with mean $z\in\mathbb{R}^d$ and covariance A, and let $\alpha=\beta=1$ in Corollary 6.2 for the homogeneous Landau equation with Maxwell molecules. According to [3, Theorem 1.1] (see also [6]), there exists a constant $p>0$ such that if

$$\int_{\mathbb{R}^d}|x|^p\mu_0(\mathrm{d}x)+\int_{\mathbb{R}^d}|\xi|^p|\hat\mu_0(\xi)|^2\mathrm{d}\xi<\infty,$$

where $\hat\mu_0$ is the fourier transform of μ_0, then

$$\|P_t^*\mu_0-N(z_0,\gamma_0^2I)\|_{var}\le c_1\mathrm{e}^{-c_2t},\quad t\ge0$$

holds for some constants $c_1, c_2 > 0$ depending on μ_0, where $z_0 := \int_{\mathbb{R}^3} x\mu_0(\mathrm{d}x)$ and $\gamma_0^2 := \int_{\mathbb{R}^d} |x - z_0|^2 \mu_0(\mathrm{d}x)$. See [4] for exponential convergence in the case that $\gamma \in (0, 1]$. Therefore, P_t^* is not ergodic since the limit distribution varies in the initial one. This fits the inequality (6.6) where the upper bound does not go to 0 as $t \to \infty$. However, it seems that the sharp upper bound in (6.6) should be bounded in t.

6.2 The case with hard potentials: $\gamma \in [0, 1]$

When $\gamma \in [0, 1]$, the weak existence and uniqueness have been proved in [8]. To prove the same assertion for strong solutions, we first present a result for the equivalence of the weak existence/uniqueness and the strong existence/uniqueness.

Theorem 6.3. *Let $\theta \geq 1$. Assume that for any $\mu \in C([0, \infty) \to \mathscr{P}_\theta)$ the SDE*

$$\mathrm{d}X_t = b_t(X_t, \mu_t)\mathrm{d}t + \sigma_t(X_t, \mu_t)\mathrm{d}W_t \tag{6.7}$$

has strong existence and uniqueness for X_0 with $\mathscr{L}_{X_0} = \mu_0$. Then for initial distribution $\mu_0 \in \mathscr{P}_\theta$, the DDSDE (1.6) *has weak existence (respectively uniqueness) if and only if it has strong existence* (*respectively uniqueness*).

Proof. (a) Since the strong existence implies the weak one, it suffices to prove the strong existence from the weak one. For any initial distribution $\mu_0 \in \mathscr{P}_\theta$, let $(\bar{X}_t, \bar{W}_t)$ be a weak solution under probability $\bar{\mathbb{P}}$. We have

$$\mathrm{d}\bar{X}_t = b_t(\bar{X}_t, \mu_t)\mathrm{d}t + \sigma_t(\bar{X}_t, \mu_t)\mathrm{d}\bar{W}_t, \tag{6.8}$$

where $\mu_t := \mathscr{L}_{\bar{X}_t}|_{\bar{\mathbb{P}}}$. Now, given a Brownian motion under the probability $\mathbb{P}$, let X_t be a strong solution to (6.7) with $\mathscr{L}_{X_0} = \mu_0$. By Yamada-Watanabe's principle for SDE, the strong existence and uniqueness of (6.7) imply the weak uniqueness, so that $\mathscr{L}_{X_t} = \mu_t$ so that (6.7) reduces to the DDSDE (1.6). Then the strong solution to (6.7) is also a strong solution to (1.6).

(b) Obviously, the weak uniqueness implies the strong uniqueness. On the other hand, let (1.6) has strong uniqueness, we aim to prove the weak uniqueness. Let $(X_t^{(i)}, W_t^{(i)})$ under probability $\mathbb{P}^i (i = 1, 2)$ be two weak solutions to (1.6) with $\mathscr{L}_{X_0^{(1)}}|_{\mathbb{P}^1} = \mathscr{L}_{X_0^{(2)}}|_{\mathbb{P}^2} = \mu_0$, we aim to prove

$$\mathscr{L}_{X^{(1)}}|_{\mathbb{P}^1} = \mathscr{L}_{X^{(2)}}|_{\mathbb{P}^2}. \tag{6.9}$$

Let $\mu_t = \mathscr{L}_{X_t^{(1)}}|_{\mathbb{P}^1}$. By assumption, the SDE

$$\mathrm{d}X_t = b_t(X_t, \mu_t)\mathrm{d}t + \sigma_t(X_t, \mu_t)\mathrm{d}W_t^{(2)}, \quad X_0 = X_0^{(2)} \tag{6.10}$$

has a unique strong solution $X := (X_t)_{t \geq 0}$. By Yamada-Watanabe's principle, (6.10) also has weak uniqueness. So,

$$\mathscr{L}_X|_{\mathbb{P}^2} = \mathscr{L}_{X^{(1)}}|_{\mathbb{P}^1}. \tag{6.11}$$

In particular, $\mathscr{L}_{X_t}|_{\mathbb{P}^2} = \mu_t$, so that X_t is also a strong solution to (1.6) with the given Brownian motion $W_t^{(2)}$ replacing W_t. Since $X_t^{(2)}$ solves the same DDSDE, by the strong uniqueness of (1.6) we have $X = X^{(2)}$. Combining this with (6.11), we prove (6.9). $\square$

Now, we consider the DDSDE (1.6) with b_t and σ_t in (6.3) for $\gamma \in (0,1]$.

Corollary 6.4. *Let with b_t and σ_t in* (6.3) *for $\gamma \in (0,1]$. Then for any $\mathscr{F}_0$-measurable X_0 with density f_0 satisfying* (1.5)*, the DDSDE* (1.6) *has a unique strong solution such that $\mathbb{E}e^{|X_t|^\alpha} < \infty$ for any $t > 0$.*

Proof. By [8, Theorem 2], the SDDE has a unique weak solution such that $\mathbb{E}e^{|X_t|^\alpha} < \infty$ for any $t > 0$. According to Theorem 6.3, the same holds for the strong solution. $\square$

Acknowledgement. The author would like to thank Jianhai Bao, Arnaud Guillin and Xing Huang for corrections and helpful comments.

References

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

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

[3] E. A. Carlen, E. Gabetta, G. Toscani, *Propagation of smoothness and the rate of exponential convergence to equilibrium for a spatially homogeneous Maxwellian gas,* Commun. Math. Phys. 305(1999), 521–546.

[4] K. Carrapatoso, *Exponential convergence to equilibrium for the homogeneous Landau equation with hard potentials,* Bull. Sci. Math. 139(2015), 777–805.

[5] L. Desvillettes, C. Villani, *On the spatially homogeneous Landau equation for hard potentials, Part I : existence, uniqueness and smothness,* Comm. Part. Diff. Equat. 25(2000), 179–259.

[6] L. Desvillettes, C. Villani, *On the spatially homogeneous Landau equation for hard potentials, Part II: H-Theorem and Applications,* Comm. Part. Diff. Equat. 25(2000), 261–298.

[7] S. Fang, T. Zhang, *A study of a class of stochastic differential equations with non-Lipschitzian coefficients,* Probab. Theory Relat. Fields 132(2005), 356–390.

[8] N. Fournier, A. Guillin, *From a Kac-like particle system to the Landau equation for hard potentials and Maxwell molecules,* arXiv:1510.01123.

[9] T. Funaki, *The diffusion approximation of the spatially homogeneous Boltzman equation,* Duke J. Math. 521(1985), 1–23.

[10] H. Guérin, *Existence and regularity of a weak function-solution for some Landau equations with a stochastic approach,* Stoch. Proc. Appl. 101(2002), 303–325.

[11] I. Ikeda, S. Watanabe, *Stochastic differential equations and Diffusion processes,* North-Holland, Amsterdam, 1981.

[12] M. Kac, *Foundations of kinetic theory,* Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III, University of California Press, 171–197.

[13] A. N. Kolmogoroff, *Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung,* Math. Ann. 104(1931), 415–458.

[14] J. R. MacCann, *Existence and uniqueness of monotone measure-preserving maps,* Duke Math. J. 80(1995), 309–323.

[15] M. Röckner, F.-Y. Wang, *Log-harnack inequality for stochastic differential equations in Hilbert spaces and its consequences*, Infin. Dim. Anal. Quat. Probab. Relat. Top. 13(2010), 27–37.

[16] D.W. Stroock, S.R.S. Varadhan, *Multidimensional Diffusion Processes,* Springer, New York, 1979.

[17] Sznitman, Weak solution for controllded nonlinear SDEs, 1984.

[18] C. Villani, *Optimal Transport, Old and New,* Springer, 2009.

[19] C. Villani, *On the spatially homogeneous Landau equation for Maxwellian Mocecules,* Math. Mod. Meth. Appl. Sci. 8(1998), 957–983.

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

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

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

[23] F.-Y. Wang, *Harnack inequality for SDE with multiplicative noise and extension to Neumann semigroup on nonconvex manifolds,* Ann. Probab. 39(2011), 1449–1467.

[24] F.-Y.Wang, *Integration by parts formula and shift Harnack inequality for stochastic equations,* Ann. Probab. 42(2014), 994–1019.

[25] T. Yamada, S. Watanabe *On the uniqueness of solutions of stochastic differential equations I & II,* J. Math. Kyoto Univ. 11(1971), 155–167; 553–563.

[26] X. Zhang, *Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients,* Electr. J. Probab. 16(2011), 1096–1116.