Singular Degenerate SDEs: Well-Posedness and Exponential Ergodicity *

Panpan Ren[c)], **Martin Grothaus**[b)], **Feng-Yu Wang**[a)]

[a)] Center for Applied Mathematics, Tianjin University, Tianjin 300072, China

[b)] Mathematics Department, University of Kaiserslautern-Landau, Kaiserslautern 67653, Germany

[c)] Mathematics Department, City University of Hong Kong, Hong Kong, Kowloon, China.

grothaus@mathematik.uni-kl.de, panparen@cityu.edu.hk, wangfy@tju.edu.cn

July 14, 2023

Abstract

The well-posedness and exponential ergodicity are proved for stochastic Hamiltonian systems containing a singular drift term which is locally integrable in the component with noise. As an application, the well-posedness and uniform exponential ergodicity are derived for a class of singular degenerated McKean-Vlasov SDEs.

AMS subject Classification: 60B05, 60B10.
Keywords: Singular degenerate SDE, well-posedness, exponential ergodicity.

1 Introduction

Let $d_1, d_2 \in \mathbb{N}$. For fixed $T \in (0, \infty]$, consider the following degenerate SDE for $(X_t, Y_t) \in \mathbb{R}^{d_1+d_2} = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$:

E1 (1.1)
$$\begin{cases} \mathrm{d}X_t = Z_t^{(1)}(X_t, Y_t)\mathrm{d}t, \\ \mathrm{d}Y_t = \big(Z_t^{(2)}(X_t, Y_t) + b_t(Y_t)\big)\mathrm{d}t + \sigma_t(Y_t)\mathrm{d}W_t, \quad t \in [0, T], \end{cases}$$

where $[0, T] := [0, \infty)$ when $T = \infty$, $(W_t)_{t\in[0,T]}$ is an m-dimensional Brownian motion on a complete filtrated probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\in[0,T]}, \mathbb{P})$, and

$$Z^{(i)} : [0, T] \times \mathbb{R}^{d_1+d_2} \to \mathbb{R}^{d_i}, \quad i = 1, 2,$$

*Feng-Yu Wang is supported in part by the National Key R&D Program of China (2022YFA1006000, 2020YFA0712900) and NNSFC (11831014, 11921001). Martin Grothaus and Panpan Ren acknowledge support by the Alexander von Humboldt Foundation.

$$b : [0,T] \times \mathbb{R}^{d_2} \to \mathbb{R}^{d_2}, \quad \sigma : [0,T] \times \mathbb{R}^{d_2} \to \mathbb{R}^{d_2} \otimes \mathbb{R}^m$$

are measurable. We assume that $Z_t^{(i)}(x,y)$ and $\sigma_t(y)$ are continuous in $(x,y) \in \mathbb{R}^{d_1+d_2}$ as "regular" coefficients, but $b_t(y)$ only satisfies a local integrability condition in (t,y) and is regarded as a "singular" term. See assumptions (A_1)-(A_3) below for details.

SDE (1.1) is known as stochastic Hamiltonian system when the drifts are given by the gradients of a Hamiltonian functional. The associated Fokker-Planck equation is called the Langevin equation or kinetic Fokker-Planck equation. In particular, when $Z_t^{(1)}(x,y) = y$, X_t and Y_t stand for the location and speed of a fluid flow at time t respectively.

In this paper, we investigate the well-posedness and exponential ergodicity of (1.1) with drift $Z_t^{(2)}(x,y) + b_t(y)$ discontinuous in (t,y). These two properties have been intensively studied when the coefficients are regular enough or the invariant probability measure is known, which we summary as follows.

When $Z_t^{(1)}(x,y)$ and $Z_t^{(2)}(x,y) + b_t(y)$ satisfy a Dini-Hölder continuity condition in (x,y), the well-posedness of (1.1) has been proved in [2, 18], see also the recent paper [8] for the weak well-posedness of (1.1) for $d_1 = d_2, Z^{(1)}(x,y) = y, \sigma = I_{d_2}$(the $d_2 \times d_2$ identity matrix) and $Z^{(2)} + b$ being in a weighted anisotropic Besov space.

When the SDE is time independent, the exponential ergodicity for special versions of (1.1) has been studied in many references. When the unique invariant probability measure is given, the hypocoercivity introduced by Villani [15] has attracted a lot of attentions , and has been further developed in a series of papers such as [1, 6, 7] based on an abstract analytic framework built up by Dolbeaut, Mouhot and Schmeiser [3]. However, the study in this direction heavily relies on the explicit formulation of the invariant probability measure, for which the drifts are given by (weighted) gradient of a Hamiltonian functional. A crucial motivation in the ergodicity theory is to simulate the invariant probability measure by using the stochastic system. In this spirit, we study the ergodicity for the above general model with unknown invariant probability measure. On the other hand, when the coefficients are regular enough, the exponential ergodicity follows from some modified dissipativity conditions, see for instance [19] for the hypercontractivity implying the exponential ergodicity in related entropy, and [14] for an extension to McKean-Vlasov SDEs.

In Section 2 and Section 3, we investigate the well-posedness and exponential ergodicity for (1.1) with $Z_t^{(2)}(x,y)$ only satisfying a local integrability condition in (t,y), such that the corresponding results derived in the recent papers [13, 16, 21] for non-degenerate SDEs are extended to the present degenerate setting. Finally, in Section 4 we extend the main results to McKean-Vlasov SDEs.

2 Well-posedness

In this part we let $T \in (0,\infty)$ be finite. For any $r > 0$ and $(x,y) \in \mathbb{R}^{d_1+d_2}$, let

$$B_r(x,y) := \big\{(x',y') \in \mathbb{R}^{d_1+d_2} : |x-x'| + |y-y'| \le r\big\}, \quad B_r(y) := \big\{y' \in \mathbb{R}^{d_2} : |y-y'| \le r\big\}.$$

For any

$$(p,q) \in \mathscr{K} := \Big\{(p,q) : \ p,q \in (2,\infty), \frac{d_2}{p} + \frac{2}{q} < 1\Big\},$$

define the $\tilde{L}_q^p$-norm of a (vector or real valued) measurable function f on $[0,T]\times\mathbb{R}^{d_2}$:

$$\|f\|_{\tilde{L}_q^p} := \sup_{y\in\mathbb{R}^{d_2}} \bigg(\int_0^T \|1_{B_1(y)} f_t\|_{L^p(\mathbb{R}^{d_2})}^q \mathrm{d}t \bigg)^{\frac{1}{q}}.$$

We write $f \in \tilde{L}_q^p$ if $\|f\|_{\tilde{L}_q^p} < \infty$ and $\tilde{L}^p$ represents the $\tilde{L}_q^p$-norm independent of q. We use $\nabla^{(1)}$ and $\nabla^{(2)}$ to denote the gradient operators in $x \in \mathbb{R}^{d_1}$ and $y \in \mathbb{R}^{d_2}$ respectively, so that $(\nabla^{(i)})^2$ is the corresponding Hessian operator. In case only one variable is concerned, we simply denote the gradient by ∇.

(A_1) For any $n \geq 1$ there exists a constant $0 < K_n < \infty$ such that

$$\max_{i=1,2} |Z_t^{(i)}(x,y) - Z_t^{(i)}(x',y')| \leq K_n(|x-x'| + |y-y'|), \quad t \in [0,T], (x,y), (x',y') \in B_n(0).$$

Moreover,

$$\sup_{t\in[0,T], i=1,2} |Z_t^{(i)}(0)| < \infty.$$

(A_2) $\sigma\sigma^*$ is invertible with $\|\sigma\|_\infty + \|(\sigma\sigma^*)^{-1}\|_\infty < \infty$, and

$$\lim_{\varepsilon\downarrow 0} \sup_{t\in[0,T], |y-y'|\leq\varepsilon} \|\sigma_t(y) - \sigma_t(y')\| = 0.$$

Moreover, there exist $l \in \mathbb{N}$, $\{(p_i,q_i)\}_{0\leq i\leq l}$ and $1 \leq f_i \in \tilde{L}_{q_i}^{p_i}, 0 \leq i \leq l$ such that

$$|b| \leq f_0, \quad \|\nabla\sigma\| \leq \sum_{i=1}^l f_i.$$

(A_3) There exist $\varepsilon \in (0,1)$ and $1 \leq V \in C^2(\mathbb{R}^{d_1+d_2})$ with

LY (2.1) $$\lim_{|x|+|y|\to\infty} V(x,y) = \infty, \quad \limsup_{|x|+|y|\to\infty} \sup_{y'\in B_\varepsilon(y)} \frac{|\nabla^{(2)}V(x,y')| + \|(\nabla^{(2)})^2 V(x,y')\|}{V(x,y)} < \infty,$$

such that

$$\varepsilon \sup_{y'\in B_\varepsilon(y)} \big\{ |Z_t^{(1)}(x,y)| \|\nabla^{(1)}\nabla^{(2)}V(x,y')\| + |Z_t^{(2)}(x,y)| \big(\|\nabla^{(2)}V(x,y')\| + \|(\nabla^{(2)})^2 V(x,y')\| \big) \big\}$$
$$+ \langle Z_t^{(1)}(x,y), \nabla V(\cdot,y)(x)\rangle + \langle Z_t^{(2)}(x,y), \nabla V(x,\cdot)(y)\rangle \leq \eta_t V(x,y)$$

holds for some $0 \leq \eta \in L^1([0,T])$, all $t \in [0,T]$ and $(x,y) \in \mathbb{R}^{d_1+d_2}$.

T1 **Theorem 2.1.** *Assume (A_1) and (A_2). Then for any initial value (X_0,Y_0) the SDE (1.1) has a unique strong solution up to the life time ζ. If (A_3) holds then $\zeta = T$ and*

EST (2.2) $$\mathbb{E}\Big[\sup_{t\in[0,T]} V(X_t,Y_t) \Big| \mathscr{F}_0 \Big] \leq cV(X_0,Y_0)$$

holds for some constant $0 < c < \infty$.

By a truncation argument, we first consider the case where $\{K_n\}_{n\geq 1}$ is bounded. In this case, we may follow the line of [20] to prove the well-posedness by using Zvonkin's transform. The only difference is that in the present degenerate setting we apply this transform for the following time-dependent elliptic operators on $\mathbb{R}^{d_2}$:

$$L_t := \frac{1}{2}\mathrm{tr}\big(\sigma_t\sigma_t^*\nabla^2\big) + b_t\cdot\nabla, \quad t\in[0,T].$$

Moreover, since Y_t also depends on X_t, we have to reprove the Khasminskii estimate for Y_t which is crucial in the proof of pathwise uniqueness.

L1 **Lemma 2.2.** *Assume (A_2) and that $\|Z^{(2)}\|_\infty<\infty$. Then for any $(p,q)\in\mathscr{K}$ there exists an increasing function $H:[0,\infty)\to[0,\infty)$ such that for any strong solution $(X_t,Y_t)_{t\in[0,T]}$ of (1.1),*

$$\mathbb{E}\big[\mathrm{e}^{\int_0^T|f_t(Y_t)|^2\mathrm{d}t}\big]\leq H(\|f\|_{\tilde L_q^p}), \quad f\in\tilde L_q^p.$$

Proof. Let

$$\gamma_s := \big(\sigma_s^*(\sigma_s\sigma_s^*)^{-1}Z_s^{(2)}\big)(X_s,Y_s), \quad s\in[0,T].$$

By (A_2) and the boundedness of $Z^{(2)}$, we have

BP0 (2.3) $$K := \int_0^T \big\|\gamma_s\big\|_\infty^2\mathrm{d}s<\infty.$$

So, by Girsanov's theorem,

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

is an m-dimensional Brownian motion under the probability measure $\mathbb{Q}:=R\mathbb{P}$, where

$$R := \mathrm{e}^{-\int_0^T\langle\gamma_s,\mathrm{d}W_s\rangle-\frac{1}{2}\int_0^T|\gamma_s|^2\mathrm{d}s}.$$

So, Y_t solves the SDE

$$\mathrm{d}Y_t = b_t(Y_t)\mathrm{d}t+\sigma_t(Y_t)\mathrm{d}\tilde W_t, \quad t\in[0,T].$$

By [20, Lemma 4.1], (A_2) implies that

$$\mathbb{E}_\mathbb{Q}\mathrm{e}^{\int_0^T|f_t(Y_t)|^2\mathrm{d}t}\leq H_0(\|f\|_{\tilde L_q^p})$$

holds for some increasing function $H_0:[0,\infty)\to[0,\infty)$. Thus,

$$\begin{aligned}&\mathbb{E}\mathrm{e}^{\int_0^T|f_t(Y_t)|^2\mathrm{d}t} = \mathbb{E}_\mathbb{Q}\big[R^{-1}\mathrm{e}^{\int_0^T|f_t(Y_t)|^2\mathrm{d}t}\big]\\ &\leq\big(\mathbb{E}[R^{-2}]\big)^{\frac{1}{2}}\Big(\mathbb{E}_\mathbb{Q}\mathrm{e}^{2\int_0^T|f_t(Y_t)|^2\mathrm{d}t}\Big]\Big)^{\frac{1}{2}} = \big(\mathbb{E}[R^{-2}]\big)^{\frac{1}{2}}\sqrt{H_0(2\|f\|_{\tilde L_q^p})}.\end{aligned}$$

Then the proof is finished since (2.3) implies

$$\begin{aligned}\mathbb{E}[R^{-2}] &= \mathbb{E}\big[\mathrm{e}^{-2\int_0^T\langle\gamma_s,\mathrm{d}W_s\rangle+\int_0^T|\gamma_s|^2\mathrm{d}s}\big]\\ &\leq\mathbb{E}\big[\mathrm{e}^{-2\int_0^T\langle\gamma_s,\mathrm{d}W_s\rangle-2\int_0^T|\gamma_s|^2\mathrm{d}s+3K}\big] = \mathrm{e}^{3K}<\infty.\end{aligned}$$

□

Proof of Theorem 2.1. (a) The well-posedness up to life time. By a truncation argument, in stead of (A_1) we may and do assume that $Z_t^{(i)}$ are bounded and Lipschitz continuous in (x,y) uniformly in $t\in[0,T]$, so that the life time of (X_t,Y_t) is T. By [20, Themore 3.2], (A_2) implies that for any $\lambda\geq 0$ the PDE

PDE (2.4) $$(\partial_t+L_t)u_t=-\lambda u_t-b_t,\quad t\in[0,T], u_T=0$$

for $u:[0,T]\times\mathbb{R}^{d_2}\to\mathbb{R}^{d_2}$ has a unique solution with

$$\|u\|_\infty+\|\nabla u\|_\infty+\|\nabla^2 u\|_{\tilde L^{p_0}_{q_0}}<\infty,$$

where ∇ and ∇^2 are the gradient and Hessian operators on $\mathbb{R}^{d_2}$. Moreover, for $\varepsilon\in(0,1)$ there exits $\lambda>0$ is large enough such that,

W1 (2.5) $$\|u\|_\infty+\|\nabla u\|_\infty<\varepsilon.$$

We now make the Zvonkin's transform for Y_t:

$$\tilde Y_t:=\Theta_t(Y_t),\quad \Theta_t(y):=y+u_t(y),\quad t\in[0,T], y\in\mathbb{R}^{d_2}.$$

By (2.4) and generalised Itô's formula (see [20]), (1.1) becomes

E1' (2.6) $$\begin{cases}\mathrm{d}X_t=\tilde Z_t^{(1)}(X_t,\tilde Y_t)\mathrm{d}t,\\ \mathrm{d}\tilde Y_t=\tilde Z_t^{(2)}(X_t,\tilde Y_t)\mathrm{d}t+\tilde\sigma_t(\tilde Y_t)\mathrm{d}W_t,\quad t\in[0,T],\end{cases}$$

where

$$\begin{aligned}&\tilde Z_t^{(1)}(x,y):=Z_t^{(1)}\big(x,\Theta_t^{-1}(y)\big),\\ &\tilde Z_t^{(2)}(x,y):=\big((\nabla\Theta_t)\circ\Theta_t^{-1}(y)\big)Z_t^{(2)}\big(x,\Theta_t^{-1}(y)\big)-\lambda u_t\circ\Theta_t^{-1}(y),\\ &\tilde\sigma_t(y):=\big((\nabla\Theta_t)\circ\Theta_t^{-1}(y)\big)\sigma_t\circ\Theta_t^{-1}(y),\quad t\in[0,T],(x,y)\in\mathbb{R}^{d_1+d_2}.\end{aligned}$$

Since $\|\nabla u\|_\infty<1$, Θ_t is diffeomorphism so that the well-posedness of (1.1) is equivalent to that of (2.6). Noting that the coefficients of (2.6) are bounded and continuous in (x,y), this SDE has a weak solution. By the Yamada-Watanable principle, it remains to prove the pathwise uniqueness of (2.6). This can be done as in [20] by using Khasminskii's estimate in Lemma 2.2. Below we present a detailed proof for completeness.

For any nonnegative measurable function f on $\mathbb{R}^{d_2}$, consider its maximal functional

$$\mathscr{M}f(x):=\sup_{r\in(0,1)}\frac{1}{|B_r(0)|}\int_{B_r(0)}f(x+y)\mathrm{d}y,\quad x\in\mathbb{R}^{d_2}.$$

By [20, Lemma 2.1], there exists a constant $0<c<\infty$ such that for any function $f\in L^\infty(\mathbb{R}^d)$ with $\nabla f\in L^1_{loc}(\mathbb{R}^d)$,

W2 (2.7) $$\begin{aligned}&|f(x)-f(y)|\leq c|x-y|\big(\|f\|_\infty+\mathscr{M}|\nabla f|(x)+\mathscr{M}|\nabla f|(y)\big),\quad x,y\in\mathbb{R}^{d_2},\\ &\|\mathscr{M}|f|\|_{\tilde L^p_q}\leq c\|f\|_{\tilde L^p_q}.\end{aligned}$$

Now, let $(X_t^{(i)}, \tilde{Y}_t^{(i)})_{t\in[0,T]}, i = 1, 2$, be two solutions of (2.6) with $(X_0^{(1)}, \tilde{Y}_0^{(1)}) = (X_0^{(2)}, \tilde{Y}_0^{(2)})$. By the Lipschitz continuity of $Z_t^{(i)}$ uniformly in t, (2.7) and $\|\nabla^2 u\|_{\tilde{L}_q^p} < \infty$, we find functions $1 \le g_j \in \tilde{L}_{q_j}^{p_j}$ $(0 \le j \le l)$ such that

$$\sum_{i=1,2} |\tilde{Z}_t^{(i)}(x,y) - \tilde{Z}_t^{(i)}(x',y')|^2 + \|\tilde{\sigma}_t(y) - \tilde{\sigma}_t(y')\|^2$$

$$\le \big(|x-x'|^2 + |y-y'|^2\big) \sum_{j=0}^{l} \big(g_j(t,y)^2 + g_j(t,y')^2\big), \quad t \in [0,T], (x,y), (x',y') \in \mathbb{R}^{d_1+d_2}.$$

Then by Itô's formula, we find a constant $0 < c < \infty$ and a martingale M_t such that

$$\xi_t := |X_t^{(1)} - X_t^{(2)}|^2 + |\tilde{Y}_t^{(1)} - \tilde{Y}_t^{(2)}|^2, \quad t \in [0,T]$$

satisfies

W3 (2.8) $$\mathrm{d}\xi_t \le c\xi_t \sum_{j=0}^{l} \sum_{i=1,2} |g_j(t, \tilde{Y}_t^{(i)})|^2 + \mathrm{d}M_t, \quad t \in [0,T], \xi_0 = 0.$$

By Lemma 2.2, for all $0 < \theta < \infty$, we have

$$\mathbb{E}\big[\mathrm{e}^{\theta \int_0^T |g_j(t,\tilde{Y}_t^{(i)})|^2 \mathrm{d}t}\big] < \infty, \quad 0 \le j \le l, \ i = 1, 2.$$

So, by the stochastic Gronwall inequality, see [21, Lemma 3.7], (2.8) implies $(X_t^{(1)}, \tilde{Y}_t^{(1)}) = (X_t^{(2)}, \tilde{Y}_t^{(2)})$ for all $t \in [0,T]$.

(b) Now, let (A_3) hold, we aim to prove (2.2). By (2.5), we have $|Y_t - \tilde{Y}_t| < \varepsilon$. Combining this with (2.1) it suffices to prove (2.2) for $\tilde{Y}_t$ replacing Y_t, i.e.

EST' (2.9) $$\mathbb{E}\Big[\sup_{t\in[0,T]} V(X_t, \tilde{Y}_t)\Big|\mathscr{F}_0\Big] \le cV(X_0, \tilde{Y}_0)$$

holds for some constant $0 < c < \infty$. By (2.5), (2.6), and Itô's formula, the boundedness of $\tilde{\sigma}$ and (A_3) imply that for some constant $0 < C < \infty$,

E0* (2.10) $$\mathrm{d}V(X_t, \tilde{Y}_t) \le C(1+\eta_t)V(X_t, \tilde{Y}_t)\mathrm{d}t + \mathrm{d}M_t, \quad t \in [0,T]$$

holds for some martingale M_t with

$$\mathrm{d}\langle M\rangle_t \le CV(X_t, \tilde{Y}_t)^2 \mathrm{d}t, \quad t \in [0,T].$$

Let $\tau_n := T \wedge \inf\{t \ge 0, |X_t^x| + |\tilde{Y}_t^y| \ge n\}$ and the life time

$$\zeta := \lim_{n\to\infty} \tau_n, \quad n \ge 1.$$

Then by the Gronwall Inequality we have

$$\mathbb{E}\big(V(X_{\tau_n}, \tilde{Y}_{\tau_n})1_{\tau_n<T}\big|\mathscr{F}_0\big) \le V(X_0, \tilde{Y}_0)\mathrm{e}^{c\int_0^T (1+\eta_s)\mathrm{d}s},$$

which implies

$$\mathbb{E}\big(1_{\tau_n<T}\big|\mathscr{F}_0\big) \le \frac{V(X_0,\tilde{Y}_0)\mathrm{e}^{c\int_0^T(1+\eta_s)\mathrm{d}s}}{\inf_{|x|+|y|\ge n}V(x,y)}.$$

Then by Fatou's Lemma, we have

$$\mathbb{E}\big(1_{\zeta\le T}\big|\mathscr{F}_0\big) = 0,$$

which further implies

$$\mathbb{P}(\zeta\le T) = \mathbb{E}\big(1_{\zeta\le T}\big) = \mathbb{E}\Big(\mathbb{E}\big(1_{\zeta\le T}\big|\mathscr{F}_0\big)\Big) = 0.$$

With the definition of the life time ζ, when $n\to\infty$, we have $\zeta = T$. Finally, by a standard argument using the Burkholder-Davis-Gundy and Gronwall inequalities, we prove (2.9) for some constant $c>0$. □

3 Exponential ergodicity

In this part we consider the time-homogeneous case such that (1.1) becomes

E3 (3.1)
$$\begin{cases}\mathrm{d}X_t = Z^{(1)}(X_t,Y_t)\mathrm{d}t,\\ \mathrm{d}Y_t = \big\{Z^{(2)}(X_t,Y_t)+b(Y_t)\big\}\mathrm{d}t+\sigma_t(Y_t)\mathrm{d}W_t,\quad t\ge 0,\end{cases}$$

where

$$Z^{(i)}:\mathbb{R}^{d_1+d_2}\to\mathbb{R}^{d_i},\quad i=1,2,\quad b:\mathbb{R}^{d_2}\to\mathbb{R}^{d_2},\quad \sigma:\mathbb{R}^{d_2}\to\mathbb{R}^{d_2}\otimes\mathbb{R}^m$$

are measurable. We investigate the ergodicity of the associated Markov process. To this end, we make the following assumption, which, according to Theorem 2.1, implies the well-posedness and non-explosion of this SDE.

(B_1) For any $1\le n<\infty$ there exists a constant $0<K_n<\infty$ such that

$$\sup_{i=1,2}|Z^{(i)}(x,y)-Z^{(i)}(x',y')| \le K_n(|x-x'|+|y-y'|),\quad (x,y),(x',y')\in B_n(0).$$

(B_2) $\sigma\sigma^*$ is invertible with $\|\sigma\|_\infty+\|(\sigma\sigma^*)^{-1}\|_\infty<\infty$, and there exists $p>(2\vee d)$ such that $|b|+\|\nabla\sigma\|\in\tilde{L}^p$.

(B_3) There exist constants $\varepsilon\in(0,1), 0<K<\infty$, an increasing function $\Phi:[1,\infty)\to(0,\infty)$ with $\Phi(n)\to\infty$ as $n\to\infty$, and $1\le V\in C^2(\mathbb{R}^{d_1+d_2})$ with

LY11 (3.2)
$$\lim_{|x|+|y|\to\infty}V(x,y)=\infty,\quad \limsup_{|x|+|y|\to\infty}\sup_{y'\in B_\varepsilon(y)}\frac{|\nabla V(x,\cdot)(y')|+\|\nabla^2V(x,\cdot)(y')\|}{\big(V\wedge\Phi(V)\big)(x,y)}=0,$$

such that

$$\varepsilon\sup_{y'\in B_\varepsilon(y)}\big\{|Z_t^{(1)}(x,y)|\|\nabla^{(1)}\nabla^{(2)}V(x,y')\|+|Z_t^{(2)}(x,y)|(|\nabla^{(2)}V(x,y')|+\|\nabla^{(2)}\nabla^{(2)}V(x,y')\|)\big\}$$
$$+\langle Z_t^{(1)}(x,y),\nabla V(\cdot,y)(x)\rangle+\langle Z_t^{(2)}(x,y),\nabla V(x,\cdot)(y)\rangle\le K-\Phi(V(x,y)).$$

By Theorem 2.1, under (B_1)-(B_3) the SDE (3.1) is well-posed. Let $\{P_t\}_{t\geq 0}$ be the associated Markov semigroup, i.e.

$$P_t f(x,y) = \mathbb{E}[f(X_t^{x,y}, Y_t^{x,y})], \quad f \in \mathscr{B}_b(\mathbb{R}^{d_1+d_2}), t \geq 0, (x,y) \in R^{d_1+d_2},$$

where $(X_t^{x,y}, Y_t^{x,y})$ solves (3.1) with initial value (x,y). We investigate the ergodicity of P_t, i.e. it has a unique invariant probability measure μ such that

$$\lim_{t\to\infty} P_t^*\nu = \mu, \quad \nu \in \mathscr{P},$$

where $\mathscr{P}$ is the space of all probability measures on $\mathbb{R}^{d_1+d_2}$, and

$$(P_t^*\nu)(f) := \nu(P_t f) = \int_{\mathbb{R}^{d_1+d_2}} P_t f \mathrm{d}\nu, \quad f \in \mathscr{B}_b(\mathbb{R}^{d_1+d_2}), \nu \in \mathscr{P}, t \geq 0.$$

In terms of (B_3), we consider the convergence under the V-variation norm

$$\|\mu_1 - \mu_2\|_V := \sup_{f\in\mathscr{B}_b(\mathbb{R}^{d_1+d_2}), |f|\leq V} |\mu_1(f) - \mu_2(f)|.$$

Under this norm, the space

$$\mathscr{P}_V := \big\{\mu \in \mathscr{P} : \ \mu(V) < \infty\big\}$$

is a complete metric sapce. When $V \equiv 1$, we denote the norm by $\|\cdot\|_{var}$ which is known as the total variation norm. The Lyapunov condition (B_3) implies the existence of invariant probability measure.

3.1 Main results and example

To prove the ergodicity, we need the following assumption that any compact set is a petite set of P_t:

(B_4) Any compact set D of $\mathbb{R}^{d_1+d_2}$ is P_t-petite, i.e. there exists $t_0 > 0$ and a non-trivial finite measure ν such that

$$\inf_{x\in D} P_{t_0}(x,\cdot) \geq \nu,$$

where $P_t(x,\cdot)$ is the transition probability kernel of P_t at $x \in \mathbb{R}^{d_1+d_2}$.

T1.6.1 **Theorem 3.1.** *Assume (B_1)-(B_3). Then the following assertions hold.*

(1) *P_t has an invariant probability measure μ such that*

$$\mu\big(\Phi(\varepsilon_0 V)\big) = \int_{\mathbb{R}^{d_1+d_2}} \Phi\big(\varepsilon_0 V(x,y)\big)\mu(\mathrm{d}x, \mathrm{d}y) < \infty$$

holds for some $\varepsilon_0 > 0$.

(2) *If P_t is t_0-regular for some $t_0 > 0$, i.e. $\{P_{t_0}(x,\cdot) : x \in \mathbb{R}^{d_1+d_2}\}$ are mutually equivalent, then*

EGD (3.3) $$\lim_{t\to\infty} P_t f(x,y) = \mu(f), \quad \mu \in \mathscr{P},\ f \in \mathscr{B}(\mathbb{R}^{d_1+d_2}).$$

(3) *If (B_4) holds and $\Phi(r) \geq \delta r$ for some constant $\delta > 0$ and all $r \geq 0$, then there exist constants $1 < c < \infty, \lambda > 0$ such that*

EX1 (3.4) $$\|P_t^*\mu_1 - P_t^*\mu_2\|_V \leq c\mathrm{e}^{-\lambda t}\|\mu_1 - \mu_2\|_V, \quad \mu_1, \mu_2 \in \mathscr{P}_V, t \geq 0.$$

Consequently, $\mu \in \mathscr{P}_V$ is the unique invariant probability measure of P_t, and

$$\|P_t^*\nu - \mu\|_V \leq c\mathrm{e}^{-\lambda t}\|\nu - \mu\|_V, \quad \nu \in \mathscr{P}_V, t \geq 0.$$

(4) *Let (B_4) hold and $H(r) := \int_0^r \frac{\mathrm{d}s}{\Phi(s)} < \infty$ for $r \geq 0$. If Φ is convex, then there exist constants $1 < k < \infty, \lambda > 0$ such that*

EX0 (3.5) $$\|P_t^*\delta_{(x,y)} - \mu\|_V \leq k\big\{1 + H^{-1}(H(V(x,y)) - k^{-1}t)\big\}\mathrm{e}^{-\lambda t}, \quad (x,y) \in \mathbb{R}^{d_1+d_2}, t \geq 0,$$

where H^{-1} is the inverse of H with $H^{-1}(r) := 0$ for $r \leq 0$. Consequently, if $H(\infty) < \infty$ then there exist constants $0 < c, \lambda, t^* < \infty$ such that

EX2 (3.6) $$\|P_t^*\mu_1 - \mu_2\|_V \leq c\mathrm{e}^{-\lambda t}\|\mu_1 - \mu_2\|_{var}, \quad \forall t \geq t^*,\ \mu_1, \mu_2 \in \mathscr{P}.$$

In general, (B_4) follows from a Hörmander condition. In this spirit, we use the following explicit condition replacing (B_4).

(B_4') $d_1 = d_2 = d$, $\nabla^{(2)}Z^{(1)}$ is invertible with

$$\|\nabla^{(2)}Z^{(1)}\|_\infty + \|(\nabla^{(2)}Z^{(1)})^{-1}\|_\infty < \infty,$$

$\nabla^{(2)}Z^{(1)}$ is Hölder continuous, and $\|\nabla^{(1)}Z^{(1)}\| + \|(\nabla^{(2)})^2 Z^{(1)}\|$ is locally bounded.

TN **Theorem 3.2.** *Assume (B_1)-(B_3) and (B_4'), then (B_4) holds and P_t is t_0-regular for any $t_0 > 0$, so that all assertions of Theorem 3.1 apply.*

Example 3.1. Simply consider $d_1 = d_2 = d$ and that $\sigma_t(y) = I_{d\times d}$ is the identity matrix. We make the following choices of $b, Z^{(1)}$ and $Z^{(2)}$:

- b satisfies $\|b\|_{\tilde{L}^p} < \infty$ for some $p > d$. For example, it is easy to see that this is true when

$$b(x) := \int_{\mathbb{R}^d} \frac{x-y}{|x-y|^{\alpha+1}}\nu(\mathrm{d}y), \quad x \in \mathbb{R}^d$$

for some $\alpha \in (0,1)$ and a finite measure ν on $\mathbb{R}^d$. This type drifts are of interests in statistical physics, see [11] and references therein. Here, we extend the existing study to degenerate setting.

- For some constants c_1, c_2, c_3, δ with $0 < c_1, |c_2|, c_3 < \infty$ and $0 \le \delta < \infty$,

$$Z^{(1)}(x,y) := -c_1(1+|x|)^\delta x + c_2 y, \quad Z^{(2)}(x,y) := Z(x,y) - c_3(1+|y|)^\delta y,$$

where $Z : \mathbb{R}^{2d} \to \mathbb{R}^d$ is a locally Lispchitz continuous with

$$\lim_{|(x,y)|\to\infty} \frac{|Z(x,y)|}{|(x,y)|} = 0.$$

Take, for some constant $\theta \in (0,\infty)$,

$$V(x,y) = (1+|x|^2+|y|^2)^\theta, \quad x, y \in \mathbb{R}^d.$$

Then (B_1), (B_2) and (B_4') hold, so that by Theorem 3.2, we have the following assertions.

(1) When $\delta = 0$ and $|c_2|$ is small enough, we find a constant $0 < c_0 < \infty$ such that (B_3) holds for $\Phi(r) = c_0 r$. Assertions (1)-(3) in Theorem 3.1 imply that P_t has a unique invariant probability measure μ such that $\mu(|\cdot|^2) < \infty$, (3.3) and (3.4) for some constants $0 < c, \lambda < \infty$ hold.

(2) When $\delta > 0$, then (B_3) holds for $\Phi(r) = c_0(1 + r^{1+\delta/(2\theta)})$ for some constant $0 < c_0 < \infty$, so that Theorem 3.1 (4) implies (3.6) for some constants $c, \lambda > 0$.

3.2 Proofs

Proof of Theorem 3.1. Once (1) is proved, (2) follows from Doob's Theorem [4]. So, below we only prove (1), (3) and (4).

(a) We prove the existence of invariant probability measure by using Zvonkin's transform. Let

$$L^0 = \frac{1}{2}\mathrm{tr}\Big(\sigma\sigma^*\nabla^2\Big) + b\cdot\nabla.$$

According to [17, Lemma 2.5], (B_2) implies that there exists $\lambda > 0$ such that the PDE

THT0 (3.7) $$(L^0 - \lambda)u = -b$$

has a unique solution satisfying

THT' (3.8) $$\|u\|_\infty + \|\nabla u\|_\infty < \varepsilon, \quad \|\nabla^2\|_{\tilde{L}^p} < \infty,$$

where ε defined in (B_2).
Then $\Theta(y) := y + u(y),\ y \in \mathbb{R}^{d_2}$, gives rise to a diffeomorphism on $\mathbb{R}^{d_2}$. So,

THT (3.9) $$\tilde{\Theta}(x,y) = (x, \Theta(y)), \quad (x,y) \in \mathbb{R}^{d_1+d_2}$$

is a diffeomorphism on $\mathbb{R}^{d_1+d_2}$.

Let $(X_t^{x,y}, Y_t^{x,y})_{t\ge 0}$ solve (3.1) with initial value $(x,y) \in \mathbb{R}^{d_1+d_2}$, and let

$$\hat{P}_t f(x,y) := \mathbb{E}[(X_t^{x,y}, \Theta(Y_t^{x,y}))], \quad t \ge 0, f \in \mathscr{B}_b(\mathbb{R}^{d_1+d_2}), (x,y) \in \mathbb{R}^{d_1+d_2}.$$

Then $(P_t)_{t\geq 0}$ and $(\hat{P}_t f)_{t\geq 0}$ satisfy

PHT (3.10) $$(\hat{P}_t f)(x,y) = \big(P_t(f\circ\tilde{\Theta})\big)(\tilde{\Theta}^{-1}(x,y)), \quad (x,y)\in\mathbb{R}^{d_1+d_2}.$$

So, μ is an invariant probability measure of P_t if and only if

pm (3.11) $$\hat{\mu} := \mu\circ\tilde{\Theta}^{-1}$$

is $\hat{P}_t$-invariant. Therefore, it is sufficient to show that $\hat{P}_t$ has an invariant probability measure. By the Bogoliov-Krylov theorem, we only need to verify the tightness of

pm1 (3.12) $$\hat{\mu}_n := \frac{1}{n}\int_0^n \hat{P}_s((0,0),\cdot)\mathrm{d}s, \quad n\geq 1.$$

By (B_3) and (3.8), we find a constant $K\in(0,\infty)$ such that

$$|V(x,\Theta(y)) - V(x,y)| \leq K\big\{V(x,\Theta(y))\wedge V(x,y)\big\}, \quad (x,y)\in\mathbb{R}^{d_1+d_2},$$

so that for $r_0 := \frac{1}{1+K}\in(0,1)$,

W2' (3.13) $$\gamma_0(V\circ\tilde{\Theta})(x,y) \leq V(x,y) \leq \gamma_0^{-1}(V\circ\tilde{\Theta})(x,y).$$

Combining these with (B_3), (3.7) and applying Itô's formula, we find a constant $0<c_1<\infty$ such that

W3' (3.14) $$\mathrm{d}(V\circ\tilde{\Theta})(X_t,Y_t) \leq \Big(K - c_1\Phi\big(\gamma_0(V\circ\tilde{\Theta})(X_t,Y_t)\big)\Big)\mathrm{d}t + \mathrm{d}M_t$$

for some martingale $(M_t)_{t\geq 0}$. Letting $(X_0,Y_0)=(0,0)$, we deduce from (3.14) that

$$\begin{aligned}&\int_{\mathbb{R}^{d_1+d_2}} \Phi\big(\gamma_0(V\circ\tilde{\Theta})\big)\mathrm{d}\hat{\mu}_n = \frac{1}{n}\int_0^n \mathbb{E}\Big[\Phi\big(\gamma_0(V\circ\tilde{\Theta})(X_t^{(0,0)},Y_t^{(0,0)})\big)\Big]\mathrm{d}t\\ &\leq \frac{K+\ V\circ\tilde{\Theta}(0)/n}{c_1} < \infty, \quad n\geq 1.\end{aligned}$$

Since $\Phi\big(\gamma_0(V\circ\tilde{\Theta})\big)$ has compact level sets, this implies the tightness of $\{\hat{\mu}_n\}_{n\geq 1}$, and the weak limit $\hat{\mu}$ of a convergent subsequence gives an invariant probability measure of $\hat{P}_t$. Moreover,

$$\int_{\mathbb{R}^{d_1+d_2}} \Phi\big(\gamma_0(V\circ\tilde{\Theta})\big)\mathrm{d}\hat{\mu} \leq \frac{K}{c_1} < \infty.$$

Therefore, by (3.11) and (3.8), $\mu := \hat{\mu}\circ\tilde{\Theta}$ is an invariant probability measure of P_t and $\mu(\Phi(\varepsilon_0 V)) < \infty$ holds for some constant $0<\varepsilon_0<\infty$.

(b) In the situations of (3) and (4), we have $\Phi(r)\geq c_0 r$ for some constant $0<c_0<\infty$ and all $r\geq 1$, so that (1) implies that $\mu(V)<\infty$. By (3.10) and the definition of weighted total variation norm, we obtain

$$\|P_t^*\mu_1 - P_t^*\mu_2\|_V = \|\hat{P}_t^*(\mu_1\circ\tilde{\Theta}^{-1}) - \hat{P}_t^*(\mu_2\circ\tilde{\Theta}^{-1})\|_{V\circ\tilde{\Theta}^{-1}},$$

$$|P_t^*\mu_1-\mu\|_V=\|\hat{P}_t^*(\mu_1\circ\tilde{\Theta}^{-1})-\hat{\mu}\|_{V\circ\tilde{\Theta}^{-1}},\quad \mu_1,\mu_2\in\mathscr{P}_V,$$

where μ and $\hat{\mu}$ are the above constructed invariant probability measures of $(P_t)_{t\geq 0}$ and $(\hat{P}_t)_{t\geq 0}$, respectively. Combining this with (3.13), we derive

$$\begin{aligned}\|P_t^*\mu_1-P_t^*\mu_2\|_V&\leq\gamma_0^{-1}\|\hat{P}_t^*(\mu_1\circ\tilde{\Theta}^{-1})-\hat{P}_t^*(\mu_2\circ\tilde{\Theta}^{-1})\|_V,\\ \|P_t^*\mu_1-\mu\|_V&\leq\gamma_0^{-1}\|\hat{P}_t^*(\mu_1\circ\tilde{\Theta}^{-1})-\hat{\mu}\|_V,\quad \mu_1,\mu_2\in\mathscr{P}_V.\end{aligned}$$

So, it remains to verify (3.4) and (3.5) for $\hat{P}_t^*$ replacing P_t^*.

By (3.14) and $\Phi(r)\geq c_0 r$, we have

$$\mathrm{d}V(x,\Theta(y))\leq\big(K-c_1c_0\gamma_0V(x,\Theta(y))\big)\mathrm{d}t+\mathrm{d}M_t.$$

By invoking (B_4) and (3.10), any compact subset of $\mathbb{R}^{d_1+d_2}$ is $\hat{P}_t$-petite. Furthermore, from (3.14) we deduce that $(\hat{P}_t^*)_{t\geq 0}$ admits the Lyapunov condition: for some constant $0<k_1,\ k_2=c_1c_0\gamma_0<\infty$ and for any $(x,y)\in\mathbb{R}^{d_1+d_2}$,

$$\hat{P}_tV(x,\Theta(y))=\mathbb{E}\big(V(x,\Theta(y))\big)\leq\frac{k_1}{k_2}+\mathrm{e}^{-k_2t}V(x,\Theta(y)).$$

Consequently, the Harris theorem [9, Theorem 1.3] and (3.13) yield that there exist constants $0<c,\lambda<\infty$ such that

$$\|\hat{P}_t^*\delta_{(x,y)}-\hat{\mu}\|_V\leq\mathrm{e}^{-\lambda t}\|\delta_{(x,y)}-\hat{\mu}\|_V\leq c\mathrm{e}^{-\lambda t}V(x,y),\quad t\geq 0.$$

Thus, by following the line of part (c) in the proof of [17, Theorem 2.1], we find constants $c,\lambda>0$ such that

W6 (3.15) $$\|\hat{P}_t^*\mu_1-\hat{P}_t^*\mu_2\|_V\leq c\mathrm{e}^{-\lambda t}\|\mu_1-\mu_2\|_V,\quad \mu_1,\mu_2\in\mathscr{P}_V.$$

This immediately implies (3.4) for $\hat{P}_t^*$ replacing P_t^*.

Next, by (3.15), the semigroup property of $(\hat{P}_t)_{t\geq 0}$ and the invariance of $\hat{\mu}$, we have

$$\|\hat{P}_t^*\delta_{(x,y)}-\hat{\mu}\|_V=\|\hat{P}_{t/2}^*\hat{P}_{t/2}^*\delta_{(x,y)}-\hat{P}_{t/2}^*\hat{\mu}\|_V\leq c\mathrm{e}^{-\lambda t/2}\|\hat{P}_{t/2}^*\delta_{(x,y)}-\hat{\mu}\|_V.$$

On the other hand, by the proof of [17, (2.35)], (B_4) implies

$$\hat{P}_{2/t}V(x,y)\leq c_2\Big(1+H^{-1}(H(V(x,y))-\frac{t}{2c_2}\Big),\quad 0<c_2<\infty$$

so that

$$\begin{aligned}\|\hat{P}_t^*\delta_{(x,y)}-\hat{\mu}\|_V&\leq c\mathrm{e}^{\frac{-\lambda t}{2}}\|\hat{P}_{2/t}^*\delta_{(x,y)}-\hat{\mu}\|_V\leq c_3\mathrm{e}^{\frac{-\lambda t}{2}}\big(\hat{P}_{2/t}^*\delta_{(x,y)}(V)+\hat{\mu}(V)\big)\\ &\leq c_4\Big(1+H^{-1}(H(V(x,y))-\frac{t}{2c_2}\Big)\mathrm{e}^{-\lambda t},0<c_3,c_4<\infty.\end{aligned}$$

Therefore, (3.5) holds for $\hat{P}_t^*$ replacing P_t^*. □

Proof of Theorem 3.2. By Proposition 3.3 below, (B_1)-(B_3) and (B_4') imply (B_4) for any $t_0 > 0$ and $\nu(\mathrm{d}x, \mathrm{d}y) := \inf_{(x',y'),(x'',y'')\in D} p_{t_0}(x, y; x', y')1_D(x, y)\mathrm{d}x\mathrm{d}y$. . □

In the following Proposition 3.3, (B_3) is weakened as

(B_3') There exist constants $\varepsilon \in (0, 1), K > 0$, and $1 \le V \in C^2(\mathbb{R}^{d_1+d_2})$ with

LY1 (3.16) $$\lim_{|x|+|y|\to\infty} V(x, y) = \infty, \quad \limsup_{|x|+|y|\to\infty} \sup_{y'\in B_\varepsilon(y)} \frac{|\nabla V(x, \cdot)(y')| + \|(\nabla^2)^2 V(x, y')\|}{V(x, y)} < \infty,$$

such that

$$\begin{aligned}&\varepsilon \sup_{y'\in B_\varepsilon(y)} \big(|Z^{(1)}(x, y)|\|\nabla^{(1)}\nabla^{(2)}V(x, y')\| + |Z^{(2)}(x, y)|(|\nabla^{(2)}V(x, y')| + \|(\nabla^2)^2 V(x, y')\|)\big)\\ &+ \langle Z^{(1)}(x, y), \nabla V(\cdot, y)(x)\rangle + \langle Z^{(2)}(x, y), \nabla V(x, \cdot)(y)\rangle \le KV(x, y).\end{aligned}$$

P1 **Proposition 3.3.** *Assume that $(B_1), (B_2), (B_3')$ and (B_4') hold. If σ is Hölder continuous, then P_t has a heat kernel $p_t(x, y; x', y')$ with respect to the Lebesgue measure such that*

$$\inf_{(x,y),(x',y')\in B_k(0)} p_t(x, y; x', y') > 0, \ \forall t, k > 0.$$

To prove this result, we apply the Harnack inequality presented in [5] for the PDE

PDE'' (3.17) $$\begin{aligned}\partial_t f_t(x, y) =& -y \cdot \nabla^{(1)} f_t(x, y) - \big(Z^{(2)} \cdot \nabla^{(2)} f_t\big)(x, y)\\ &+ \mathrm{div}^{(2)}\big(a\nabla^{(2)} f_t\big)(x, y) + (Uf_t)(x, y), \quad t \ge 0, x, y \in \mathbb{R}^d,\end{aligned}$$

where $\mathrm{div}^{(2)}$ is the divergence operator in the second component y, and

$$Z^{(2)} : \mathbb{R}^{2d} \to \mathbb{R}^d, \quad U : \mathbb{R}^{2d} \to \mathbb{R}, \quad a : \mathbb{R}^{2d} \to \mathbb{R}^{d\otimes d}$$

are measurable satisfying the following assumption.

(B_2') a is invertible, $\nabla^{(2)}a$ exists, such that $|Z^{(2)}| + |U| + \|a\| + \|a^{-1}\|$ is locally bounded in $\mathbb{R}^{2d}$.

The following Harnack inequality is essentially due to [5].

LN1 **Lemma 3.4.** *Assume (B_2'). Then there exists a constant $r_0 \in (0, 1)$ such that for any $t > 0$ and $r \in (0, r_0]$, there exists a locally bounded function*

$$\varphi_{t,r} : \mathbb{R}^{2d} \times \mathbb{R}^{2d} \to (0, \infty)$$

such that any positive weak solution (in the sense of integration by parts) f_t of (3.17) satisfies

$$f_t(x', y') \le \varphi_{t,r}(x, y; x', y')f_{t+r}(x, y), \quad (x, y), (x', y') \in \mathbb{R}^{2d}.$$

Proof. By [5, Theorem 3 and Remark 4], there exist small constants $r_0, r_1 \in (0,1)$, such that for any $t > 0$ there exists a locally bounded function

$$C_t : (0, r_0] \times \mathbb{R}^{2d} \to (0, \infty)$$

such that any positive solution f of (3.17) satisfies

$$f_t(x,y) \le C_{t,r}(x,y) \inf_{(x',y') \in B(x,y;r_1)} f_{t+r}(x',y'), \quad r \in (0, r_0],$$

where $B(x,y;r_1) := \{(x',y') \in \mathbb{R}^{2d} : \ |(x-x', y-y')| \le r_1\}$. For any $(x,y), (x',y') \in \mathbb{R}^{2d}$, let

$$n = n(x,y;x',y') := \inf\big\{n \in \mathbb{N} : \ |(x-x', y-y')| \le n r_1\big\},$$

and denote

$$(x_i, y_i) := (x,y) + \frac{i}{n}(x'-x, y'-y), \quad 0 \le i \le n.$$

Then

$$f_t(x,y) \le C_{t,\frac{r}{n}}(x_0,y_0) f_{t+\frac{r}{n}}(x_1,y_1) \le \cdots \le f_{t+r}(x',y') \prod_{i=0}^{n-1} C_{t+\frac{ir}{n}, \frac{r}{n}}(x_i, y_i).$$

Therefore, the desired estimate holds for

$$\varphi_{t,r}(x,y;x',y') := \prod_{i=0}^{n-1} C_{t+\frac{ir}{n}, \frac{r}{n}}(x_i, y_i), \quad n := n(x,y;x',y').$$

□

Next, we extend Lemma 3.4 to the following more general PDE:

PDE' (3.18)
$$\begin{aligned} \partial_t f_t(x,y) = &- \big(Z^{(1)} \cdot \nabla^{(1)} f_t + Z^{(2)} \cdot \nabla^{(2)} f_t\big)(x,y) \\ &+ \mathrm{div}^{(2)}\big(a \nabla^{(2)} f_t\big)(x,y) + (U f_t)(x,y), \quad t \ge 0, \ x, y \in \mathbb{R}^d. \end{aligned}$$

LN2 **Lemma 3.5.** *Assume (B_2') and (B_4'). Then there exists a constant $r_0 \in (0,1)$ such that for any $t > 0$ and $r \in (0, r_0]$, there exists a locally bounded function*

$$\varphi_{t,r} : \mathbb{R}^{2d} \times \mathbb{R}^{2d} \to (0, \infty)$$

such that any positive weak solution (in the sense of integration by parts) f_t of (3.18) satisfies

$$f_t(x',y') \le \varphi_{t,r}(x,y;x',y') f_{t+r}(x,y), \quad (x,y), (x',y') \in \mathbb{R}^{2d}.$$

Proof. To transform (3.18) into (3.17), we make the change of variable

$$(x,y) \mapsto (x, \tilde{y}) := \big(x, (Z^{(1)}(x,y))^{-1}\big).$$

Let $\phi(x,\cdot) := \big(Z^{(1)}(x,\cdot)\big)^{-1}$ and

$$\tilde{f}_t(x,\tilde{y}) := f_t(x, \phi(x,\tilde{y})), \quad \tilde{U}(x,\tilde{y}) := U(x, \phi(x,\tilde{y})), \quad t \ge 0, x, \tilde{y} \in \mathbb{R}^d.$$

Then (3.18) implies

R1 (3.19)
$$\begin{aligned}\partial_t \tilde{f}_t(x,\tilde{y}) = &-\Big(Z^{(1)}\cdot\nabla^{(1)} f_t + Z^{(2)}\cdot\nabla^{(2)} f_t\Big)\big(x,\phi(x,\tilde{y})\big)\\ &+\mathrm{div}^{(2)}\Big(a\nabla^{(2)} f_t\Big)\big(x,\phi(x,\tilde{y})\big) + (Uf_t)\big(x,\phi(x,\tilde{y})\big),\quad t\ge 0,\ x,y\in\mathbb{R}^d.\end{aligned}$$

By chain rule, we obtain

$$\begin{aligned}\Big(\nabla^{(1)}\tilde{f}_t\Big)(x,\tilde{y}) &= (\nabla^{(1)} f_t)\big(x,\phi(x,\tilde{y})\big) + \Big(\nabla^{(1)}\phi(x,\tilde{y})\Big)\big(\nabla^{(2)} f_t\big)\big(x,\phi(x,\tilde{y})\big),\\ \Big(\nabla^{(2)}\tilde{f}_t\Big)(x,\tilde{y}) &= \Big(\nabla^{(2)}\phi(x,\tilde{y})\Big)\big(\nabla^{(2)} f_t\big)\big(x,\phi(x,\tilde{y})\big),\\ \Big(\big(\nabla^{(2)}\big)^2\tilde{f}_t\Big)(x,\tilde{y}) &= \Big(\nabla^{(2)}\phi(x,\tilde{y})\Big)^2\Big(\big(\nabla^{(2)}\big)^2 f_t\Big)\big(x,\phi(x,\tilde{y})\big)\\ &\quad+\Big((\nabla^{(2)})^2\phi(x,\tilde{y})\Big)\big(\nabla^{(2)} f_t\big)\big(x,\phi(x,\tilde{y})\big).\end{aligned}$$

So,

R2 (3.20)
$$\begin{aligned}\big(\nabla^{(2)} f_t\big)\big(x,\phi(x,\tilde{y})\big) &= \Big(\nabla^{(2)}\phi(x,\tilde{y})\Big)^{-1}\big(\nabla^{(2)}\tilde{f}_t\big)(x,\tilde{y}),\\ \big(\nabla^{(1)} f_t\big)\big(x,\phi(x,\tilde{y})\big) &= \big(\nabla^{(1)}\tilde{f}_t\big)(x,\tilde{y}) - \Big(\nabla^{(1)}\phi(x,\tilde{y})\Big)\Big(\nabla^{(2)}\phi(x,\tilde{y})\Big)^{-1}\big(\nabla^{(2)}\tilde{f}_t\big)(x,\tilde{y}).\end{aligned}$$

So, letting

$$A^{x,\tilde{y}} := \Big(\nabla^{(2)}\phi(x,\tilde{y})\Big)^{-1},$$

we derive

R3 (3.21)
$$\begin{aligned}\mathrm{div}^{(2)}\big(a\nabla^{(2)} f_t\big)\big(x,\phi(x,\tilde{y})\big) &= \sum_{i,j=1}^d \big[\partial_i^{(2)}\big(a_{ij}\partial_j^{(2)} f_t\big)\big](x,\phi(x,\tilde{y}))\\ &= \sum_{i,j,k=1}^d A_{ik}^{x,\tilde{y}}\partial_k^{(2)}\big((a_{ij}\partial_j^{(2)} f_t)(x,\phi(x,\tilde{y}))\big)\\ &= \sum_{i,j,k,l=1}^d A_{ik}^{x,\tilde{y}}\partial_k^{(2)}\big(a_{ij}(x,\phi(x,\tilde{y}))A_{jl}^{x,\tilde{y}}\partial_l^{(2)}\tilde{f}_t(x,\tilde{y})\big)\\ &= \sum_{k,l=1}^d \partial_k^{(2)}\Big(\big[(A^{x,\tilde{y}})^* a(x,\phi(x,\tilde{y}))A^{x,\tilde{y}}\big]_{kl}\partial_l^{(2)}\tilde{f}_t(x,\tilde{y})\Big)\\ &\quad-\sum_{i,j,k,l=1}^d \big(\partial_k^{(2)}A_{ik}^{x,\tilde{y}}\big)a_{ij}(x,\phi(x,\tilde{y}))A_{jl}^{x,\tilde{y}}\partial_l^{(2)}\tilde{f}_t(x,\tilde{y})\\ &= \mathrm{div}^{(2)}\big(\tilde{a}\nabla^{(2)}\tilde{f}_t\big)(x,\tilde{y}) - \big(\hat{Z}\cdot\nabla^{(2)}\tilde{f}_t\big)(x,\tilde{y}),\end{aligned}$$

where

$$\tilde{a}(x,\tilde{y}) := \big(A^{x,\tilde{y}}\big)^* a(x,\phi(x,\tilde{y}))A^{x,\tilde{y}},\quad \hat{Z}_l := \sum_{i,j,k=1}^d \big(\partial_k^{(2)}A_{ik}^{x,\tilde{y}}\big)a_{ij}(x,\phi(x,\tilde{y}))A_{jl}^{x,\tilde{y}},\ 1\le l\le d.$$

Substituting (3.20) and (3.21) into (3.19), and noting that $Z_1(x,\phi(x,\tilde{y}))=\tilde{y}$, we derive

$$\begin{aligned}\partial_t\tilde{f}_t(x,\tilde{y}) =& -\big(\tilde{y}\cdot\nabla^{(1)}\tilde{f}_t+\tilde{Z}^{(2)}\cdot\nabla^{(2)}\tilde{f}_t\big)(x,\tilde{y})\\ &+\mathrm{div}^{(2)}\big(\tilde{a}\nabla^{(2)}f_t\big)(x,\tilde{y})+(\tilde{U}\tilde{f}_t)(x,y),\quad t\geq 0,\ x,y\in\mathbb{R}^d,\end{aligned}$$

for the above defined $\tilde{U}$, $\tilde{a}$ and

$$\tilde{Z}^{(2)}(x,\tilde{y}):=\hat{Z}(x,\tilde{y})+\big(\nabla^{(1)}\phi(x,\tilde{y})\big)^*\tilde{y}+\big(\nabla^{(2)}\phi(x,\tilde{y})\big)^*Z^{(2)}(x,\phi(x,\tilde{y})),\quad x,\tilde{y}\in\mathbb{R}^d.$$

Combining this with (B_2') and (B_4'), we may apply Lemma 3.4 to this PDE to derive the desired estimate. □

We also need the following result for the existence of heat kernel.

LN3 **Lemma 3.6.** *Assume that $(B_1),(B_2)$ and (B_4') hold, and the solution to (3.1) is non-explosive. Then (3.1) has heat kernel p_t; namely, for any $t>0$ and the solution (X_t,Y_t) starting at (x_0,y_0), the distribution of (X_t,Y_t) has a density $p_t(x_0,y_0;\cdot)$ with respect to the Lebesgue measure.*

Proof. (a) We first assume $b=0$ and $Z^{(1)}(x,y)=y$, but allow σ also depends on x such that σ is Hölder continuous and $\|\sigma^*\|_\infty+\|(\sigma\sigma^*)^{-1}\|_\infty<\infty$. When $Z^{(2)}$ is bounded, the existence of heat kernel follows from [12, Theorem 1.5]. In general, for any $n\geq 1$ let

$$Z^{2,n}(x,y):=Z^{(2)}(\varphi_n(x),\varphi_n(y)),\quad \varphi_n(x):=x1_{\{|x|\leq n\}}+\frac{nx}{|x|}1_{\{|x|>n\}}.$$

Let $(X_t^{(n)},Y_t^{(n)})$ solve the SDE

$$\begin{cases}\mathrm{d}X_t^{(n)}=Y_t^{(n)}dt,\\ \mathrm{d}Y_t^{(n)}=Z^{2,n}(X_t^{(n)},Y_t^{(n)})\mathrm{d}t+\sigma(Y_t)\mathrm{d}W_t,\quad (X_0,Y_0)=(x_0,y_0).\end{cases}$$

Then for any $t>0$ and $n\geq 1$, the distribution of $(X_t^{(n)},Y_t^{(n)})$ is absolutely continuous with respect to the Lebesgue measure; i.e. for any null set $A\subset\mathbb{R}^{2d}$, $\mathbb{P}((X_t^{(n)},Y_t^{(n)})\in A)=0$. Letting

$$\tau_n:=\inf\{t\geq 0:|X_t|\vee|Y_t|\geq n\},$$

we have $(X_t,Y_t)=(X_t^n,Y_t^n)$ for $t\leq\tau_n$. By the non-explosion we have $\tau_n\to\infty$ as $n\to\infty$, so that

$$\mathbb{P}((X_t,Y_t)\in A)\leq\lim_{n\to\infty}\big\{\mathbb{P}((X_t^{(n)},Y_t^{(n)})\in A)+\mathbb{P}(\tau_n<t)\big\}=0$$

holds for all null set A. Thus, the heat kernel exists.

(b) Let $b=0$ and for $Z^{(1)}$ satisfying (B_4'). As shown in the proof of Lemma 3.5, with the transform $(x,y)\mapsto\big(x,(Z^{(1)}(x,y))^{-1}\big)$ we reduce the situation (a), so that the heat kernel exists. Finally, when $\|b\|_{\tilde{L}^p}<\infty$ for some $p>d$, by [17, Lemma 2.5], when $\lambda>0$ is large enough, the PDE

$$\frac{1}{2}\mathrm{tr}\big\{\sigma\sigma^*\nabla^2\big\}u+b\cdot\nabla u=-b+\lambda u$$

for $u:\mathbb{R}^d\to\mathbb{R}^d$ has a unique solution such that

ASS (3.22) $$\|\nabla^2 u\|_{\tilde{L}^p}<\infty,\quad \|\nabla u\|_\infty\le\frac{1}{2}.$$

Moreover, by the Sobolev embedding theorem, $\|\nabla^2 u\|_{\tilde{L}^p}<\infty$ implies that ∇u is Hölder continuous. By Itô's formula, we see that $(X_t,\tilde{Y}_t)=(X_t,\Theta(Y_t))$ solves the SDE

$$\begin{cases}\mathrm{d}X_t=\tilde{Z}^{(1)}(X_t,\tilde{Y}_t)\mathrm{d}t,\\ \mathrm{d}\tilde{Y}_t=\tilde{Z}^{(2)}(X_t,\tilde{Y}_t)\mathrm{d}t+\tilde{\sigma}(Y_t)\mathrm{d}W_t,\quad (X_0,\tilde{Y}_0)=(x_0,\Theta(y_0)),\end{cases}$$

where

$$\begin{aligned}&\tilde{Z}^{(1)}(x,\tilde{y}):=Z^{(1)}(x,\Theta^{-1}(\tilde{y})),\\ &\tilde{Z}^{(2)}(x,\tilde{y}):=\big((\nabla\Theta)Z^{(2)}(x,\cdot)\big)(\Theta^{-1}(\tilde{y}))+\lambda u(\Theta^{-1}(\tilde{y})),\\ &\tilde{\sigma}(\tilde{y}):=\big((\nabla\Theta)\sigma\big)(\Theta^{-1}(\tilde{y})),\quad x,\tilde{y}\in\mathbb{R}^d.\end{aligned}$$

Thus, $(X_t,\tilde{Y}_t):=(X_t,\Theta(Y_t))$ solves the SDE of type (3.1) with $b=0$, so that by Step (a) it has a heat kernel. By (3.22), this implies that (X_t,Y_t) has heat kernel as well. □

Proof of Proposition 3.3. By Lemma 3.6, for any $t>0$, (X_t,Y_t) has a distribution density function (heat kernel) $p_t(x_0,y_0;\cdot)$. Let $a=\frac{1}{2}\sigma\sigma^*$, so that

$$\frac{1}{2}\mathrm{tr}\big(\sigma\sigma^*\nabla^{(2)}\big)f=\mathrm{div}^{(2)}\big(a\nabla^{(2)}f\big)+\big(\mathrm{div}^{(2)}a\big)\cdot\nabla^{(2)}f,\quad f\in C^2$$

where $(\mathrm{div}^{(2)}a)_i:=\sum_{j=1}^d\partial_j^{(2)}a_{ij}$. As shown in Step (b) in the proof of Lemma 3.6, with Zvonkin's transform $(X_t,\tilde{Y}_t):=(X_t,\theta(Y_t))$ for $b+\mathrm{div}^{(2)}a$ replacing b, we may and do assume that $b+\mathrm{div}^{(2)}a=0$, so that the generator of (X_t,Y_t) becomes

$$Lf:=\mathrm{div}^{(2)}\big(a\nabla^{(2)}f\big)+Z^{(1)}\cdot\nabla^{(1)}f+Z^{(2)}\cdot\nabla^{(2)}f.$$

It is easy to see that the adjoint operator of L in $L^2(\mathbb{R}^{2d})$ is

$$L^*f=\mathrm{div}^{(2)}\big(a\nabla^{(2)}f\big)-Z^{(1)}\cdot\nabla^{(1)}f-Z^{(2)}\cdot\nabla^{(2)}f-\big(\mathrm{div}^{(1)}Z^{(1)}+\mathrm{div}^{(2)}Z^{(2)}\big)f,$$

Hence, that the heat kernel $f_t:=p_t(x_0,y_0;\cdot)$ solves the equation (3.18) for

$$U:=\big(\mathrm{div}^{(1)}Z^{(1)}+\mathrm{div}^{(2)}Z^{(2)}\big).$$

So, by Lemma 3.5 with $\beta=\beta_0\wedge\frac{t}{2}$, we derive

$$p_t(x_0,y_0;x,y)\ge C_t(x,y;x',y')p_{t-\beta}(x_0,y_0;x',y'),\quad (x,y),(x',y')\in\mathbb{R}^{2d}$$

for some function $C_t:\mathbb{R}^{2d}\times\mathbb{R}^{2d}\to(0,\infty)$ satisfying

$$c_t(\gamma):=\inf_{(x,y),(x',y')\in B_\gamma(0)}c_t(x,y;x',y')>0.$$

Let $k>0$ be a constant. For any $\gamma\geq k$, we obtain

$$\begin{aligned}&\inf_{(x,y)\in B_k(0)} p_t(x_0,y_0;x,y)\geq c_t(\gamma)\sup_{(x',y')\in B_\gamma(0)} p_{t-\beta}(x_0,y_0;x',y')\\&\geq \tilde{c}_t(\gamma)\int_{B_N(0)} p_{t-\beta}(x_0,y_0;x',y')\mathrm{d}x'\mathrm{d}y'=\tilde{c}_t(\gamma)\mathbb{P}\big((X_t,Y_t)\in B_\gamma(0)\big),\end{aligned}$$

where $\tilde{c}_t(\gamma):=\frac{c_t(N)}{|B_\gamma(0)|}>0$, for $|B_\gamma(0)|$ the volume of $B_\gamma(0)$. On the other hand, by Theorem 2.1, $(B_1),(B_2)$ and (B_3') imply (2.2), so that for $k_\gamma:=\inf_{(x,y)\notin B_\gamma(0)}U(x,y)$,

$$\mathbb{P}\big((X_t,Y_t)\notin B_\gamma(0)\big)\leq\mathbb{P}\big(U(X_t,Y_t)\geq k_\gamma\big)\leq\frac{c}{k_\gamma}U(x_0,y_0).$$

Taking large enough N such that $\frac{c}{k_\gamma}\sup_{(x_0,y_0)\in K}U(x_0,y_0)\leq\frac{1}{2}$, we obtain

$$\inf_{(x_0,y_0),(x,y)\in K} p_t(x_0,y_0;x,y)\geq\tilde{c}_t(\gamma)\inf_{(x_0,y_0)\in K}\Big(1-\frac{c}{k_\gamma}U(x_0,y_0)\Big)\geq\frac{1}{2}\tilde{c}_t(\gamma)>0.$$

So, the desired assertion holds.

□

4 Extension to McKean-Vlasov SDEs

In this part we extend Theorem 2.1 and Theorem 3.2 to McKean-Vlasov SDEs. Consider the following distribution dependent SDE on $\mathbb{R}^{d_1+d_2}$:

E41 (4.1)
$$\begin{cases}\mathrm{d}X_t=Z_t^{(1)}(X_t,Y_t)\mathrm{d}t,\\ \mathrm{d}Y_t=\big(Z_t^{(2)}(X_t,Y_t,\mathscr{L}_{(X_t,Y_t)})+b_t(Y_t)\big)\mathrm{d}t+\sigma(Y_t)\mathrm{d}W_t,\quad t\in[0,T],\end{cases}$$

where $\mathscr{L}_{(X_t,Y_t)}\in\mathscr{P},\ t\in[0,T]$ is the law of (X_t,Y_t), $Z^{(1)},b,\sigma$ and W are as in (1.1), and

$$Z^{(2)}:[0,T]\times\mathbb{R}^{d_1+d_2}\times\mathscr{P}\to\mathbb{R}^{d_2}$$

is measurable. We first study the well-posedness of (4.1), then investigate the uniform ergodicity for the time-homogeneous model

EQ4.1 (4.2)
$$\begin{cases}\mathrm{d}X_t=Z^{(1)}(X_t,Y_t)\mathrm{d}t,\\ \mathrm{d}Y_t=\big(Z^{(2)}(X_t,Y_t,\mathscr{L}_{(X_t,Y_t)})+b(Y_t)\big)\mathrm{d}t+\sigma(Y_t)\mathrm{d}W_t,\quad t\geq0.\end{cases}$$

4.1 Well-posedness of (4.1)

Let $\hat{\mathscr{P}}$ be a sub-space of $\mathscr{P}$. We call (4.1) well-posed for distributions in $\hat{\mathscr{P}}$, if for any $\mathscr{F}_0$-measurable random variable (X_0,Y_0) with $\mathscr{L}_{(X_0,Y_0)}\in\hat{\mathscr{P}}$ (respectively, any initial distribution $\gamma\in\hat{\mathscr{P}}$), (4.1) has a unique strong solution (respectively, unique weak solution) (X_t,Y_t) such that

$$[0,T]\ni t\to\mathscr{L}_{(X_t,Y_t)}\in\hat{\mathscr{P}}$$

is continuous in the weak topology.

To extend Theorem 2.1, let δ_0 be the Dirac measure at $0 \in \mathbb{R}^{d_1+d_2}$ and denote

$$Z_t^{(2)}(x,y) := Z_t^{(2)}(x,y,\delta_0), \quad t \in [0,T], (x,y) \in \mathbb{R}^{d_1+d_2}.$$

Let

$$C_b^w([0,T];\hat{\mathscr{P}}) := \big\{\mu \in [0,T] \to \hat{\mathscr{P}} \text{ is weakly continous, } \sup_{t\in[0,T]} \mu_t(V) < \infty\big\}.$$

($\hat{A}$) (A_1)-(A_3) hold for the above defined $Z_t^{(2)}(x,y)$. Moreover, for any $n \geq 1$ and any $\mu \in C_b^w([0,T];\hat{\mathscr{P}})$, there exists a constant $K_{n,\mu} > 0$ such that

$$|Z_t^{(2)}(x,y,\mu_t) - Z_t^{(2)}(x',y',\mu_t)| \leq K_{n,\mu}|(x-x',y-y')|, \quad (x,y),(x',y') \in B_n(0).$$

The following result extends Theorem 2.1 to the distribution dependent setting as well as [13, Theorem 1.1] to the present degenerate case.

T4.1 **Theorem 4.1.** *Assume* ($\hat{A}$) *for* $\hat{\mathscr{P}} = \mathscr{P}$ *or* $\hat{\mathscr{P}} = \mathscr{P}_V$.

(1) *If* $\hat{\mathscr{P}} = \mathscr{P}$, *and there exists* $0 \leq K \in L^2([0,T])$ *such that*

CD (4.3) $$\big|Z_t^{(2)}(x,y,\mu) - Z_t^{(2)}(x,y,\nu)\big| \leq K_t\|\mu-\nu\|_{var}, \quad \mu,\nu \in \mathscr{P},$$

then (4.1) *is well-posed for distributions in* $\mathscr{P}$.

(2) *Let* $\hat{\mathscr{P}} = \mathscr{P}_V$. *If*

CD' (4.4) $$\begin{aligned} &|Z_t^{(2)}(x,y,\mu) - Z_t^{(2)}(x,y,\nu)| \sup_{y'\in B_\varepsilon(y)} \big(1 + |\nabla^{(2)}V(x,y')| + \|(\nabla^{(2)})^2V(x,y')\|\big) \\ &\leq K_t\|\mu-\nu\|_V, \quad (x,y) \in \mathbb{R}^{d_1+d_2}, \ \mu,\nu \in \mathscr{P}_V, \end{aligned}$$

then (4.1) *is well-posed for distributions in* $\mathscr{P}_V$.

Proof. Let (X_0,Y_0) be $\mathscr{F}_0$-measurable with $\gamma := \mathscr{L}_{(X_0,Y_0)} \in \hat{\mathscr{P}}$. For any

$$\mu \in \mathscr{C}^\gamma := \{\mu \in C_b^w([0,T];\hat{\mathscr{P}}) : \mu_0 = \gamma\},$$

assumption ($\hat{A}$) together with (4.3) or (4.4) implies (A_1)-(A_3) for $Z_t^{(2)}(x,y,\mu_t)$ replacing $Z_t^{(2)}$, so that by Theorem 2.1, the SDE

DE (4.5) $$\begin{cases} \mathrm{d}X_t^\mu = Z_t^{(1)}(X_t^\mu, Y_t^\mu)\mathrm{d}t, \\ \mathrm{d}Y_t^\mu = \{Z_t^{(2)}(X_t^\mu, Y_t^\mu, \mu_t) + b_t(Y_t^\mu)\}\mathrm{d}t + \sigma(Y_t^\mu)\mathrm{d}W_t, \quad t \in [0,T], \end{cases}$$

is well-posed, where $(X_0^\mu, Y_0^\mu) = (X_0, Y_0)$. By [10, Theorem 3.1], it suffices to show that

$$\Psi : \mathscr{C}^\gamma \to \mathscr{C}^\gamma, \quad \Psi_t(\mu) := \mathscr{L}_{(X_t^\mu, Y_t^\mu)}$$

has a unique fixed point in $\mathscr{C}^\gamma$. Below we prove this for $\hat{\mathscr{P}} = \mathscr{P}$ and $\hat{\mathscr{P}} = \mathscr{P}_V$ respectively.

(1) Let $\hat{\mathscr{P}} = \mathscr{P}$ and (4.3) holds. For $\mu, \nu \in \mathscr{C}^\gamma$, we reformulate (4.5) as

DE' (4.6)
$$\begin{cases} \mathrm{d}X_t^\mu = Z_t^{(1)}(X_t^\mu, Y_t^\mu)\mathrm{d}t, \\ \mathrm{d}Y_t^\mu = \big(Z_t^{(2)}(X_t^\mu, Y_t^\mu, \nu_t) + b_t(Y_t^\mu)\big)\mathrm{d}t + \sigma(Y_t^\mu)\mathrm{d}\tilde{W}_t, \quad t \in [0,T], \end{cases}$$

where $(X_0^\mu, Y_0^\mu) = (X_0, Y_0)$ and

$$\begin{aligned} \tilde{W}_t &:= W_t - \int_0^t \xi_s^{\mu,\nu}\mathrm{d}s, \\ \xi_s^{\mu,\nu} &:= \big(\sigma_s^*(\sigma_s\sigma_s^*)^{-1}\big)(Y_s^\mu)\big(Z_s^{(2)}(X_s^\mu, Y_s^\mu, \nu_s) - Z_s^{(2)}(X_s^\mu, Y_s^\mu, \mu_s)\big). \end{aligned}$$

By (A_2) and (4.3), there exists a constant $0 < c_1 < \infty$ such that

XIS (4.7)
$$|\xi_s^{\mu,\nu}|^2 \le c_1 K_s^2 \|\mu_s - \nu_s\|_{var}^2, \quad s \in [0,T].$$

Since $[0,T] \ni s \mapsto \|\mu_s - \nu_s\|_{var}$ is measurable and bounded, by Girsanov's theorem, $\tilde{W}$ is a Brownian motion under the weighted probability measure $\mathbb{Q} := R_T\mathbb{P}$, where

$$R_t := \mathrm{e}^{\int_0^t \langle \xi_s^{\mu,\nu}, \mathrm{d}W_s\rangle - \frac{1}{2}\int_0^t |\xi_s^{\mu,\nu}|^2\mathrm{d}s}, \quad t \in [0,T]$$

is a martingale. Then by the weak uniqueness of (4.5), the law of (X_t^μ, Y_t^μ) under $\mathbb{Q}$ satisfies

CD2 (4.8)
$$\mathscr{L}_{(X_t^\mu, Y_t^\mu)|\mathbb{Q}} = \mathscr{L}_{(X_t^\nu, Y_t^\nu)} = \Psi_t(\nu), \quad t \in [0,T].$$

Combining this with the martingale property of R_t, Pinsker's inequality, (4.7) and letting $\mathbb{E}_\mathbb{Q}$ be the expectation with respect to $\mathbb{Q}$, we derive

CDN (4.9)
$$\begin{aligned} \|\Psi_t(\mu) - \Psi_t(\nu)\|_{var}^2 &:= \sup_{|f|\le 1} \big|\mathbb{E}\big[f(X_t^\mu, Y_t^\mu)(1 - R_t)\big]\big|^2 \le \big(\mathbb{E}[|1 - R_t|]\big)^2 \\ &\le 2\mathbb{E}[R_t \log R_t] = 2\mathbb{E}_\mathbb{Q}[\log R_t] = \int_0^t \mathbb{E}_\mathbb{Q}[|\xi_s^{\mu,\nu}|^2]\mathrm{d}s \le c_1 \int_0^t K_s^2 \|\mu_s - \nu_s\|_{var}^2 \mathrm{d}s. \end{aligned}$$

This implies that when $\lambda > 0$ is large enough, Ψ is contractive under the complete metric

$$\rho_\lambda(\mu,\nu) := \sup_{t\in[0,T]} \mathrm{e}^{-\lambda t}\|\mu_t - \nu_t\|_{var}, \quad \mu, \nu \in \mathscr{C}^\gamma,$$

and hence has a unique fixed point.

(2) Let $\hat{\mathscr{P}} = \mathscr{P}_V$ and (4.3) holds. The following argument is similar to the proof of [13, Theorem 1.1 (1)], we include here for completeness. Let

$$\mathscr{C}_N^\gamma := \Big\{\mu \in \mathscr{C}^\gamma : \sup_{t\in[0,T]} \mu_t(V)\mathrm{e}^{-Nt} \le N(1 + \gamma(V))\Big\}, \quad N \ge 1.$$

It suffices to find a constant $N \ge 1$ such that $\Psi\mathscr{C}_N^\gamma \subset \mathscr{C}_N^\gamma$ and Ψ has a unique fixed point in $\mathscr{C}_N^\gamma$.

Firstly, by (A_3) for $Z_t^{(2)}(x,y) := Z_t^{(2)}(x,y,\delta_0)$ and (4.4), we find a constant $0 < c_0 < \infty$ such that

$$\varepsilon \sup_{y' \in B_\varepsilon(y)} \Big(|Z_t^{(1)}(x,y)| \|\nabla^{(1)}\nabla^{(2)} V(x,y')\| + |Z_t^{(2)}(x,y,\mu_t)| \big(\|\nabla^{(2)} V(x,y')\| + \|(\nabla^{(2)})^2 V(x,y')\| \big) \Big)$$
$$+ \langle Z_t^{(1)}(x,y), \nabla V(\cdot,y)(x) \rangle + \langle Z_t^{(2)}(x,y,\mu_t), \nabla V(x,\cdot)(y) \rangle \le \eta_t V(x,y) + c_0 K_t \mu_t(V)$$

holds for all $t \in [0,T]$, $(x,y) \in \mathbb{R}^{d_1+d_2}$ and $\mu \in \mathscr{C}^\gamma$. As in (2.10), by combining this with (A_2) and Itô's formula for

$$(X_t^\mu, \tilde{Y}_t^\mu) := (X_t^\mu, \Theta_t(Y_t^\mu)), \quad t \in [0,T],$$

we find a constant $0 < c_1 < \infty$ such that

CDO (4.10) $$\mathrm{d}\big(V(X_t^\mu, \tilde{Y}_t^\mu)\big)^2 \le c_1 \Big((1+\eta_t) \big(V(X_t^\mu, \tilde{Y}_t^\mu)^2 + K_t^2 \mu_t(V)^2 \big) \Big) \mathrm{d}t + \mathrm{d}M_t, \quad t \in [0,T],$$

for some martingale M_t. Then there exists a constant $0 < c_2 < \infty$ such that

$$\mathbb{E}[V(X_t^\mu, \tilde{Y}_t^\mu)^2 | \mathscr{F}_0] \le \mathrm{e}^{c_1 \int_0^t (1+\eta_r)\mathrm{d}r} V(X_0, Y_0)^2 + \int_0^t c_1 K_s^2 \mathrm{e}^{c_1 \int_s^t (1+\eta_r)\mathrm{d}r} \mu_s(V)^2 \mathrm{d}s$$

CDS (4.11) $$\le c_2 V(X_0,Y_0)^2 + c_2 N^2 \big(1+\gamma(V)\big)^2 \int_0^t K_s^2 \mathrm{e}^{2Ns} \mathrm{d}s$$
$$\le c_2 V(X_0,Y_0)^2 + c_2 \big(N(1+\gamma(V)) \mathrm{e}^{Nt} \big)^2 \int_0^t K_s^2 \mathrm{e}^{-2N(t-s)} \mathrm{d}s, \quad t \in [0,T], \quad \mu \in \mathscr{C}_N^\gamma.$$

Noting that $K \in L^2([0,T])$ implies

$$\lim_{N \to \infty} \sup_{t \in [0,T]} \int_0^t K_s^2 \mathrm{e}^{-2N(t-s)} \mathrm{d}s = 0,$$

while (A_3), (2.5) and Jensen's inequality yield

$$\big(\Psi_t(\mu)\big)(V) := \mathbb{E}[V(X_t^\mu, Y_t^\mu)] \le c_3 \mathbb{E}[V(X_t^\mu, \tilde{Y}_t^\mu)] \le c_3 \mathbb{E}\Big[\big(\mathbb{E}(V(X_t^\mu, \tilde{Y}_t^\mu)^2 | \mathscr{F}_0) \big)^{\frac{1}{2}} \Big]$$

holds for some constant $0 < c_3 < \infty$, we find a constant $1 \le N_0 < \infty$ such that (4.11) yields

$$\sup_{t \in [0,T]} \{\Psi_t(\mu)\}(V) \mathrm{e}^{-Nt} \le c_2 \mathbb{E}[V(X_0,Y_0)] + c_2 N \big(1+\gamma(V)\big) \int_0^t K_s^2 \mathrm{e}^{-2N(t-s)} \mathrm{d}s$$
$$= c_2 \gamma(V) + c_2 N \big(1+\gamma(V)\big) \int_0^t K_s^2 \mathrm{e}^{-2N(t-s)} \mathrm{d} \le N(1+\gamma(V)), \quad N \ge N_0.$$

Therefore, $\Psi \mathscr{C}_N^\gamma \subset \mathscr{C}_N^\gamma$ for $N \ge N_0$.

Next, let $N \ge N_0$. We intend to prove that Ψ has a unique fixed point in $\mathscr{C}_N^\gamma$, by using the Girsanov transform defined in (1) to show that for large $\lambda > 0$, Ψ is contractive in the following complete metric on $\mathscr{C}_N^\gamma$:

$$\rho_{V,\lambda}(\mu,\nu) := \sup_{t \in [0,T]} \mathrm{e}^{-\lambda t} \|\mu_t - \nu_t\|_V.$$

By (4.4) and (A_2), we find a constant $c_0 > 0$ such that instead of (4.7),

XIS' (4.12) $$|\xi_s^{\mu,\nu}|^2 \le c_0 K_s^2 \|\mu_s - \nu_s\|_V^2, \quad s \in [0,T],\ \mu,\nu \in \mathscr{C}_N^\gamma.$$

So, this together with (4.8), (4.11) and (2.2) yields that for some constants $0 < c_1(N), c_2(N) < \infty$,

$$\begin{aligned}\|\Phi_t(\mu) - \Phi_t(\nu)\|_V &= \sup_{|f|\le V} \big|\mathbb{E}\big[f(X_t^\mu, Y_t^\mu)(1-R_t)\big]\big| \le \mathbb{E}\big[V(X_t^\mu, Y_t^\mu)|1-R_t|\big]\\
\text{RPP2}\quad (4.13)\qquad &\le \mathbb{E}\Big[\big(\mathbb{E}(V(X_t^\mu, Y_t^\mu)^2|\mathscr{F}_0)\big)^{\frac12}\big(\mathbb{E}[|R_t-1|^2|\mathscr{F}_0]\big)^{\frac12}\Big]\\
&\le c_1(N)\mathbb{E}\Big[V(X_0,Y_0)\big(\mathbb{E}[R_t^2-1|\mathscr{F}_0]\big)^{\frac12}\Big],\quad \mu,\nu\in\mathscr{C}_N^\gamma,\end{aligned}$$

and due to $\mathrm{e}^r - 1 \le r\mathrm{e}^r$ for $r \in \mathbb{R}$,

$$\begin{aligned}&\mathbb{E}[R_t^2 - 1|\mathscr{F}_0] \le \mathbb{E}\bigg[\mathrm{e}^{2\int_0^t\langle \xi_s^{\mu,\nu}, \mathrm{d}W_s\rangle - \int_0^t |\xi_s^{\mu,\nu}|^2\mathrm{d}s} - 1\bigg|\mathscr{F}_0\bigg]\\
&\le \mathbb{E}\bigg[\mathrm{e}^{2\int_0^t\langle \xi_s^{\mu,\nu}, \mathrm{d}W_s\rangle - 4\int_0^t |\xi_s^{\mu,\nu}|^2\mathrm{d}s + \int_0^t 3c_0K_s^2\|\mu_s-\nu_s\|_V^2\mathrm{d}s}\bigg|\mathscr{F}_0\bigg] - 1\\
&= \mathrm{e}^{\int_0^t 3c_0K_s^2\|\mu_s-\nu_s\|_V^2\mathrm{d}s} - 1 \le \mathrm{e}^{\int_0^t 3c_0K_s^2\|\mu_s-\nu_s\|_V^2\mathrm{d}s}\int_0^t 3c_0K_s^2\|\mu_s-\nu_s\|_V^2\mathrm{d}s\\
&\le c_2(N)\int_0^t K_s^2\|\mu_s-\nu_s\|_V^2\mathrm{d}s,\quad \mu,\nu\in\mathscr{C}_N^\gamma,\ t\in[0,T].\end{aligned}$$

Combining this with (4.13), we find a constant $0 < c_3(N) < \infty$ such that

$$\begin{aligned}&\rho_\lambda(\Phi(\mu),\Phi(\nu)) = \sup_{t\in[0,T]} \mathrm{e}^{-\lambda t}\|\Phi_t(\mu) - \Phi_t(\nu)\|_V\\
&\le c_3(N)(1+\gamma(V))\rho_\lambda(\mu,\nu)\sup_{t\in[0,T]}\bigg(\int_0^t K_s^2\mathrm{e}^{-2\lambda(t-s)}\mathrm{d}s\bigg)^{\frac12},\quad \mu,\nu\in\mathscr{C}_N^\gamma,\ t\in[0,T].\end{aligned}$$

Since

$$\lim_{\lambda\to\infty}\sup_{t\in[0,T]}\bigg(\int_0^t K_s^2\mathrm{e}^{-2\lambda(t-s)}\mathrm{d}s\bigg)^{\frac12} = 0,$$

when $\lambda > 0$ is large enough, Ψ is contractive on the complete metric space $(\mathscr{C}_N^\gamma, \rho_\lambda)$, so that it has a unique fixed point in $\mathscr{C}_N^\gamma$ as desired.

□

4.2 Uniform ergodicity of (4.2)

Assume that

$(\hat{B})$ (B_1)-(B_3) and (B_4') hold for $Z^{(2)}(x,y) := Z^{(2)}(x,y,\delta_0)$. Moreover, there exists a constant $0 < \kappa < \infty$ such that for all $(x,y) \in \mathbb{R}^{d_1+d_2}$ and $\gamma_1, \gamma_2 \in \mathscr{P}$,

$$|Z^{(2)}(x,y,\gamma_1) - Z^{(2)}(x,y,\gamma_2)| \le \kappa\|\gamma_1 - \gamma_2\|_{var}.$$

To investigate the uniform exponential ergodicity, we consider the following reference SDE for $\gamma \in \mathscr{P}$,

EQ4.2 (4.14)
$$\begin{cases} \mathrm{d}X_t^\gamma = Z^{(1)}(X_t^\gamma, Y_t^\gamma)\mathrm{d}t, \\ \mathrm{d}Y_t^\gamma = \big(Z^{(2)}(X_t^\gamma, Y_t^\gamma, \gamma) + b(Y_t^\gamma)\big)\mathrm{d}t + \sigma(Y_t^\gamma)\mathrm{d}W_t. \end{cases}$$

Let $(P_t^\gamma)^*\nu = \mathscr{L}_{(X_t^\gamma, Y_t^\gamma)}$ with $\mathscr{L}_{(X_0^\gamma, Y_0^\gamma)} = \nu$. Denote by $(X_t^{\gamma,x,y}, Y_t^{\gamma,x,y})$ the solution to (4.14) with the initial value $(x, y) \in \mathbb{R}^{d_1+d_2}$.

Thm4 **Theorem 4.2.** *Assume $(\hat{B})$. If κ is small enough and Φ in (B_3) is convex with $\int_0^\infty \frac{1}{\Phi(s)}\mathrm{d}s < \infty$, then P_t^* associated with (4.2) has a unique invariant probability measure $\hat{\mu} \in \mathscr{P}$ such that $\hat{\mu}(\Phi(\varepsilon_0 V)) < \infty$ for some $\varepsilon_0 > 0$, and there exists constants $c, \lambda > 0$ such that*

$$\|P_t^*\nu - \hat{\mu}\|_{var} \le ce^{-\lambda t}\|\hat{\mu} - \nu\|_{var}, \quad t \ge 0, \nu \in \mathscr{P}.$$

Proof. According to [17, Lemma 3.3], it is sufficient to find constants $0 < k_0, c, \lambda < \infty$ such that when $\kappa < k_0$, for any $\mu \in \mathscr{P}$, $(P_t^\gamma)_{t\ge 0}$ has a unique invariant measure μ_γ satisfying

EQ4.3 (4.15)
$$\|(P_t^\gamma)^*\mu - \mu_\gamma\|_{var} \le ce^{-\lambda t}\|\mu - \mu_\gamma\|_{var}.$$

By the uniform Harris type theorem [17, Lemma 3.3], we only need find $t_0, t_1 > 0$ and a measurable set $B \in \mathscr{B}(\mathbb{R}^{d_1+d_2})$, such that

EQ4.4 (4.16)
$$\inf_{\gamma\in\mathscr{P}, z\in\mathbb{R}^{d_1+d_2}} P_{t_0}^\gamma(z, B) > 0,$$

EQ4.5 (4.17)
$$\sup_{\gamma\in\mathscr{P}; z, z'\in B} \|(P_{t_1}^\gamma)^*\delta_z - (P_{t_1}^\gamma)^*\delta_{z'}\|_{var} < 2.$$

Below we prove these two estimates for $t_1 = 1, B = B_k(0)$ for large enough $k > 0$ and some $t_0 > 1$.

(a) Proof of (4.16). Let $Z_t^{(2)}(x, y) = Z_t^{(2)}(x, y, \delta_0)$, $(X_t^{\delta_0,x,y}, Y_t^{\delta_0,x,y})_{t\ge 0}$ solve (3.1) with initial value $(x, y) \in \mathbb{R}^{d_1+d_2}$, and let $(P_t^{\delta_0})_{t\ge 0}$ be the associated Markov semigroup. By Theorem 3.1 (1) and (3), $P_t^{\delta_0}$ has a unique invariant probability measure μ such that (3.6) holds for some constants $0 < c, \lambda < \infty$. Consequently, there exists a constant $1 \le t_0 < \infty$ such that

$$\|(P_t^{\delta_0})^*\nu - \mu\|_{var} \le \frac{1}{4}, \quad t \ge t_0, \ \nu \in \mathscr{P}.$$

Taking $k > 0$ such that $\mu(B_k(0)) > \frac{3}{4}$, this implies that for $B := B_k(0)$,

FY1 (4.18)
$$P_{t_0}^{\delta_0} 1_B(x, y) \ge 1 - P_{t_0}^{\delta_0} 1_{B^c}(x, y) \ge 1 - \mu(B^c) - \frac{1}{4} \ge \frac{1}{2}, \quad (x, y) \in \mathbb{R}^{d_1+d_2}.$$

Now, for any $\gamma \in \mathscr{P}$, let

$$\xi_s^\gamma := \big(\sigma^*(\sigma\sigma^*)^{-1}\big)(X_s^{\delta_0,x,y}, Y_s^{\delta_0,x,y})\big(Z_2^{(2)}(X_s^{\delta_0,x,y}, Y_s^{\delta_0,x,y}, \gamma) - Z_2^{(2)}(X_s^{\delta_0,x,y}, Y_s^{\delta_0,x,y}, \delta_0)\big).$$

Then (B_1) and (B_2) imply

$$|\xi_s^\gamma| \le c_1, \quad s \ge 0,\ \gamma \in \mathscr{P}$$

for some constant $0 < c_1 < \infty$. Let

$$R := \mathrm{e}^{\int_0^{t_0} \langle \xi_s^\gamma, \mathrm{d}W_s\rangle - \frac{1}{2}\int_0^{t_0} |\xi_s|^2 \mathrm{d}s}.$$

Thus,

$$\mathbb{E}[R^{-1}] = \mathbb{E}[\mathrm{e}^{-\int_0^{t_0} \langle \xi_s^\gamma, \mathrm{d}W_s\rangle + \frac{1}{2}\int_0^{t_0} |\xi_s|^2 \mathrm{d}s}] \le \mathrm{e}^{c_1^2 \frac{t_0}{2}},$$

and by Girsanov's theorem,

FY0 (4.19) $$\mathscr{L}_{(X_t^{\delta_0,x,y}, Y_t^{\delta_0,x,y})|\mathbb{Q}} = \mathscr{L}_{(X_t^{\gamma,x,y}, Y_t^{\gamma,x,y})}, \quad t \in [0, t_0].$$

So that Schwarz inequality and (4.18) yield

$$P_{t_0}^\gamma 1_B(x,y) = \mathbb{E}[1_B(X_{t_0}^{\delta_0,x,y}, Y_{t_0}^{\delta_0,x,y})R] \ge \frac{(\mathbb{E}[1_B(X_{t_0}^{\delta_0,x,y}, Y_{t_0}^{\delta_0,x,y})])^2}{\mathbb{E}[R^{-1}]} \ge \frac{1}{2}\mathrm{e}^{-c_1^2 t_0} > 0,$$

for any $(x,y) \in \mathbb{R}^{d_1+d_2}$ and any $\gamma \in \mathscr{P}$. Hence, (4.16) holds.

(b) Proof of (4.17): Recall that $B = B_k(0)$. Let $p_1(x,y;\cdot)$ be the distribution density of $P_1^{\delta_0}(x,y;\cdot)$. By Proposition 3.3, we have

$$\delta := \inf_{(x,y),(x',y') \in B} p_1(x,y;x',y') > 0,$$

where p_1 is the heat kernel of P_1. Then for any $z, z' \in B$,

FY2 (4.20)
$$\begin{aligned} &\|(P_1^{\delta_0})^*\delta_z - (P_1^{\delta_0})^*\delta_{z'}\|_{var} = \int_{\mathbb{R}^{2d}} |p_1(z;x,y) - p_1(z';x,y)| \mathrm{d}x\mathrm{d}y \\ &\le \int_{\mathbb{R}^{2d}} \big(p_1(z;x,y) + p_1(z';x,y) - 2[p_1(z;x,y) \wedge p_1(z';x,y)]\big)\mathrm{d}x\mathrm{d}y \\ &\le 2 - 2\delta \mathrm{vol}(B) := \delta' < 2. \end{aligned}$$

On the other hand, by $(\hat{B})$, (4.19) and Pinsker's inequality as in (4.9), we find a constant $0 < c_2 < \infty$ such that

$$\|(P_1^{\delta_0})^*\delta_z - (P_1^\gamma)^*\delta_z\|_{var} \le c_2\kappa, \quad z \in \mathbb{R}^{d_1+d_2}.$$

Combining this with (4.20) we conclude that when κ is small enough, (4.17) holds for $t_1 = 1$. □

References

CHSG [1] E. Camrud, D. P.Herzog, G. Stoltz, M.Gordina *Weighted L^2-congractivity of Langevin dynamics with singular potentials,* Nonlinearity, 35(2022), 998-1035.

C16 [2] P. E. Chaudru de Raynal, *Strong existence and uniqueness for stochastic differential equation with Hölder drift and degenerate noise,* Ann. Inst. Henri Poincare Probab. Stat. 53(2017), 259-286.

DMS [3] J. Dolbeault, C. Mouhot, C. Sschmeiser, *Hypocoercivity for linear kinetic equations conserving mass,* Trans. Amer. Math. Soc. 367(2015), 3807-3828.

Doob [4] J. L. Doob, *Asymptotic properties of Markoff transition probabilities,* Trans. Amer. Math. Soc. 63(1948), 394–421.

GIMV [5] F. Golse, C. Imbert, C. Mouhot, F. Vasseur, *Harnack inequality for kinetic Fokker-Planck equations with rough coefficients and applications to the Landau equation,* Ann. della Scuola Norm. Sup. di Pisa 19(2019), 253-295.

GS [6] M. Grothaus, P. Stilgenbauer, *A hypocoercivity related ergodicity method for singularly distorted non-symmetric diffusions,* Integral Equations and Operator Theory 83(2015), 331-379.

GW19 [7] M. Grothaus, F.-Y. Wang, *Weak poincaré inequalities for convergence rate of degenerate diffusion processes,* Ann. Probab. 47(2019), 2930-2952.

ZZZ [8] Z. Hao, X. Zhang, R. Zhu, X. Zhu, *Singular Kinetic equations and applications,* arXiv:2108.05042.

HM [9] M. Hairer, J.C. Mattingly, *Yet another look at Harris' ergodic theorem for Markov chains,* Progress in Probability 63(2011), 109-117.

HRW [10] X. Huang, P. Ren, F.-Y. Wang, *Distribution dependent stochastic differential equations,* Front. Math. China 16(2021), 257-301.

12 [11] N. V. Krylov, M. Röckner, *Strong solutions of stochastic equations with singular time dependent drift,* Probab. Theory Related Fields 131(2005), 154-196.

LPP [12] A. Lanconelli, S. Pagliarani, A. Pascucci, *Local densities for a class of degenerate diffusions,* Ann. Inst. H. Poincaré Probab. Statist. 56(2020), 1440-1464.

R21 [13] P. Ren, *Singular McKean-Vlasov SDEs: well-posedness, regularities and Wang's Harnack inequality,* Stoc. Proc. Appl. 156(2023), 291-311.

RW21 [14] P. Ren, F.-Y. Wang, *Exponential convergence in entropy and Wasserstein for McKean-Vlasov SDEs,* Nolinear Analysis 206(2021), 112259.

V [15] C. Villani, *Hypocoercivity,* Mem. Amer. Math. Soc. 202(2009).

W21a [16] F.-Y. Wang, *Exponential ergodicity for fully non-dissipative McKean-Vlasov SDEs,* Bernoulli. 29(2023), 1035-1062.

W21c [17] F.-Y. Wang, *Exponential ergodicity for singular reflecting McKean-Vlasov SDEs,* Stoc. Proc. Appl. 160(2023), 265-293.

WZ16 [18] F.-Y. Wang, X. Zhang, *Degenerate SDE with Hölder-Dini drift and non-Lipschitz noise coefficient,* SIAM J. Math. Anal. 48(2016), 2189-2226.

W17 [19] F.-Y. Wang, *Hypercontractivity and applications for stochastic Hamiltonian systems,* J. Funct. Anal. 272(2017), 5360–5383.

XXZZ [20] P. Xia, L. Xie, X. Zhang, G. Zhao, *$L^q(L^p)$-theory of stochastic differential equations,* Stoch. Proc. Appl. 130(2020), 5188-5211.

XZ [21] L. Xie, X. Zhang, *Ergodicity of stochastic differential equations with jumps and singular coefficients,* Ann. Inst. Henri Poincaré Probab. Stat. 56(2020), 175-229.