

# Exponential Ergodicity for Non-Dissipative McKean-Vlasov SDEs \*

Feng-Yu Wang<sup>a),b)</sup>

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

June 24, 2021

## Abstract

Under Lyapunov and monotone conditions, the exponential ergodicity in the induced Wasserstein quasi-distance is proved for a class of non-dissipative McKean-Vlasov SDEs, which strengthen some recent results established under dissipative conditions in long distance. Moreover, when the SDE is order-preserving, the exponential ergodicity is derived in the Wasserstein distance induced by one-dimensional increasing functions chosen according to the coefficients of the equation.

AMS subject Classification: 60H10, 60G65.

Keywords: Exponential ergodicity, McKean-Vlasov SDEs, fully non-dissipative condition, Lyapunov condition, coupling method.

## 1 Introduction

Consider the following second order differential operator on  $\mathbb{R}^d$ :

$$L := \frac{1}{2} \sum_{i,j=1}^d a_{ij} \partial_i \partial_j + \sum_{i=1}^d b_i \partial_i,$$

where  $a := (a_{ij})_{1 \leq i,j \leq d}$  is positive definite and  $C^2$ -smooth,  $b := (b_i)_{1 \leq i \leq d}$  is  $C^1$ -smooth. The Harris theorem says that if there exists a Lyapunov function  $0 \leq V \in C^2(\mathbb{R}^d)$  with  $\lim_{|x| \rightarrow \infty} V(x) = \infty$  such that

$$(1.1) \quad LV \leq C_0 - C_1 V$$

---

\*Supported in part by the National Key R&D Program of China (No. 2020YFA0712900) and NNSFC (11771326, 11831014, 11921001).

holds for some constants  $C_0, C_1 > 0$ , then the diffusion process generated by  $L$  is exponentially ergodic, see [12, Theorem 1.5] for a more general assertion, and see [9, Theorem 2.1] for an explicit estimate on the exponential convergence rate when  $a = I_d$ , the  $d \times d$ -identity matrix. A typical example satisfying (1.1) is  $L = \frac{1}{2}\Delta + b \cdot \nabla$  with  $b \in C^1$  such that

$$(1.2) \quad b(x) = -|x|^{p-2}x, \quad |x| \geq 1$$

holds for some  $p \geq 1$ . It is easy to see that when  $p \geq 2$ , this operator is dissipative in long distance, i.e.

$$(1.3) \quad \langle x - y, b(x) - b(y) \rangle \leq C_3|x - y| - C_4|x - y|^2$$

holds for some constants  $C_3, C_4 > 0$ . However, when  $p \in [1, 2)$ , it is fully non-dissipative in the sense that

$$(1.4) \quad \sup_{|x-y|=r} \langle x - y, b(x) - b(y) \rangle \geq 0, \quad r \geq 0.$$

On the other hand, when  $p \in (0, 1)$ , the diffusion process is not exponential ergodic since the Poincaré inequality fails (see for instance [22, Corollary 1.4]). Therefore, in this example,  $p = 1$  is critical for the exponential ergodicity.

In this paper, we aim to extend the above mentioned Harris theorem to McKean-Vlasov SDEs (also called distribution dependent or mean field SDEs), for which the time-marginal distribution  $\mu_t$  of the solution satisfies the following nonlinear Fokker-Planck equation on  $\mathcal{P}$ , the space of probability measures on  $\mathbb{R}^d$ :

$$(1.5) \quad \partial_t \mu_t = L_{\mu_t}^* \mu_t$$

in the sense that  $\mu_t$  is continuous in  $t$  under the weak topology and

$$\mu_t(f) := \int_{\mathbb{R}^d} f d\mu_t = \mu_0(f) + \int_0^t \mu_s(L_{\mu_s} f) ds, \quad t \geq 0, f \in C_0^\infty(\mathbb{R}^d),$$

where for any  $\mu \in \mathcal{P}$ , the operator  $L_\mu$  is defined by

$$(1.6) \quad L_\mu := \frac{1}{2} \sum_{i,j=1}^d a_{ij} \partial_i \partial_j + \sum_{i=1}^d b_i(\cdot, \mu) \partial_i$$

for the above mentioned  $a$  and a distribution dependent drift

$$b : \mathbb{R}^d \times \mathcal{P} \rightarrow \mathbb{R}^d.$$

This nonlinear Fokker-Planck equation can be characterized by the following McKean-Vlasov SDE on  $\mathbb{R}^d$ :

$$(1.7) \quad dX_t = b(X_t, \mathcal{L}_{X_t}) dt + \sigma(X_t) dW_t,$$

where  $\sigma : [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^m$  such that  $\sigma \sigma^* = a$ ,  $W_t$  is the  $m$ -dimensional Brownian motion on a complete filtration probability space  $(\Omega, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ , and  $\mathcal{L}_\xi$  is the distribution

of a random variable  $\xi$ . Indeed, according to [1], for any solution of (2.1) with  $\mu_t := \mathcal{L}_{X_t}$  satisfying

$$(1.8) \quad \int_0^T \mu_t(\|a\| + |b(\cdot, \mu_t)|) dt < \infty, \quad T \geq 0,$$

$\mu_t$  solves (1.5); while a solution of (1.5) satisfying (1.8) coincides with  $\mathcal{L}_{X_t}$  for a weak solution to (1.7). Therefore, if (1.7) is weakly well-posed in a subspace  $\hat{\mathcal{P}} \subset \mathcal{P}$  (i.e. for any initial distribution in  $\hat{\mathcal{P}}$  it has a unique solution with  $\mathcal{L}_{X_t} \in \hat{\mathcal{P}}$ ), continuous in  $t$  under the weak topology, the exponential ergodicity of (1.5) in  $\hat{\mathcal{P}}$  is equivalent to that of the SDE (1.7) with initial distributions in  $\hat{\mathcal{P}}$ .

In recent years, different types of exponential ergodicity have been investigated for solutions to (1.5) under the dissipative condition (1.3) in long distance and that the dependence of  $b(x, \mu)$  on  $\mu$  is weak enough. When  $a = I_d$ , see [18] for the exponential ergodicity in  $\mathbf{W}_2$ , [21] for the ergodicity under the polynomial mixing property for the associated mean-field particle systems, [10] for the exponential convergence in the total variation norm for Dirac initial measures, [11] for exponential ergodicity in the “mean field entropy”, [17] for exponential ergodicity in the  $L^1$ -Wasserstein distance. See also [20] for the exponential ergodicity in the relative entropy where  $a$  may be non-constant. However, as already mentioned above that the condition (1.3) excludes fully non-dissipative examples of  $b$  satisfying (1.2) for  $p \in [1, 2)$ . On the other hand, in this case the diffusion process generated by  $L := \Delta + b \cdot \nabla$  is exponential ergodic according to the Harris theorem, so that in the spirit of stable perturbations, when  $b(x, \mu) = -\nabla|x|^p + b_0(x, \mu)$  for  $p \in [1, 2)$  and large  $|x|$  and  $b_0$  is small enough, the exponential ergodicity for (1.5) with  $a = I_d$  should also hold. This has been confirmed in [2, Theorem 3.1] for  $b(x, \mu) := b_0(x) + \varepsilon b_1(x, \mu)$  with small  $\varepsilon > 0$ , where  $b_0$  and  $b_1$  satisfies  $\langle b_0(x), x \rangle \leq -c_1|x|$  for some constant  $c_1 > 0$  and large  $|x|$ ,  $\|b_1\|_\infty \leq c_2$  and

$$|b_0(x) - b_0(y)| + |b_1(x, \mu) - b_1(x, \nu)| \leq c_2(|x - y| + \mathbf{W}_2(\mu, \nu))$$

for some constant  $c_2 > 0$  and the  $L^2$ -Wasserstein distance  $\mathbf{W}_2$ . In this paper, we will prove a general version of such a result for (1.5), which includes non-constant diffusion coefficient  $a$  and non  $\mathbf{W}_2$ -Lipschitz  $b_1(x, \cdot)$ , see Example 1.2 below.

The main idea of the present study is to decompose  $a$  into  $a = \lambda I_d + \hat{\sigma} \hat{\sigma}^*$  for some constant  $\lambda > 0$  and Lipschitz continuous  $\hat{\sigma}$  as in [19], then for the corresponding McKean-Vlasov SDE we adopt the coupling by reflection for the noise with coefficient  $\sqrt{\lambda} I_d$ , and the coupling by parallel displacement for the noise with coefficient  $\hat{\sigma}$ .

The coupling by reflection was applied in [5, 6, 7] to estimate the first eigenvalue on Riemannian manifolds as well as the spectral gap for elliptic diffusions, and has been developed in the study of SDEs and SPDEs. Unlike in the study of classical SDEs (or diffusion processes) for which we may let two marginal processes move together after the first meeting time (i.e. coupling time), in the distribution dependent setting this is no-longer practicable since the difference of marginal distributions may separate the marginal processes after the coupling time. To fix this problem, after the coupling time we will take the coupling by parallel displacement for all noises, so that the marginal processes will not move too far away each other.

The remainder of the paper is organized as follows. In Section 2, we investigate the exponential ergodicity of (1.5) under Lyapunov and monotone conditions, which apply to a class of fully non-dissipative models (see Examples 2.1 and 2.2). In Section 3, we prove the exponential ergodicity under the dissipative condition in long time, which extends some existing results to non-constant  $a$  (see Example 3.1). Finally, Section 4 concerns with the exponential ergodicity for order-preserving McKean-Vlasov SDEs.

## 2 Under Lyapunov and monotone conditions

We will consider the following more general version of (1.7) where the coefficients may also depend on the time parameter:

$$(2.1) \quad dX_t = b_t(X_t, \mathcal{L}_{X_t})dt + \sigma_t(X_t)dW_t,$$

where

$$\sigma : [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^m, \quad b : [0, \infty) \times \mathbb{R}^d \times \mathcal{P} \rightarrow \mathbb{R}^d$$

are measurable and  $W_t$  is the  $m$ -dimensional Brownian motion. Recall that the SDE (2.1) is called strongly (respectively, weakly) well-posed in a subspace  $\hat{\mathcal{P}} \subset \mathcal{P}$ , if for any  $s \geq 0$  and  $\mathcal{F}_s$ -measurable initial value  $X_s$  with  $\mathcal{L}_{X_s} \in \hat{\mathcal{P}}$  (respectively, any  $\mu \in \hat{\mathcal{P}}$ ), (2.1) has a unique solution from time  $s$  (respectively, a unique weak solution with initial distribution  $\mu$  from time  $s$ ) such that the time-marginal of the solution is continuous in  $\hat{\mathcal{P}}$  under the weak topology. We call (2.1) well-posed if it is both strongly and weakly well-posed. In this case, we denote  $P_{s,t}^* \mu = \mathcal{L}_{X_t}$  for  $X_t$  solving (2.1) from time  $s$  with  $\mathcal{L}_{X_s} = \mu \in \hat{\mathcal{P}}$ , so that  $t \mapsto P_{s,t}^* \mu$  is continuous in  $t \geq s$  and

$$(2.2) \quad P_{s,t}^* = P_{r,t}^* P_{s,r}^*, \quad 0 \leq s \leq r \leq t.$$

When  $b$  and  $a$  do not depend on  $t$ , we have  $P_{s,t}^* = P_{t-s}^* := P_{0,t-s}^*, t \geq s$ .

### 2.1 Main result

For any  $t \geq 0$  and  $\mu \in \mathcal{P}$ , consider the second-order differential operator

$$(2.3) \quad L_{t,\mu} := \frac{1}{2} \text{tr}\{\sigma_t \sigma_t^* \nabla^2\} + b_t(\cdot, \mu) \cdot \nabla.$$

For any probability measure  $\mu$  and a measurable function  $f$ , we denote  $\mu(f) = \int f d\mu$  is the integral exists. We assume the following Lyapunov condition. For any positive measurable function  $V$  on  $\mathbb{R}^d$ , let

$$\mathcal{P}_V := \{\mu \in \mathcal{P} : \mu(V) < \infty\}.$$

(H<sub>1</sub>) **(Lyapunov)** There exists a function  $0 \leq V \in C^2(V)$  with  $\lim_{|x| \rightarrow \infty} V(x) = \infty$  and

$$(2.4) \quad \sup_{t \geq 0, x \in \mathbb{R}^d} \frac{|\sigma_t(x) \nabla V(x)|}{1 + V(x)} < \infty,$$

such that for some  $K_0, K_1 \in L_{loc}^1([0, \infty); \mathbb{R})$

$$(2.5) \quad L_{t,\mu} V \leq K_0(t) - K_1(t)V, \quad t \geq 0, \mu \in \mathcal{P}_V.$$

We remark that the existence of invariant probability measure has been studied in [13] under an integrated Lyapunov condition weaker than (2.5), see also [15] for a recent survey on this topic.

For any  $l > 0$ , consider the class

$$\Psi_l := \{\psi \in C^2([0, l]; [0, \infty)) : \psi(0) = \psi'(l) = 0, \psi'|_{[0, l]} > 0\}.$$

For each  $\psi \in \Psi_l$ , we extend it to the half line by setting  $\psi(r) = \psi(r \wedge l)$ , so that  $\psi'$  is non-negative and Lipschitz continuous with compact support, with

$$(2.6) \quad c_\psi := \sup_{r>0} \frac{r\psi'(r)}{\psi(r)} < \infty.$$

When  $\psi'' \leq 0$ , we have  $\|\psi'\|_\infty := \sup |\psi'| = \psi'(0)$  and  $c_\psi = \lim_{r \downarrow 0} \frac{r\psi'(r)}{\psi(r)} = 1$ .

For any constant  $\beta > 0$ , the weighted Wasserstein distance (also called transportation cost) is given by

$$\mathbf{W}_{\psi, \beta V}(\mu, \nu) := \inf_{\pi \in \mathcal{C}(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \psi(|x - y|) (1 + \beta V(x) + \beta V(y)) \pi(dx, dy), \quad \mu, \nu \in \mathcal{P}_V.$$

In general,  $\mathbf{W}_{\psi, \beta V}$  is only a quasi-distance on  $\mathcal{P}_V$  as the triangle inequality may not hold. But it is complete in the sense that any  $\mathbf{W}_{\psi, \beta V}$ -Cauchy sequence in  $\mathcal{P}_V$  is convergent. For any  $\mu, \nu \in \mathcal{P}_V$ , we introduce

$$(2.7) \quad \hat{W}_{\psi, \beta V}(\mu, \nu) := \inf_{\pi \in \mathcal{C}(\mu, \nu)} \frac{\int_{\mathbb{R}^d \times \mathbb{R}^d} \psi(|x - y|) (1 + \beta V(x) + \beta V(y)) \pi(dx, dy)}{\int_{\mathbb{R}^d \times \mathbb{R}^d} \psi'(|x - y|) (1 + \beta V(x) + \beta V(y)) \pi(dx, dy)},$$

which will come naturally from Itô's formula for the process

$$\psi(|X_t - Y_t|) (1 + \beta V(X_t) + \beta V(Y_t))$$

for a coupling  $(X_t, Y_t)$  of the SDE. We observe that

$$\sup_{\pi \in \mathcal{C}(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \psi'(|x - y|) (1 + \beta V(x) + \beta V(y)) \pi(dx, dy) \leq 1 + \beta \mu(V) + \beta \nu(V),$$

so that  $\hat{\mathbf{W}}_{\psi, \beta V} \geq \frac{\mathbf{W}_{\psi, \beta V}(\mu, \nu)}{1 + \beta \mu(V) + \beta \nu(V)}$ . As shown in Example 1.2 below that in many cases

$$\hat{\mathbf{W}}_{\psi, \beta V} \geq \frac{c \mathbf{W}_{\psi, \beta V}(\mu, \nu)}{1 + \beta [\mu(V) \wedge \nu(V)]}$$

holds for some constant  $c > 0$ .

Moreover, let  $\|\nabla f\|_\infty$  be the Lipschitz constant of a real function  $f$  on  $\mathbb{R}^d$ . We need the following non-degenerate and monotone conditions.

(H<sub>2</sub>) **(Non-degeneracy)** There exist  $\alpha \in L^1_{loc}([0, \infty); (0, \infty))$  and measurable

$$\hat{\sigma} : [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^d$$

with  $\int_0^T \|\nabla \hat{\sigma}_t\|_\infty dt < \infty$  for  $T \in (0, \infty)$ , such that

$$(2.8) \quad a_t(x) := (\sigma_t \sigma_t^*)(x) = \alpha_t I_d + (\hat{\sigma}_t \hat{\sigma}_t^*)(x), \quad t \geq 0, x \in \mathbb{R}^d.$$

( $H_3$ ) (**Monotonicity**)  $b$  is bounded on bounded set in  $[0, \infty) \times \mathbb{R}^d \times \mathcal{P}_V$ . Moreover, there exist  $l > 0$ ,  $K, \theta, q_l \in L^1_{loc}([0, \infty); [0, \infty))$  and  $\psi \in \Psi_l$ , such that

$$(2.9) \quad 2\alpha_t \psi''(r) + K_t \psi'(r) \leq -q_l(t) \psi(r), \quad r \in [0, l], t \geq 0,$$

$$(2.10) \quad \begin{aligned} & \langle b_t(x, \mu) - b_t(y, \nu), x - y \rangle + \frac{1}{2} \|\hat{\sigma}_t(x) - \hat{\sigma}_t(y)\|_{HS}^2 \\ & \leq K_t |x - y|^2 + \theta_t |x - y| \hat{\mathbf{W}}_{\psi, \beta V}(\mu, \nu), \quad x, y \in \mathbb{R}^d, \mu, \nu \in \mathcal{P}_V, t \geq 0. \end{aligned}$$

**Remark 1.1.** (1) Since  $V \geq 0$  with  $V(x) \rightarrow \infty$  as  $|x| \rightarrow \infty$ , we have

$$(2.11) \quad \kappa_{l, \beta}(t) := \inf_{|x-y|>l} \frac{K_1(t)V(x) + K_1(t)V(y) - 2K_0(t)}{\beta^{-1} + V(x) + V(y)} \in \mathbb{R}, \quad l > 0,$$

and  $\kappa_{l, \beta}(t) > 0$  for large enough  $l > 0$  and  $K_1(t) > 0$ .

(2) Consider the one-dimensional differential operator  $L = 2\lambda \frac{d^2}{dr^2} + K \frac{d}{dr}$  on  $[0, l]$ . In (2.9) one may take  $\psi$  to be the first eigenfunction of  $L$  with Dirichlet boundary at 0 and Neumann boundary at  $l$ . In this case,  $q_l > 0$  is the first mixed eigenvalue.

(3) (2.4) and ( $H_2$ ) imply that

$$(2.12) \quad \begin{aligned} \alpha_{l, \beta}(t) := c_\psi \sup_{|x-y| \in (0, l)} & \left\{ \alpha_t \frac{|\nabla V(x) - \nabla V(y)|}{|x-y| \{\beta^{-1} + V(x) + V(y)\}} \right. \\ & \left. + \frac{|\{\hat{\sigma}_t(x) - \hat{\sigma}_t(y)\}[(\hat{\sigma}_t(\cdot)^* \nabla V)(x) + (\hat{\sigma}_t(\cdot)^* \nabla V)(y)]|}{|x-y| \{\beta^{-1} + V(x) + V(y)\}} \right\} < \infty \end{aligned}$$

for any  $\beta, l > 0$ . In many cases, we have  $\alpha_{l, \beta} \downarrow 0$  as  $\beta \downarrow 0$ . For instance, it is the case when  $V(x) = e^{|x|^p}$  for  $p \in (0, 1)$  and large  $|x|$ , and  $\hat{\sigma}$  is Lipschitz continuous with  $\|\hat{\sigma}(x)\| \leq c(1+|x|^q)$  for some constants  $c > 0$  and  $q \in (0, 1-p)$ , or  $V(x) = |x|^k$  for some  $k > 0$  and large  $|x|$ .

For  $K_0, q_l, \kappa_{l, \beta}$  and  $\alpha_{l, \beta}$  given in ( $H_1$ ), ( $H_3$ ), (2.11) and (2.12) respectively, let

$$(2.13) \quad \lambda_{l, \beta}(t) := \min \{ \kappa_{l, \beta}(t), q_l(t) - 2K_0(t)\beta - \alpha_{l, \beta}(t) \}.$$

Since  $\alpha_{l, \beta}(t) \rightarrow 0$  as  $\beta \rightarrow 0$ , and since  $\kappa_{l, \beta}(t) > 0$  for  $K_1(t) > 0$  and large  $l > 0$ , when  $K_1(t) > 0$  we may take large  $l > 0$  and small  $\beta > 0$  such that  $\lambda_{l, \beta}(t) > 0$ . The main result in this part is the following.

**Theorem 2.1.** Assume ( $H_1$ )-( $H_3$ ), with  $\psi'' \leq 0$  when  $\hat{\sigma}_t(\cdot)$  is non-constant for some  $t \geq 0$ . Then the SDE (2.1) is well-posed in  $\mathcal{P}_V$ , and  $P_t^* := P_{0,t}^*$  satisfies

$$(2.14) \quad \mathbf{W}_{\psi, \beta V}(P_t^* \mu, P_t^* \nu) \leq e^{-\int_0^t \{\lambda_{l, \beta}(s) - \theta_s\} ds} \mathbf{W}_{\psi, \beta V}(\mu, \nu), \quad t \geq 0, \mu, \nu \in \mathcal{P}_V.$$

Consequently, if  $(a, b)$  does not depend on  $t$  and  $\lambda_{l, \beta} > \theta$ , then  $P_t^*$  has a unique invariant probability measure  $\bar{\mu} \in \mathcal{P}_V$  such that

$$(2.15) \quad \mathbf{W}_{\psi, \beta V}(P_t^* \mu, \bar{\mu}) \leq e^{-(\lambda_{l, \beta} - \theta)t} \mathbf{W}_{\psi, \beta V}(\mu, \bar{\mu}), \quad t \geq 0, \mu \in \mathcal{P}_V.$$

## 2.2 An example

In the following example, the drift  $b_0$  is fully non-dissipative in the sense of (1.4). As mentioned in Introduction that a critical model for the exponential ergodicity is the diffusion process generated by  $\Delta - (\nabla H) \cdot \nabla$  with  $H(x) = |x|$  for large  $|x|$ , which is now covered by this example for  $p = 1$ . Moreover, the following example is not covered by [2] since (??) does not hold even for  $\hat{\sigma} = 0$ , because  $\log \mu(V)$  is not  $\mathbf{W}_2$ -Lipschitz continuous in  $\mu$ .

**Example 2.1.** Let  $a = I_d + \hat{\sigma} \hat{\sigma}^*$  for some Lipschitz continuous matrix valued function  $\hat{\sigma}$ ,  $V(x) = e^{(1+|x|^2)^{p/2}}$  for some  $p \in (0, 1]$ , and

$$b(x, \mu) := b_0(x) + \varepsilon \Phi(x, \log \mu(V))$$

for some  $\varepsilon \in [0, 1)$ ,  $b_0 \in C^1(\mathbb{R}^d)$  with  $b_0(x) = -|x|^{-p}x$  for  $|x| \geq 1$ , and  $\Phi \in C_b^1(\mathbb{R}^d \times [0, \infty); \mathbb{R}^d)$ . Let

$$(2.16) \quad \tilde{\mathbf{W}}_V(\mu, \nu) := \inf_{\pi \in \mathcal{C}(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \{1 \wedge |x - y|\} \cdot \{1 + V(x) + V(y)\} \pi(dx, dy).$$

Then when  $\varepsilon > 0$  is small enough,  $P_t^*$  has a unique invariant probability measure  $\bar{\mu} \in \mathcal{P}_V$ , and there exist constants  $c, q > 0$  such that

$$\tilde{\mathbf{W}}_V(P_t^* \mu, \bar{\mu}) \leq c e^{-qt} \tilde{\mathbf{W}}_V(\mu, \bar{\mu}), \quad t \geq 0, \mu \in \mathcal{P}_V.$$

*Proof.* It is easy to see that  $(H_1)$  holds for some constants  $K_0, K_1 > 0$ ,  $(H_2)$  holds for  $\alpha = 1$ . Since  $V(x) \rightarrow \infty$  as  $|x| \rightarrow \infty$ , we take  $l > 0$  such that

$$\inf_{|x-y| \geq l} \{K_1 V(x) + K_1 V(y) - 2K_0\} \geq 1.$$

So, in (2.11) the constant  $\kappa_{l,\beta} > 0$  for all  $\beta > 0$ . Next, take  $\psi \in \Psi_l$  such that (2.9) holds for some  $q_l > 0$ , for instance  $\psi$  is the first mixed eigenfunction of  $2\frac{d^2}{dr^2} + K\frac{d}{dr}$  on  $[0, l]$  with Dirichlet condition at 0 and Neumann condition at  $l$ . Then there exists a constant  $c_0 > 0$  such that

$$(2.17) \quad |V(x) - V(y)| \leq c_0 \psi(|x - y|)(V(x) + V(y)), \quad x, y \in \mathbb{R}^d.$$

Next, since for any  $\pi \in \mathcal{C}(\mu, \nu)$  we have

$$\begin{aligned} & \int_{\mathbb{R}^d \times \mathbb{R}^d} \psi'(|x - y|)(1 + \beta V(x) + \beta V(y)) \pi(dx, dy) \\ & \leq \|\psi'\|_\infty \int_{\{|x-y| \leq l\}} \{1 + (1 + e)\beta[V(x) \wedge V(y)]\} \pi(dx, dy) \\ & \leq (2 + e)[\mu(V) \wedge \nu(V)], \quad \beta \in (0, 1], \end{aligned}$$

(2.7) implies

$$\hat{W}_{\psi, \beta V}(\mu, \nu) \geq \frac{\mathbf{W}_{\psi, \beta V}(\mu, \nu)}{(2 + e)[\mu(V) \wedge \nu(V)]}, \quad \beta \in (0, 1].$$

Combining this with  $\Phi \in C_b^1$  and noting that (2.17) implies

$$|\mu(V) - \nu(V)| \leq \inf_{\pi \in \mathcal{C}(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |V(x) - V(y)| \pi(dx, dy) \leq c_0 \beta^{-1} \mathbf{W}_{\psi, \beta V}(\mu, \nu)$$

for some constant  $c_0 > 0$ , we find a constant  $c_1 > 0$  such that

$$\begin{aligned} |b(x, \mu) - b(x, \nu)| &\leq \varepsilon \|\nabla \Phi(x, \cdot)\|_\infty |\log \mu(V) - \log \nu(V)| \\ &\leq \frac{\varepsilon \|\nabla \Phi(x, \cdot)\|_\infty |\mu(V) - \nu(V)|}{\mu(V) \wedge \nu(V)} \leq c_1 \varepsilon \beta^{-1} \hat{\mathbf{W}}_{\psi, \beta V}(\mu, \nu), \quad \beta \in (0, 1]. \end{aligned}$$

Noting that  $\|\nabla b_0\|_\infty + \|\nabla \Phi\|_\infty + \|\nabla \hat{\sigma}\|_\infty < \infty$ , this implies  $(H_3)$  holds for some constant  $K > 0$  and  $\theta = c_1 \varepsilon \beta^{-1}$ ,  $\beta \in (0, 1]$ .

Finally, as observed in Remark 1.1(3) that for the present  $V$  we have  $\alpha_{l, \beta} \downarrow 0$  as  $\beta \downarrow 0$ . Then in (2.13),  $\lambda_{l, \beta} > 0$  for small  $\beta \in (0, 1]$ . Therefore, by Theorem 2.1, when  $\varepsilon > 0$  is small enough,  $P_t^*$  has a unique invariant probability measure  $\bar{\mu} \in \mathcal{P}_V$ , such that

$$\mathbf{W}_{\psi, \beta V}(P_t^* \mu, \bar{\mu}) \leq e^{-qt} \mathbf{W}_{\psi, \beta V}(\mu, \bar{\mu}), \quad t \geq 0$$

holds for some constant  $q > 0$ . This completes the proof since

$$C^{-1} \tilde{\mathbf{W}}_V \leq \mathbf{W}_{\psi, \beta V} \leq C \tilde{\mathbf{W}}_V$$

holds for some constant  $C > 1$ . □

## 2.3 Proof of Theorem 2.1

Since  $\psi(r) := \psi(r \wedge l)$  for  $\psi \in \Psi_l$  is not second order differentiable at  $l$ , we introduce the following lemma ensuring Itô's formula for  $\psi$  of a semi-martingale which will be used frequently in the sequence.

**Lemma 2.2.** *Let  $\xi_t$  be a non-negative continuous semi-martingale satisfying*

$$d\xi_t \leq A_t dt + dM_t$$

*for a local martingale  $M_t$  and an integrable adapted process  $A_t$ . Then for any  $\psi \in C^1([0, \infty))$  with  $\psi'$  non-negative and Lipschitz continuous, we have*

$$d\psi(\xi_t) \leq \psi'(\xi_t) A_t dt + \frac{1}{2} \psi''(\xi_t) d\langle M \rangle_t + \psi'(\xi_t) dM_t,$$

where

$$\psi''(r) := \limsup_{s \downarrow r} \limsup_{\varepsilon \downarrow 0} \frac{\psi'(s + \varepsilon) - \psi'(s)}{\varepsilon}, \quad r \geq 0$$

is a bounded measurable function on  $[0, \infty)$ .

*Proof.* By restricting before a stopping time, we may and do assume that  $\xi_t, \int_0^t A_s ds$  and  $M_t$  are bounded processes. For any  $n \geq 1$ , let

$$\psi_n(r) = n \int_0^\infty \psi(r+s) e^{-ns} ds, \quad r \geq 0.$$

Then each  $\psi_n$  is  $C^\infty$ -smooth, with  $\psi'_n \geq 0$ ,  $(\psi_n, \psi'_n) \rightarrow (\psi, \psi')$  locally uniformly,  $\{\|\psi''_n\|_\infty\}_{n \geq 1}$  uniformly bounded, and by Fatou's lemma,

$$\begin{aligned} \limsup_{n \rightarrow \infty} \psi''_n(r) &\leq \limsup_{n \rightarrow \infty} \int_0^\infty \limsup_{\varepsilon \downarrow 0} \frac{\psi'(r+s+\varepsilon) - \psi'(r+s)}{\varepsilon} n e^{-ns} ds \\ &\leq \limsup_{s \downarrow 0} \limsup_{\varepsilon \downarrow 0} \frac{\psi'(r+s+\varepsilon) - \psi'(r+s)}{\varepsilon} = \psi''(r), \quad r \geq 0. \end{aligned}$$

Therefore, by applying Itô's formula to  $\psi_n(\xi_t)$  and letting  $n \rightarrow \infty$ , we finish the proof.  $\square$

**A. The well-posedness.** For any  $T > 0$  and a subspace  $\hat{\mathcal{P}} \subset \mathcal{P}$ , let  $C_w([0, T]; \hat{\mathcal{P}})$  be the class of all continuous maps from  $[0, T]$  to  $\hat{\mathcal{P}}$  under the weak topology.

**Lemma 2.3.** *Assume that for some  $K \in L^1_{loc}([0, \infty); (0, \infty))$*

$$(2.18) \quad L_{t,\mu} V(x) \leq \zeta_t(1 + \mu(V) + V(x)), \quad t \geq 0, x \in \mathbb{R}^d, \mu \in \mathcal{P}_V,$$

$$(2.19) \quad \|\sigma_t \nabla V(x)\| \leq \zeta_t(1 + V(x)), \quad t \geq 0, x \in \mathbb{R}^d,$$

$$(2.20) \quad \begin{aligned} &2\langle b_t(x, \mu) - b_t(y, \nu), x - y \rangle^+ + \|\sigma_t(x) - \sigma_t(y)\|_{HS}^2 \\ &\leq \zeta_t |x - y| \{|x - y| + \mathbf{W}_{\psi, V}(\mu, \nu)\}, \quad t \geq 0, x, y \in \mathbb{R}^d, \mu, \nu \in \mathcal{P}_V. \end{aligned}$$

Then (2.1) is well-posed for distributions in  $\mathcal{P}_V$  with

$$(2.21) \quad \mathbb{E}V(X_t) \leq e^{2 \int_0^t \zeta_s ds} \mathbb{E}V(X_0) \int_0^t \zeta_s e^{2 \int_s^t \zeta_r dr} ds.$$

*Proof.* It is easy to see that (2.21) follows from (2.18) and Itô's formula. To prove the well-posedness for distributions in  $\mathcal{P}_V$ , we adopt a fixed point theorem in distributions. For any  $T > 0$ ,  $\gamma := \mathcal{L}_{X_0} \in \mathcal{P}_V$ , and

$$\mu \in \mathcal{P}_{T,V}^\gamma := \{\mu \in C_w([0, T]; \mathcal{P}_V) : \mu_0 = \gamma\},$$

consider the following SDE

$$(2.22) \quad dX_t^\mu = b_t(X_t^\mu, \mu_t) + \sigma_t(X_t^\mu) dW_t, \quad X_0^\mu = X_0, t \in [0, T].$$

It is well known that the monotone condition (2.10) in  $(H_3)$  implies the well-posedness of this SDE up to life time, while the Lyapunov condition (2.18) implies

$$\sup_{t \in [0, T]} \mathbb{E}[V(X_t^\mu)] < \infty.$$

Then by the continuity of  $X_t^\mu$  in  $t$  we conclude that

$$H(\mu)(\cdot) := \mathcal{L}_{X^\mu} \in C_w([0, T]; \mathcal{P}_V).$$

It remains to prove that  $H$  has a unique fixed point  $\bar{\mu} \in \mathcal{P}_{V,T}$ , so that  $X_t^{\bar{\mu}}$  is the unique solution of (2.1) up to time  $T$ , and by the modified Yamada-Watanabe principle [?, Lemma 2.1], this also implies the weak well-posedness of (2.1) up to time  $T$ .

To prove the existence and uniqueness of the fixed point of  $H$ , we introduce

$$\mathcal{P}_{V,T}^{\gamma,N} := \left\{ \mu \in C_w([0, T]; \mathcal{P}_V) : \mu_0 = \gamma, \sup_{t \in [0, T]} e^{-Nt} \mu_t(V) \leq N(1 + \gamma(V)) \right\}, \quad N \geq 1.$$

Then as  $N \uparrow \infty$ , we have  $\mathcal{P}_{V,T}^{\gamma,N} \uparrow \mathcal{P}_{V,T}^\gamma$  as  $N \uparrow \infty$ . So, it suffices to find  $N_0 \geq 1$  such that for any  $N \geq N_0$ ,  $H\mathcal{P}_{T,V}^{\gamma,N} \subset \mathcal{P}_{T,V}^{\gamma,N}$  and  $H$  has a unique fixed point in  $\mathcal{P}_{T,V}^{\gamma,N}$ . We prove this in the following two steps.

(a) Construction of  $N_0$ . Let

$$c := e^{\int_0^T \zeta_s ds}, \quad N_0 := 3c.$$

By Itô's formula and (2.18), for any  $N \geq N_0$  and  $\mu \in \mathcal{P}_{T,V}^{\gamma,N}$ , we have

$$\begin{aligned} e^{-Nt} \mathbb{E}V(X_t^\mu) &\leq \gamma(V) e^{\int_0^t \zeta_s ds - Nt} + \int_0^t \zeta_s \{1 + N(1 + \gamma(V))\} e^{\int_s^t \zeta_r dr - N(t-s)} ds \\ &\leq c\gamma(V) + 2cN(1 + \gamma(V)) \sup_{t \in [0, T]} \int_0^t e^{-N(t-s)} ds \leq c\gamma(V) + 2c(1 + \gamma(V)) \leq N(1 + \gamma(V)). \end{aligned}$$

So,  $H\mathcal{P}_{T,V}^{\gamma,N} \subset \mathcal{P}_{T,V}^{\gamma,N}$  for  $N \geq N_0$ .

(b) Let  $N \geq N_0$ . It remains to prove that  $H$  is contractive in  $\mathcal{P}_{T,V}^{\gamma,N}$  under

$$\mathbf{W}_{\psi,V,\lambda}(\mu, \nu) := \sup_{t \in [0, T]} e^{-\lambda t} \mathbf{W}_{\psi,V}(\mu_t, \nu_t), \quad \mu, \nu \in \mathcal{P}_{T,V}^{\gamma,N}$$

for large  $\lambda > 0$ .

For  $\mu, \nu \in \mathcal{P}_{T,V}^{\gamma,N}$ , by (2.20) and the Itô-Tanaka formula, we find  $C_0 \in L_{loc}^1([0, \infty); (0, \infty))$  such that

$$d|X_t^\mu - X_t^\nu| \leq C_0(t)(\mathbf{W}_{\psi,\beta V}(\mu_t, \nu_t) + |X_t^\mu - X_t^\nu|)dt + \left\langle \frac{X_t^\mu - X_t^\nu}{|X_t^\mu - X_t^\nu|}, \{\sigma_t(X_t^\mu) - \sigma_t(X_t^\nu)\}dW_t \right\rangle.$$

Since  $\psi \in \Psi_l$ , by extending to the half-line with  $\psi(r) := \psi(r \wedge l)$ , we see that  $\psi'$  is non-negative and Lipschitz continuous. By Lemma 2.2,  $\mu, \nu \in \mathcal{P}_{V,T}^{\gamma,N}$ , and noting that  $\psi'' \leq 0$  when  $\sigma_t$  is non-constant for some  $t \geq 0$ , we find  $C_1 \in L^1([0, T]; (0, \infty))$  such that

$$\begin{aligned} d\psi(|X_t^\mu - X_t^\nu|) &\leq C_1(t)\psi(|X_t^\mu - X_t^\nu|) + \mathbf{W}_{\psi,\beta V}(\mu_t, \nu_t)dt \\ (2.23) \quad &+ \psi'(|X_t^\mu - X_t^\nu|) \left\langle \frac{X_t^\mu - X_t^\nu}{|X_t^\mu - X_t^\nu|}, \{\sigma_t(X_t^\mu) - \sigma_t(X_t^\nu)\}dW_t \right\rangle \end{aligned}$$

holds for  $t \in [0, T]$ .

On the other hand, by (2.18) and  $\mu, \nu \in \mathcal{P}_{V,T}^{\gamma,N}$ , we find a constant  $K(N) > 1$  such that

$$\begin{aligned} d\{V(X_t^\mu) + V(X_t^\nu)\} &\leq \zeta_t \{1 + \mu_t(V) + \nu_t(V) + V(X_t^\mu) + V(X_t^\nu)\} dt \\ &\quad + \langle \sigma_t(X_t^\mu) \nabla V(X_t^\mu) + \sigma_t(X_t^\nu) \nabla V(X_t^\nu), dW_t \rangle \\ &\leq K(N) \zeta_t \{1 + V(X_t^\mu) + V(X_t^\nu)\} dt + \langle \sigma_t(X_t^\mu) \nabla V(X_t^\mu) + \sigma_t(X_t^\nu) \nabla V(X_t^\nu), dW_t \rangle. \end{aligned}$$

Combining this with (2.6), (2.19), and (2.23), we find  $C_2 \in L^1([0, T]; (0, \infty))$  such that

$$\xi_t := \psi(|X_t^\mu - X_t^\nu|)(1 + V(X_t^\mu) + V(X_t^\nu))$$

satisfies

$$d\xi_t \leq C_2(t) [\xi_t + (1 + V(X_t^\mu) + V(X_t^\nu)) \mathbf{W}_{\psi,V}(\mu_t, \nu_t)] dt + dM_t, \quad t \in [0, T]$$

for some local martingale  $M_t$ . Since  $H(\mu), H(\nu) \in \mathcal{P}_{V,T}^{\gamma,N}$  implies

$$\mathbb{E}V(X_t^\mu) + \mathbb{E}V(X_t^\nu) \leq N(1 + \gamma(V))e^{NT} =: D(N) < \infty, \quad t \in [0, T],$$

this and  $\xi_0 = 0$  yield

$$e^{-\lambda t} \mathbb{E} \xi_t \leq (1 + D(N)) \mathbf{W}_{\psi,V,\lambda}(\mu, \nu) \int_0^t C_2(s) e^{\int_s^t (C_2(r) - \lambda) dr} ds, \quad t \in [0, T], \lambda > 0.$$

Noting that  $\lim_{\lambda \rightarrow \infty} \sup_{t \in [0, T]} \int_0^t C_2(s) e^{\int_s^t (C_2(r) - \lambda) dr} ds = 0$ , we conclude that when  $\lambda > 0$  is large enough,

$$e^{-\lambda t} \mathbf{W}_{\psi,V}(H_t(\mu), H_t(\nu)) \leq e^{-\lambda t} \mathbb{E} \xi_t \leq \frac{1}{2} \mathbf{W}_{\psi,V,\lambda}(\mu, \nu), \quad t \in [0, T].$$

Therefore,  $H : \mathcal{P}_{V,T}^{\gamma,N} \rightarrow \mathcal{P}_{V,T}^{\gamma,N}$  is contractive in  $\mathbf{W}_{\psi,T,\lambda}$  for large enough  $\lambda > 0$ .  $\square$

**B. Construction of coupling.** Simply denote

$$\psi_{\beta V}(x, y) := \psi(|x - y|)(1 + \beta V(x) + \beta V(y)), \quad x, y \in \mathbb{R}^d.$$

For  $s \geq 0$  and  $\mu, \nu \in \mathcal{P}_V$ , let  $X_s$  and  $Y_s$  be  $\mathcal{F}_s$ -measurable random variables such that

$$(2.24) \quad \mathcal{L}_{X_s} = P_s^* \mu, \quad \mathcal{L}_{Y_s} = P_s^* \nu, \quad \mathbb{E} \psi_{\beta V}(X_s, Y_s) = \mathbf{W}_{\psi, \beta V}(P_s^* \mu, P_s^* \nu).$$

Let  $B_t^1$  and  $B_t^2$  be two independent  $d$ -dimensional Brownian motions and consider the following SDE:

$$(2.25) \quad dX_t = b_t(X_t, P_t^* \mu) dt + \sqrt{\alpha_t} dB_t^1 + \hat{\sigma}_t(X_t) dB_t^2, \quad t \geq s.$$

By  $(H_3)$ , this SDE is well-posed. Indeed, since  $b$  is locally bounded, by Girsanov's transform to the regular SDE

$$dX_t = \sqrt{\alpha_t} dB_t^1 + \hat{\sigma}_t(X_t) dB_t^2, \quad t \geq s$$

up to the exit time of a large ball, we construct a weak solution to (2.25) up to the same stopping time. On the other hand, the monotone condition in  $(H_3)$  implies the pathwise uniqueness of (2.25), then the well-posedness is implied by the Yamada-Watanabe principle. Moreover, the Lyapunov condition in  $(H_1)$  ensures the non-explosion. Since by  $(H_2)$  and the definition of  $P_t^*$  the solution to the McKean-Vlasov SDE (2.1) is a weak solution to (2.25), the weak uniqueness of (2.25) implies that  $\mathcal{L}_{X_t} = P_t^* \mu, t \geq s$ .

To construct the coupling with reflection, let

$$u(x, y) = \frac{x - y}{|x - y|}, \quad x \neq y \in \mathbb{R}^d.$$

We consider the SDE:

$$(2.26) \quad dY_t = b_t(Y_t, P_t^* \nu) dt + \sqrt{\alpha_t} \{I_d - 2u(X_t, Y_t) \otimes u(X_t, Y_t) 1_{\{t < \tau\}}\} dB_t^1 + \hat{\sigma}_t(Y_t) dB_t^2$$

for  $t \geq s$ , where

$$\tau := \inf\{t \geq s : Y_t = X_t\}$$

is the coupling time. Since the coefficients in noises are Lipschitz continuous in  $Y_t \neq X_t$ , by the same argument leading to the well-posedness of (2.25), we conclude that (2.26) has a unique solution up to the coupling time  $\tau$ . When  $t \geq \tau$ , the equation of  $Y_t$  becomes

$$(2.27) \quad dY_t = b_t(Y_t, P_t^* \nu) dt + \sqrt{\alpha_t} dB_t^1 + \hat{\sigma}_t(Y_t) dB_t^2,$$

which is well-posed as explained above. Therefore, (2.26) has a unique solution up to life time. On the other hand, the Lyapunov condition in  $(H_1)$  implies that the solution is non-explosive, and by the same reason leading to  $\mathcal{L}_{X_t} = P_t^* \mu$ , we have  $\mathcal{L}_{Y_t} = P_t^* \nu$ .

**C. Proof of (2.14).** By  $(H_3)$  and the Itô-Tanaka formula for (2.25) and (2.26), we obtain

$$\begin{aligned} d|X_t - Y_t| &\leq \{\theta_t \hat{\mathbf{W}}_{\psi, \beta V}(P_t^* \mu, P_t^* \nu) + K_t |X_t - Y_t|\} dt \\ &\quad + 2\sqrt{\alpha_t} \langle u(X_t, Y_t), dB_t^1 \rangle + \langle u(X_t, Y_t), (\hat{\sigma}_t(X_t) - \hat{\sigma}_t(Y_t)) dB_t^2 \rangle, \quad t < \tau. \end{aligned}$$

By Lemma 2.2 and noting that  $\psi'' \leq 0$  when  $\hat{\sigma}$  is non-constant, we get

$$\begin{aligned} d\psi(|X_t - Y_t|) &\leq \{\theta_t \psi'(|X_t - Y_t|) \hat{\mathbf{W}}_{\psi, \beta V}(P_t^* \mu, P_t^* \nu) + K_t |X_t - Y_t| \psi'(|X_t - Y_t|) + 2\alpha_t \psi''(|X_t - Y_t|)\} dt \\ &\quad + \psi'(|X_t - Y_t|) \left[ 2\sqrt{\alpha_t} \langle u(X_t, Y_t), dB_t^1 \rangle + \langle u(X_t, Y_t), (\hat{\sigma}_t(X_t) - \hat{\sigma}_t(Y_t)) dB_t^2 \rangle \right], \quad t < \tau. \end{aligned}$$

Therefore, (2.9) yields

$$\begin{aligned} d\psi(|X_t - Y_t|) &\leq \{\theta_t \psi'(|X_t - Y_t|) \hat{\mathbf{W}}_{\psi, \beta V}(P_t^* \mu, P_t^* \nu) - q_l(t) \psi(|X_t - Y_t|) 1_{\{|X_t - Y_t| < l\}}\} dt \\ (2.28) \quad &\quad + \psi'(|X_t - Y_t|) \left[ 2\sqrt{\alpha_t} \langle u(X_t, Y_t), dB_t^1 \rangle \right. \\ &\quad \left. + \langle u(X_t, Y_t), (\hat{\sigma}_t(X_t) - \hat{\sigma}_t(Y_t)) dB_t^2 \rangle \right], \quad t < \tau. \end{aligned}$$

By  $(H_1)$  and Itô's formula, we obtain

$$(2.29) \quad \begin{aligned} d\{V(X_t) + V(Y_t)\} &\leq \{2K_0(t) - K_1(t)V(X_t) - K_1(t)V(Y_t)\}dt \\ &+ \sqrt{\alpha_t}\langle \nabla V(X_t) + \nabla V(Y_t) - 2\langle u(X_t, Y_t), \nabla V(Y_t) \rangle u(X_t, Y_t), dB_t^1 \rangle \\ &+ \langle \hat{\sigma}_t(X_t)^* \nabla V(X_t) + \hat{\sigma}_t(Y_t)^* \nabla V(Y_t), dB_t^2 \rangle. \end{aligned}$$

This together with (2.28) yields that

$$\phi_t := \psi_{\beta V}(X_t, Y_t) = \psi(|X_t - Y_t|)\{1 + \beta V(X_t) + \beta V(Y_t)\}$$

satisfies

$$(2.30) \quad \begin{aligned} d\phi_t &\leq \left\{ \theta_t \psi'(|X_t - Y_t|) \hat{\mathbf{W}}_{\psi, \beta V}(P_t^* \mu, P_t^* \nu) [1 + \beta V(X_t) + \beta V(Y_t)] - q_t(t) \phi_t 1_{\{|X_t - Y_t| < l\}} \right. \\ &+ \beta \psi(|X_t - Y_t|) [2K_0(t) - K_1(t)V(X_t) - K_1(t)V(Y_t)] \\ &+ \beta \psi'(|X_t - Y_t|) \left( \alpha_t |\nabla V(X_t) - \nabla V(Y_t)| \right. \\ &+ \left. \left| \{\hat{\sigma}_t(X_t) - \hat{\sigma}_t(Y_t)\} [\hat{\sigma}_t(X_t)^* \nabla V(X_t) + \hat{\sigma}_t(Y_t)^* \nabla V(Y_t)] \right| \right) \Big\} dt \\ &+ dM_t, \quad t < \tau \end{aligned}$$

for some martingale  $M_t$ . Combining (2.6), (2.11) and (2.12), we derive

$$\begin{aligned} &\beta \psi(|X_t - Y_t|) \{2K_0(t) - K_1(t)V(X_t) - K_1(t)V(Y_t)\} \\ &\leq 2K_0(t) \beta \phi_t 1_{\{|X_t - Y_t| < l\}} - \kappa_{l, \beta}(t) \phi_t 1_{\{|X_t - Y_t| \geq l\}}, \end{aligned}$$

$$\begin{aligned} &\beta \psi'(|X_t - Y_t|) \left\{ \alpha_t |\nabla V(X_t) - \nabla V(Y_t)| \right. \\ &+ \left. \left| \{\hat{\sigma}_t(X_t) - \hat{\sigma}_t(Y_t)\} [\hat{\sigma}_t(X_t)^* \nabla V(X_t) + \hat{\sigma}_t(Y_t)^* \nabla V(Y_t)] \right| \right\} \leq \alpha_{l, \beta}(t) \phi_t 1_{\{|X_t - Y_t| < l\}}. \end{aligned}$$

Hence, it follows from (2.30) that

$$\begin{aligned} d\phi_t &\leq \theta_t \psi'(|X_t - Y_t|) \hat{\mathbf{W}}_{\psi, \beta V}(P_t^* \mu, P_t^* \nu) \{1 + \beta V(X_t) + \beta V(Y_t)\} dt \\ &- \{[q_t(t) - \alpha_{l, \beta}(t) - 2K_0(t)\beta] \phi_t 1_{\{|X_t - Y_t| < l\}} + \kappa_{l, \beta}(t) \phi_t 1_{\{|X_t - Y_t| \geq l\}}\} dt + dM_t \\ &\leq \{\theta \psi'(|X_t - Y_t|) \hat{\mathbf{W}}_{\psi, \beta V}(P_t^* \mu, P_t^* \nu) \{1 + \beta V(X_t) + \beta V(Y_t)\} - \lambda_{l, \beta}(t) \phi_t\} dt + dM_t, \quad t < \tau. \end{aligned}$$

Since  $\phi_{t \wedge \tau} = 0$  for  $t \geq \tau$ , this implies

$$\begin{aligned} &e^{\int_0^t \lambda_{l, \beta}(s) ds} \mathbb{E} \phi_{t \wedge \tau} = \mathbb{E} [\phi_{t \wedge \tau} e^{\int_0^{t \wedge \tau} \lambda_{l, \beta}(s) ds}] \leq e^{\int_0^s \lambda_{l, \beta}(r) dr} \mathbb{E} \phi_s \\ &+ \mathbb{E} \int_s^{t \wedge \tau} e^{\int_0^r \lambda_{l, \beta}(p) dp} \theta_r \psi'(|X_t - Y_t|) \hat{\mathbf{W}}_{\psi, \beta V}(P_r^* \mu, P_r^* \nu) \{1 + \beta V(X_r) + \beta V(Y_r)\} dr, \quad t \geq s. \end{aligned}$$

Therefore, for any  $t \geq s$ , we have

$$(2.31) \quad \begin{aligned} &\mathbb{E} \phi_{t \wedge \tau} \leq e^{-\int_s^t \lambda_{l, \beta}(r) dr} \mathbb{E} \phi_s \\ &+ e^{\int_s^t \lambda_{l, \beta}(r) dr} \mathbb{E} \int_s^{t \wedge \tau} \theta_r \hat{\mathbf{W}}_{\psi, \beta V}(P_r^* \mu, P_r^* \nu) \psi'(|X_t - Y_t|) \{1 + \beta V(X_r) + \beta V(Y_r)\} dr. \end{aligned}$$

On the other hand, for  $t \geq \tau$ , by Itô's formula for (2.25) and (2.27), and applying (2.10), we find  $C_1 \in L^1_{loc}([0, \infty); (0, \infty))$  such that

$$\begin{aligned} d\psi(|X_t - Y_t|) \leq & \{C_1(t)\psi(|X_t - Y_t|) + \theta_t\psi'(|X_t - Y_t|)\hat{\mathbf{W}}_{\psi, \beta V}(P_t^*\mu, P_t^*\nu)\}dt \\ & + \psi'(|X_t - Y_t|)\langle\{\hat{\sigma}_t(X_t) - \hat{\sigma}_t(Y_t)\}^*u(X_t, Y_t), dB_t^2\rangle. \end{aligned}$$

Combining this with (2.29), we find  $C_2 \in L^1_{loc}([0, \infty); (0, \infty))$  such that

$$d\phi_t \leq \left\{C_2(t)\phi_t + \theta_t\hat{\mathbf{W}}_{\psi, \beta V}(P_t^*\mu, P_t^*\nu)\psi'(|X_t - Y_t|)\{1 + \beta V(X_t) + \beta V(Y_t)\}\right\}dt + dM_t, \quad t \geq \tau$$

for some martingale  $M_t$ . Therefore, for any  $t \geq s$ , we have  $t \wedge \tau \geq s$  so that

$$\begin{aligned} & \mathbb{E}[1_{\{t > \tau\}}(\phi_t - \phi_{t \wedge \tau})] \\ & \leq \mathbb{E} \int_{t \wedge \tau}^t e^{\int_r^t C_2(p)dp} \theta_r \hat{\mathbf{W}}_{\psi, \beta V}(P_r^*\mu, P_r^*\nu) \psi'(|X_r - Y_r|) \{1 + \beta V(X_r) + \beta V(Y_r)\} dr \\ & \leq e^{\int_s^t C_2(p)dp} \mathbb{E} \int_{t \wedge \tau}^t \theta_r \hat{\mathbf{W}}_{\psi, \beta V}(P_r^*\mu, P_r^*\nu) \psi'(|X_r - Y_r|) \{1 + \beta V(X_r) + \beta V(Y_r)\} dr. \end{aligned}$$

This together with (2.31), (2.24) and (2.7) yields

$$\begin{aligned} \mathbb{E}\phi_t &= \mathbb{E}\phi_{t \wedge \tau} + \mathbb{E}[1_{\{t > \tau\}}(\phi_t - \phi_{t \wedge \tau})] \leq e^{-\int_0^t \lambda_{l, \beta}(r)dr} \mathbb{E}\phi_s \\ &+ e^{\int_s^t (|\lambda_{l, \beta}| + C_2)(r)dr} \mathbb{E} \int_s^t \theta_r \hat{\mathbf{W}}_{\psi, \beta V}(P_r^*\mu, P_r^*\nu) \psi'(|X_r - Y_r|) \{1 + \beta V(X_r) + \beta V(Y_r)\} dr \\ &\leq e^{-\int_s^t \lambda_{l, \beta}(r)dr} \mathbf{W}_{\psi, \beta V}(P_s^*\mu, P_s^*\nu) + e^{\int_s^t (2|\lambda_{l, \beta}| + C_2)(r)dr} \int_s^t \theta_r e^{\int_s^r \lambda_{l, \beta}(p)dp} \mathbb{E}\phi_r dr, \quad t \geq s, \end{aligned}$$

where the last step follows from the definition of  $\hat{\mathbf{W}}_{\psi, \beta V}$  which implies

$$\hat{\mathbf{W}}_{\psi, \beta V}(P_r^*\mu, P_r^*\nu) \leq \frac{\mathbb{E}\phi_r}{\mathbb{E}[\psi'(|X_r - Y_r|)\{1 + \beta V(X_r) + \beta V(Y_r)\}]}. \quad \square$$

By Gronwall's lemma, we obtain

$$e^{\int_s^t \lambda_{l, \beta}(r)dr} \mathbb{E}\phi_t \leq \mathbf{W}_{\psi, \beta V}(P_s^*\mu, P_s^*\nu) \exp \left[ e^{\int_s^t \{2|\lambda_{l, \beta}(r)| + C_2(r)\}dr} \int_s^t \theta_r dr \right], \quad t \geq s.$$

Thus, for a.e.  $s \geq 0$ ,

$$\begin{aligned} \frac{d^+}{ds} \mathbf{W}_{\psi, \beta V}(P_s^*\mu, P_s^*\nu) &:= \limsup_{t \downarrow s} \frac{\mathbf{W}_{\psi, \beta V}(P_t^*\mu, P_t^*\nu) - \mathbf{W}_{\psi, \beta V}(P_s^*\mu, P_s^*\nu)}{t - s} \\ &\leq \limsup_{t \downarrow s} \frac{\mathbb{E}\phi_t - \mathbf{W}_{\psi, \beta V}(P_s^*\mu, P_s^*\nu)}{t - s} \\ &\leq -(\lambda_{l, \beta}(s) - \theta_s) \mathbf{W}_{\psi, \beta V}(P_s^*\mu, P_s^*\nu). \end{aligned}$$

This implies (2.14).

**D. Proof of (2.15).** Let  $a, b$  be independent of the time parameter and

$$\kappa := \lambda_{l,\beta} - \theta > 0.$$

We intend to show that  $P_t^*$  has an invariant probability measure  $\bar{\mu} \in \mathcal{P}_V$ , so that (2.14) implies (2.15) and the uniqueness of the invariant probability measure. This can be done as in the proof of [23, Theorem 3.1(2)] by verifying that  $P_t^* \delta_0$  converges in  $\mathcal{P}_V$  under  $\mathbf{W}_{\psi, \beta V}$  as  $t \rightarrow \infty$ , where  $\delta_0$  is the Dirac measure at  $0 \in \mathbb{R}^d$ . Precisely, by (2.14) and the semigroup property  $P_{t+s}^* = P_t^* P_s^*$  for  $s, t \geq 0$  due to (2.2), we have

$$\begin{aligned} \sup_{s \geq 0} \mathbf{W}_{\psi, \beta V}(P_t^* \delta_0, P_{t+s}^* \delta_0) &\leq e^{-\kappa t} \sup_{s \geq 0} \mathbb{E}^0[\psi(|X_s|)\{1 + \beta V(0) + \beta V(X_s)\}] \\ &\leq \|\psi\|_\infty e^{-\kappa t} \left\{ 1 + \beta V(0) + \sup_{s \geq 0} \mathbb{E}^0 V(X_s) \right\}, \quad t \geq 0, \end{aligned}$$

where  $\mathbb{E}^0$  is the expectation taken for the solution to (1.7) with  $X_0 = 0$ . Since (2.5) yields

$$\sup_{s \geq 0} \mathbb{E}^0[V(X_s)] \leq V(0) + \frac{K_0}{K_1} < \infty,$$

we arrive at

$$\lim_{t \rightarrow \infty} \sup_{s \geq 0} \mathbf{W}_{\psi, \beta V}(P_t^* \delta_0, P_{t+s}^* \delta_0) = 0,$$

so that when  $t \rightarrow \infty$ ,  $P_t^* \delta_0$  converges to a probability measure  $\bar{\mu} \in \mathcal{P}_V$ , which is an invariant measure of  $P_t^*$ . Indeed, in this case the semigroup property and (2.14) imply

$$\mathbf{W}_{\psi, \beta V}(P_s^* \bar{\mu}, \bar{\mu}) = \lim_{t \rightarrow \infty} \mathbf{W}_{\psi, \beta V}(P_s^* \bar{\mu}, P_s^* P_t^* \delta_0) \leq \lim_{t \rightarrow \infty} \mathbf{W}_{\psi, \beta V}(\bar{\mu}, P_t^* \delta_0) = 0, \quad s \geq 0.$$

### 3 Under dissipative condition in long distance

For any  $\psi \in \Psi$ , where

$$\Psi := \left\{ \psi \in C^2([0, \infty)) : \psi(0) = 0, \psi' > 0, r\psi'(r) + r^2(\psi'')^+(r) \leq cr \text{ for some constant } c > 0 \right\},$$

the quasi-distance

$$\mathbf{W}_\psi(\mu, \nu) := \inf_{\pi \in \mathcal{C}(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \psi(|x - y|) \pi(dx, dy)$$

on the space

$$\mathcal{P}_\psi := \left\{ \mu \in \mathcal{P} : \|\mu\|_\psi := \mu(\psi(|\cdot|)) < \infty \right\}$$

is complete, i.e. a  $\mathbf{W}_\psi$ -Cauchy sequence in  $\mathcal{P}_\psi$  converges with respect to  $\mathbf{W}_\psi$ . When  $\psi$  is concave,  $\mathbf{W}_\psi$  satisfies the triangle inequality and is hence a metric on  $\mathcal{P}_\psi$ .

In this part, we do not assume the Lyapunov condition ( $H_1$ ) but use the following condition to replace ( $H_3$ ).

( $H'_3$ ) ( **$\psi$ -Monotonicity**) Let  $\psi \in \Psi$ ,  $\gamma \in C([0, \infty))$  with  $\gamma(r) \leq Kr$  for some constant  $K > 0$  and all  $r \geq 0$ , such that

$$(3.1) \quad 2\alpha_t \psi''(r) + (\gamma \psi')(r) \leq -q_t \psi(r), \quad r \geq 0$$

holds for some  $q \in L^1_{loc}([0, \infty); (0, \infty))$ . Moreover,  $b$  is locally bounded on  $[0, \infty) \times \mathbb{R}^d \times \mathcal{P}_\psi$ , and there exists  $\theta \in L^1_{loc}([0, \infty); (0, \infty))$  such that

$$(3.2) \quad \begin{aligned} & \langle b_t(x, \mu) - b_t(y, \nu), x - y \rangle + \frac{1}{2} \|\hat{\sigma}_t(x) - \hat{\sigma}_t(y)\|_{HS}^2 \\ & \leq |x - y| \{ \theta_t \mathbf{W}_\psi(\mu, \nu) + \gamma(|x - y|) \}, \quad t \geq 0, x, y \in \mathbb{R}^d, \mu, \nu \in \mathcal{P}_\psi. \end{aligned}$$

When  $a = I_d$  and

$$b(x, \mu) = b_0(x) + \int_{\mathbb{R}^d} Z(x, y) \mu(dy)$$

for a drift  $b_0$  and a Lipschitz continuous map  $Z : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ , the exponential convergence of (2.1) is presented in [9, Theorems 2.3 and 2.4] under the condition that

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

for some function  $\kappa \in C((0, \infty))$  with  $\int_0^1 r \kappa^+(r) dr < \infty$  and  $\limsup_{r \rightarrow \infty} \kappa(r) < 0$ , and that the Lipschitz constant of  $Z$  is small enough. It is clear that in this case (3.2) holds for  $\gamma(r) := r \kappa(r)$  and  $\psi(r)$  comparable with  $r$ , for which we may choose  $\psi \in \Psi$  as in (3.15) below such that (3.1) holds for  $\alpha = 1$  and some  $q > 0$ . Therefore, this situation is included in Theorem 3.1 below.

### 3.1 Main results and example

**Theorem 3.1.** *Assume  $(H_2)$  and  $(H'_3)$ , with  $\psi'' \leq 0$  if  $\hat{\sigma}_t(\cdot)$  is non-constant for some  $t \geq 0$ . Then (2.1) is well-posed in  $\mathcal{P}_\psi$ , and  $P_t^*$  satisfies*

$$(3.3) \quad \mathbf{W}_\psi(P_t^* \mu, P_t^* \nu) \leq e^{-\int_0^t \{q_s - \theta_s \|\psi'\|_\infty\} ds} \mathbf{W}_\psi(\mu, \nu), \quad t \geq 0, \mu, \nu \in \mathcal{P}_\psi.$$

Consequently,  $(b_t, a_t)$  does not depend on  $t$ ,  $q > \theta \|\psi'\|_\infty$ , and

$$(3.4) \quad \sup_{t \geq 0} \|P_t^* \delta_0\|_\psi < \infty$$

which is the case when  $\psi'' \leq 0$ , then  $P_t^*$  has a unique invariant probability measure  $\bar{\mu} \in \mathcal{P}_\psi$  such that

$$(3.5) \quad \mathbf{W}_\psi(P_t^* \mu, \bar{\mu}) \leq e^{-(q - \theta \|\psi'\|_\infty)t} \mathbf{W}_\psi(\mu, \bar{\mu}), \quad t \geq 0, \mu \in \mathcal{P}_\psi.$$

*Proof.* By  $(H_2)$  and  $(H'_3)$ , the well-posedness follows from the proof of Lemma 2.3 with  $\mathbf{W}_\psi$  replacing  $\mathbf{W}_{\psi, V}$ , and the solution satisfies

$$(3.6) \quad \sup_{t \in [0, T]} \|P_t^* \mu\|_\psi < \infty, \quad \mu \in \mathcal{P}_\psi, T > 0.$$

We omit the details to save space. It remains to prove (3.3) and the existence of the invariant probability measure  $\bar{\mu}$  in the time homogeneous case.

(1) Proof of (3.3). Let  $s \geq 0$  and  $\mu, \nu \in \mathcal{P}_\psi$ . We make use the coupling constructed by (2.25) and (2.26) for initial values  $(X_s, Y_s)$  satisfying

$$(3.7) \quad \mathcal{L}_{X_s} = P_s^* \mu, \quad \mathcal{L}_{Y_s} = P_s^* \nu, \quad \mathbf{W}_\psi(P_s^* \mu, P_s^* \nu) = \mathbb{E}\psi(X_s, Y_s).$$

By the same reason leading to (2.28), by  $(H'_3)$  for  $\psi \in \Psi$  with  $\psi'' \leq 0$  when  $\hat{\sigma}$  is non-constant, we derive

$$(3.8) \quad \begin{aligned} d\psi(|X_t - Y_t|) \leq & \{ \theta \psi'(|X_t - Y_t|) \mathbf{W}_\psi(P_t^* \mu, P_t^* \nu) - q \psi(|X_t - Y_t|) \} dt \\ & + \psi'(|X_t - Y_t|) \left[ 2\sqrt{\lambda} \left\langle u(X_t, Y_t), dB_t^1 \right\rangle \right. \\ & \left. + \left\langle u(X_t, Y_t), (\hat{\sigma}_t(X_t) - \hat{\sigma}_t(Y_t)) dB_t^2 \right\rangle \right], \quad t < \tau. \end{aligned}$$

By the same argument leading to (2.31), this implies

$$(3.9) \quad \mathbb{E}\psi(|X_{t \wedge \tau} - Y_{t \wedge \tau}|) \leq e^{-q(t-s)} \mathbb{E}\psi(|X_s - Y_s|) + \theta \|\psi'\|_\infty \int_s^{t \wedge \tau} \mathbf{W}_\psi(P_r^* \mu, P_r^* \nu) dr, \quad t \geq s.$$

On the other hand, when  $t \geq \tau$ , by  $(H'_3)$  and applying Itô's formula for (2.25) and (2.27), we find a constant  $C > 0$  such that

$$\begin{aligned} d\psi(|X_t - Y_t|) \leq & \{ C \psi(|X_t - Y_t|) dt + \theta \|\psi'\|_\infty \mathbf{W}_\psi(P_t^* \mu, P_t^* \nu) \} dt \\ & + \psi'(|X_t - Y_t|) \langle \{\hat{\sigma}_t(X_t) - \hat{\sigma}_t(Y_t)\}^* u(X_t, Y_t), dB_t^2 \rangle. \end{aligned}$$

Thus,

$$\mathbb{E}[1_{\{t > \tau\}} \psi(|X_t - Y_t|)] \leq \theta \|\psi'\|_\infty e^{C(t-s)} \mathbb{E} \int_{t \wedge \tau}^t \mathbf{W}_\psi(P_r^* \mu, P_r^* \nu) dr, \quad t \geq s.$$

Combining this with (3.9) and (3.7), we derive

$$\begin{aligned} \mathbf{W}_\psi(P_t^* \mu, P_t^* \nu) & \leq \mathbb{E}\psi(|X_t - Y_t|) = \mathbb{E}\psi(|X_{t \wedge \tau} - Y_{t \wedge \tau}|) + \mathbb{E}[1_{\{t > \tau\}} \psi(|X_t - Y_t|)] \\ & \leq e^{-q(t-s)} \mathbb{E}\psi(|X_s - Y_s|) + \theta \|\psi'\|_\infty e^{C(t-s)} \int_s^t \mathbf{W}_\psi(P_r^* \mu, P_r^* \nu) dr \\ & = e^{-q(t-s)} \mathbf{W}_\psi(P_s^* \mu, P_s^* \nu) + \theta \|\psi'\|_\infty e^{C(t-s)} \int_s^t \mathbf{W}_\psi(P_r^* \mu, P_r^* \nu) dr, \quad t \geq s. \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{d^+}{ds} \mathbf{W}_\psi(P_s^* \mu, P_s^* \nu) & := \limsup_{t \downarrow s} \frac{\mathbf{W}_\psi(P_t^* \mu, P_t^* \nu) - \mathbf{W}_\psi(P_s^* \mu, P_s^* \nu)}{t - s} \\ & \leq -(q - \theta \|\psi'\|_\infty) \mathbf{W}_\psi(P_s^* \mu, P_s^* \nu), \quad s \geq 0. \end{aligned}$$

This implies (3.3).

(2) Existence of  $\bar{\mu} \in \mathcal{P}_\psi$ . Let  $(a_t, b_t)$  do not depend on  $t$ , and

$$\lambda := q - \theta \|\psi'\|_\infty > 0.$$

Then (3.3) implies

$$\mathbf{W}_\psi(P_t^* \delta_0, P_{t+s}^* \delta_0) \leq e^{-\lambda t} \mathbf{W}_\psi(\delta_0, P_s^* \delta_0), \quad t, s \geq 0.$$

Combining this with (2.30) we see that as  $t \rightarrow \infty$ ,  $\{P_t^* \delta_0\}$  is a  $\mathbf{W}_\psi$ -Cauchy family whose limit is an invariant probability measure of  $P_t^*$ . It remains to show that (3.4) follows from  $\psi'' \leq 0$  and (3.3). Indeed, in this case  $\mathbf{W}_\psi$  satisfies the triangle inequality so that, for  $n$  being the integer part of  $t > 1$ , (3.3) and (3.6) imply

$$\begin{aligned} \|P_t^* \delta_0\|_\psi &= \mathbf{W}_\psi(\delta_0, P_t^* \delta_0) \leq \sum_{k=0}^{n-1} \mathbf{W}_\psi(P_k^* \delta_0, P_{k+1}^* \delta_0) + \mathbf{W}_\psi(P_n^* \delta_0, P_t^* \delta_0) \\ &\leq \sum_{k=0}^{n-1} e^{-\lambda k} \|P_1^* \delta_0\|_\psi + e^{-n\lambda} \sup_{s \in [0,1]} \|P_s^* \delta_0\|_\psi \leq \left( \sup_{s \in [0,1]} \|P_s^* \delta_0\|_\psi \right) \sum_{k=0}^{\infty} e^{-\lambda k} < \infty. \end{aligned}$$

Therefore, (3.4) holds.  $\square$

As a consequence of Theorem 3.1, we consider the non-dissipative case where  $\nabla b_t(\cdot, \mu)(x)$  is positive definite in a possibly unbounded set but with bounded “one-dimensional puncture mass” in the sense of (3.12) below. Let  $\mathcal{P}_1 = \{\mu \in \mathcal{P} : \mu(|\cdot|) < \infty\}$  and

$$S_b(x) := \sup \{ \langle \nabla_v b_t(\cdot, \mu)(x), v \rangle : t \geq 0, |v| \leq 1, \mu \in \mathcal{P}_1 \}, \quad x \in \mathbb{R}^d.$$

( $H_3''$ ) There exist constants  $\theta_0, \theta_1, \theta_2, \alpha \geq 0$  such that

$$(3.10) \quad \frac{1}{2} \|\hat{\sigma}_t(x) - \hat{\sigma}_t(y)\|_{HS}^2 \leq \theta_0 |x - y|^2, \quad t \geq 0, x, y \in \mathbb{R}^d;$$

$$(3.11) \quad S_b(x) \leq \theta_1, \quad |b_t(x, \mu) - b_t(x, \nu)| \leq \varphi \mathbf{W}_1(\mu, \nu), \quad t \geq 0, x \in \mathbb{R}^d, \mu, \nu \in \mathcal{P}_1;$$

$$(3.12) \quad \kappa := \sup_{x, v \in \mathbb{R}^d, |v|=1} \int_{\mathbb{R}} 1_{\{S_b(x+sv) > -\theta_2\}} ds < \infty.$$

Let  $\mathbf{W}_1 = \mathbf{W}_\psi$  and  $\mathcal{P}_1 = \mathcal{P}_\psi$  for  $\psi(r) = r$ .

**Corollary 3.2.** *Assume  $(H_2)$  and  $(H_3'')$ . Let*

$$(3.13) \quad \begin{aligned} \gamma(r) &:= (\theta_1 + \theta_2) \{ (\kappa r^{-1}) \wedge r \} - (\theta_2 - \theta_0) r, \quad r \geq 0, \\ k &:= \frac{2\lambda}{\int_0^\infty t e^{\frac{1}{2\lambda} \int_0^t \gamma(u) du} dt} - \frac{\varphi(\theta_2 - \theta_0)}{2\lambda} \int_0^\infty t e^{\frac{1}{2\lambda} \int_0^t \gamma(u) du} dt. \end{aligned}$$

Then there exists a constant  $c > 0$  such that

$$\mathbf{W}_1(P_t^* \mu, P_t^* \nu) \leq ce^{-kt} \mathbf{W}_1(\mu, \nu), \quad t \geq 0, \mu, \nu \in \mathcal{P}_1.$$

If  $\theta_2 > \theta_0$  and

$$(3.14) \quad \varphi < \frac{4\lambda^2}{(\theta_2 - \theta_0)(\int_0^\infty t e^{\frac{1}{2\lambda} \int_0^t \gamma(u) du} dt)^2},$$

then  $\kappa > 0$  and  $P_t^*$  has a unique invariant probability measure  $\bar{\mu} \in \mathcal{P}_1$  satisfying

$$\mathbf{W}_1(P_t^* \mu, \bar{\mu}) \leq ce^{-kt} \mathbf{W}_1(\mu, \bar{\mu}), \quad t \geq 0, \mu \in \mathcal{P}_1.$$

*Proof.* For  $\gamma$  in (3.13), let

$$q := \frac{2\lambda}{\int_0^\infty t e^{\frac{1}{2\lambda} \int_0^t \gamma(u) du} dt}, \quad \theta := \frac{\varphi(\theta_2 - \theta_0)}{2\lambda} \int_0^\infty t e^{\frac{1}{2\lambda} \int_0^t \gamma(u) du} dt,$$

and take

$$(3.15) \quad \psi(r) := \int_0^r e^{-\frac{1}{2\lambda} \int_0^s \gamma(u) du} \int_s^\infty t e^{\frac{1}{2\lambda} \int_0^t \gamma(u) du} dt, \quad r \geq 0.$$

By Theorem 3.1, it suffices to verify

- (a)  $\psi \in \Psi$  and  $\psi'' \leq 0$ ;
  - (b) there exists a constant  $C > 1$  such that  $C^{-1} \mathbf{W}_\psi \leq \mathbf{W}_1 \leq C \mathbf{W}_\psi$ ;
  - (c) (3.1) and (3.2) hold.
- (a) We have  $\psi(0) = 0, \psi'(r) > 0$  and

$$(3.16) \quad \psi''(r) = -\frac{\gamma(r)}{2\lambda} e^{-\frac{1}{2\lambda} \int_0^r \gamma(u) du} \int_r^\infty t e^{\frac{1}{2\lambda} \int_0^t \gamma(u) du} dt - r, \quad r \geq 0.$$

To prove  $\psi \in \Psi$ , it suffices to show  $\psi'' \leq 0$ . To this end, take

$$r_0 := \frac{\sqrt{\kappa(\theta_1 + \theta_2)}}{\sqrt{\theta_2 - \theta_0}}.$$

It is easy to see that  $\gamma$  in (3.13) satisfies

$$(3.17) \quad \gamma|_{[0, r_0]} \geq 0, \quad \gamma|_{(r_0, \infty)} < 0.$$

Combining this with (3.16) we have  $\psi''(r) \leq 0$  for  $r \leq r_0$ . On the other hand, for  $r > r_0$  we have  $\gamma(r) < 0$  and

$$\frac{r}{-\gamma(r)} = \frac{1}{(\theta_2 - \theta_0)r^{1-p} - (\theta_1 + \theta_2)\kappa r^{-(1+p)}}$$

is decreasing in  $r > r_0$ , so that

$$\begin{aligned} \int_r^\infty t e^{\frac{1}{2\lambda} \int_0^t \gamma(u) du} dt &= \int_r^\infty \frac{2\lambda t}{\gamma(t)} \left( \frac{d}{dt} e^{\frac{1}{2\lambda} \int_0^t \gamma(u) du} \right) dt \\ &= -\frac{2\lambda r}{\gamma(r)} e^{\frac{1}{2\lambda} \int_0^r \gamma(u) du} + 2\lambda \int_r^\infty \left( \frac{d}{dt} \frac{2\lambda t}{\gamma(t)} \right) e^{\frac{1}{2\lambda} \int_0^t \gamma(u) du} dt \leq -\frac{2\lambda r}{\gamma(r)} e^{\frac{1}{2\lambda} \int_0^r \gamma(u) du}, \quad r > r_0. \end{aligned}$$

This together with (3.16) yields  $\psi''(r) \leq 0$  for  $r > r_0$ . In conclusion,  $\psi \in \Psi$ .

(b) Since  $\psi \in \Psi$  with  $\psi'' \leq 0$  implies that  $\psi(r) \leq \psi'(0)r$  and  $\frac{\psi(r)}{r}$  is decreasing in  $r > 0$ , we have  $\mathbf{W}_\psi \leq \psi'(0)\mathbf{W}_1$  and

$$\begin{aligned} \inf_{r>0} \frac{\psi(r)}{r} &= \lim_{r \rightarrow \infty} \frac{\psi(r)}{r} = \lim_{r \rightarrow \infty} \psi'(r) \\ (3.18) \quad &= \lim_{r \rightarrow \infty} \frac{\int_r^\infty t \exp[\frac{1}{2\lambda} \int_0^t \gamma(u) du] dt}{\exp[\frac{1}{2\lambda} \int_0^r \gamma(u) du]} \\ &= \lim_{r \rightarrow \infty} \frac{2\lambda r}{-\gamma(r)} = \frac{2\lambda}{\theta_2 - \theta_0} \in (0, \infty). \end{aligned}$$

Thus,

$$\frac{1}{\psi'(0)} \mathbf{W}_\psi \leq \mathbf{W}_1(\mu, \nu) \leq \frac{\theta_2 - \theta_0}{2\lambda} \mathbf{W}_\psi.$$

(c) By (3.13) we have

$$2\lambda\psi''(r) + \gamma(r)\psi'(r) = -2\lambda r, \quad r \geq 0.$$

Since  $\psi(r) \leq \psi'(0)r$ , this implies

$$2\lambda\psi''(r) + \gamma(r)\psi'(r) \leq -\frac{2\lambda r}{\psi'(0)r} \psi(r) = -q\psi(r), \quad r \geq 0.$$

Therefore, (3.1) holds.

Next, for  $x \neq y$ , let  $v = \frac{x-y}{|x-y|}$ . Then (3.11) implies

$$\begin{aligned} &\langle b_t(x, \mu) - b_t(y, \nu), x - y \rangle \\ &= |x - y| \langle b_t(x, \mu) - b_t(y, \mu), v \rangle + |x - y| \langle b_t(y, \mu) - b_t(y, \nu), v \rangle \\ (3.19) \quad &\leq \varphi |x - y| \mathbf{W}_1(\mu, \nu) + |x - y| \int_0^{|x-y|} S_b(y + s(x - y)) ds \\ &= \varphi |x - y| \mathbf{W}_1(\mu, \nu) + \int_0^{|x-y|^2} S_b(y + sv) ds, \quad \mu, \nu \in \mathcal{P}_1. \end{aligned}$$

On the other hand, by (3.11) and (3.12) we obtain

$$\int_0^{|x-y|^2} S_b(y + sv) ds \leq \theta_1 \int_0^{|x-y|^2} 1_{\{S_b(x+sv) > -\theta_2\}} ds - \theta_2 \int_0^{|x-y|^2} 1_{\{S_b(x+sv) \leq -\theta_2\}} ds$$

$$= (\theta_1 + \theta_2) \int_0^{|x-y|^2} 1_{\{S_b(x+sv) > -\theta_2\}} ds - \theta_2 |x-y|^2 \leq (\theta_1 + \theta_2)(\kappa \wedge |x-y|^2) - \theta_2 |x-y|^2.$$

Combining this with (3.10) and (3.19), we derive (3.2).  $\square$

To illustrate Corollary 3.2, we consider the following nonlinear PDE for probability density functions  $(\rho_t)_{t \geq 0}$  on  $\mathbb{R}^d$ :

$$(3.20) \quad \partial_t \rho_t = \frac{1}{2} \left\{ \operatorname{div}(a \nabla \rho_t) + \sum_{i,j=1}^d \partial_j [\rho_t \partial_i a_{ij}] \right\} + \operatorname{div} \{ \rho_t \nabla (G + W \circledast \rho_t) \},$$

where  $G \in C^2(\mathbb{R}^d)$ ,  $W \in C^2(\mathbb{R}^d \times \mathbb{R}^d)$  and

$$(3.21) \quad W \circledast \rho_t := \int_{\mathbb{R}^d} W(\cdot, y) \rho_t(y) dy.$$

According to the correspondence between the nonlinear Fokker-Planck equation (1.5) and the McKean-Vlasov SDE (2.1), the exponential ergodicity of  $\mu_t(dx) := \rho_t(x)dx$  is equivalent to that of  $P_t^*$  associated with (2.1) for

$$(3.22) \quad b(x, \mu) := -\nabla G(x) - \int_{\mathbb{R}^d} \{ \nabla W(\cdot, z)(x) \} \mu(dz), \quad x \in \mathbb{R}^d, \mu \in \mathcal{P}_1.$$

When  $a = I_d$  and  $W$  is symmetric (i.e.  $W(x, y) = W(y, x)$ ), the exponential ergodicity in  $\mathbf{W}_2$  is derived for  $\nabla^2 \geq \lambda I_d$  for some  $\lambda > 0$  and a class of  $W$  with locally Lipschitz continuous  $\nabla W$ , the exponential ergodicity in the mean field entropy has been investigated in [11] under the dissipative condition in long distance for small enough  $\|\nabla_x \nabla_y W\|_\infty$ , see also [3] for a special setting, [17] for the exponential ergodicity in  $\mathbf{W}_1$ , [10] for the exponential convergence in the total variation norm, and [20] for the exponential ergodicity in relative entropy. In the following example, we consider the exponential ergodicity in  $\mathbf{W}_1$  for possibly non-constant  $a$ . Indeed, Corollary 3.2 also applies to granular type equations with non-constant diffusion coefficients.

**Example 3.1.** Let  $a$  satisfy  $(H_2)$  with  $\hat{\sigma}$  satisfying (3.10). Consider (3.20) with  $G \in C^2(\mathbb{R}^d)$  and  $W \in C^2(\mathbb{R}^d \times \mathbb{R}^d)$  such that

$$(3.23) \quad \begin{aligned} \nabla_{G+W(\cdot, z)}^2 &\geq \theta_2 1_{\{|\cdot| \geq \lambda_0\}} - \theta_1 1_{\{|\cdot| < \lambda_0\}}, \quad z \in \mathbb{R}^d, \\ \|\nabla_x \nabla_y W(x, y)\| &\leq \tilde{\theta}, \quad x, y \in \mathbb{R}^d \end{aligned}$$

holds for some constants  $\lambda_0, \theta_1, \theta_2 > 0$ . Then the assertion in Corollary 3.2 holds for  $\kappa = 4\lambda_0$  and  $(P_t^* \mu)(dx) := \rho_t(x)dx$ , where  $\rho_t$  solves (3.20) with  $\rho_0(x)dx \in \mathcal{P}_1$ .

*Proof.* It is easy to see that (3.23) implies (3.11). So, it remains to verify that  $\kappa$  in (3.12) satisfies  $\kappa \leq 4\lambda_0$ . By the second inequality in (3.23) we have

$$S_b(x) \leq -\theta_2 1_{\{|x| \geq \lambda_0\}} + \theta_1 1_{\{|x| < \lambda_0\}}, \quad x \in \mathbb{R}^d.$$

For  $x, v \in \mathbb{R}^d$  with  $|v| = 1$ , if there exists  $s_0 \in \mathbb{R}^d$  such that  $|x + s_0 v| < \lambda_0$ , then

$$|x + sv| \geq |s - s_0| - |x + s_0 v| > |s - s_0| - \lambda_0.$$

so that

$$\{s \in \mathbb{R} : |x + sv| < \lambda_0\} \subset (s_0 - 2\lambda_0, s_0 + 2\lambda_0),$$

which implies

$$\kappa := \sup_{x, v \in \mathbb{R}^d, |v|=1} \int_{\{S_b(x+sv) > -\theta_2\}} ds \leq \sup_{x, v \in \mathbb{R}^d, |v|=1} \int_{\{|x+sv| < \lambda_0\}} ds \leq 4\lambda_0.$$

□

## 4 Order-preserving McKean-Vlasov SDEs

In this part, we consider (2.1) with

$$(4.1) \quad \sigma(x) = \text{diag}\{\sigma_1(x_1), \dots, \sigma_d(x_d)\}, \quad b_t(x, \mu) = (b_1(t, x, \mu), \dots, b_d(t, x, \mu)),$$

where  $\{\sigma_i\}_{1 \leq i \leq d} \subset C(\mathbb{R})$  and

$$(4.2) \quad \begin{aligned} b_i(t, x, \mu) &:= \bar{b}_i(x_i) + \int_{\mathbb{R}^d} Z_i(t, x, y) \mu(dy), \\ \bar{b}_i &\in C(\mathbb{R}), \quad Z_i \in C([0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d; \mathbb{R}), \quad 1 \leq i \leq d. \end{aligned}$$

Then  $X_t = (X_t^1, \dots, X_t^d)$  for  $(X_t^i)_{1 \leq i \leq d}$  solving the SDEs

$$(4.3) \quad dX_t^i = \{\bar{b}_i(X_t^i) + \mathcal{L}_{X_t}(Z_i(t, X_t, \cdot))\} dt + \sigma_i(X_t^i) dW_t^i, \quad 1 \leq i \leq d,$$

where  $\mu(f) := \int_{\mathbb{R}^d} f d\mu$  for a measure  $\mu$  on  $\mathbb{R}^d$  and a measurable function  $f \in L^1(\mu)$ ,  $W_t := (W_t^i)_{1 \leq i \leq d}$  is a  $d$ -dimensional Brownian motion on a complete filtration probability space  $(\Omega, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ .

When  $Z = 0$  (i.e. without interaction), for each  $i$ ,  $X_t^i$  is a one-dimensional diffusion process generated by

$$L_i(r) = \bar{b}_i(r) \frac{d}{dr} + \frac{1}{2} \sigma_i(r)^2 \frac{d^2}{dr^2}.$$

Sharp criteria on the exponential ergodicity have been established for one-dimensional diffusion processes, see for instance [7]. These criteria also apply to the diffusion process generated by  $L(x) := \sum_{i=1}^d L_i(x_i)$  as the components are independent one-dimensional diffusion processes. We will investigate the exponential ergodicity for the solution to (4.3) by making a distribution dependent perturbation to the  $L$ -diffusion process.

To this end, we take the following class of functions as alternatives to the first eigenfunction of  $L_i$ ,  $1 \leq i \leq d$ :

$$\Phi := \left\{ \phi = (\phi_1, \dots, \phi_d) : \phi_i \in C^1(\mathbb{R}), \lim_{|r| \rightarrow \infty} |\phi_i(r)| = \infty, \right.$$

$\phi'_i > 0$  is locally Lipschitz continuous,  $1 \leq i \leq d$  }.

For any  $\phi \in \Phi$ ,  $\mathbb{R}^d$  is a Polish space under the metric

$$d_\phi(x, y) := |\phi(x) - \phi(y)|_1 = \sum_{i=1}^d |\phi_i(x_i) - \phi_i(y_i)|, \quad x, y \in \mathbb{R}^d,$$

so that

$$\mathcal{P}_\phi := \{\mu \in \mathcal{P} : \mu(|\phi|_1) < \infty\}$$

is a Polish space under the Wasserstein distance

$$\mathbf{W}_\phi(\mu, \nu) := \inf_{\pi \in \mathcal{C}(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} d_\phi(x, y) \pi(dx, dy).$$

## 4.1 Main result and example

(A) There exists  $\phi \in \Phi$  such that the following conditions hold:

(A<sub>1</sub>)  $\sigma, V$  and  $Z(t, \cdot)$  (uniformly in  $t$ ) are locally Lipschitz continuous, and there exists a constant  $K > 0$  such that

$$|\phi'_i \sigma_i| \leq K(1 + |\phi_i|), \quad L_i \phi_i^2 \leq K(1 + \phi_i^2), \quad 1 \leq i \leq d.$$

(A<sub>2</sub>)  $\phi_i \in C^2(\mathbb{R})$  and there exists a constant  $q > 0$  such that

$$L_i \phi_i(r) - L_i \phi_i(s) \leq -q|\phi_i(r) - \phi_i(s)|, \quad -\infty < s \leq r < \infty, \quad 1 \leq i \leq d.$$

(A<sub>3</sub>)  $\sup_{t \geq 0} |Z(t, 0, 0)| < \infty$ , for each  $1 \leq i \leq d$ ,  $Z_i(t, x, y)$  is increasing in  $(x_j)_{j \neq i}$  and  $y$ , and there exist constants  $\theta_1, \theta_2 \geq 0$  such that

$$\begin{aligned} & \sum_{i=1}^d |Z_i(t, x, y) \phi'_i(x_i) - Z_i(t, \bar{x}, \bar{y}) \phi'_i(\bar{x}_i)| \\ & \leq \theta_1 d_\phi(x, \bar{x}) + \theta_2 d_\phi(y, \bar{y}), \quad t \geq 0, x \geq \bar{x}, y \geq \bar{y}. \end{aligned}$$

We see that (A) implies the well-posedness of (4.3) in  $\mathcal{P}_\phi$ , and the solution is order-preserving, i.e. for any initial values  $X_0, Y_0$  with distributions in  $\mathcal{P}_\phi$  and  $\mathbb{P}(X_0 \geq Y_0) = 1$ , we have  $\mathbb{P}(X_t \geq Y_t, t \geq 0) = 1$ .

**Theorem 4.1.** *Assume (A). Then (4.3) is well-posed and order-preserving for distributions in  $\mathcal{P}_\phi$ . Moreover,*

$$(4.4) \quad \mathbf{W}_\phi(P_t^* \mu, P_t^* \nu) \leq e^{-(q-\theta_1-\theta_2)t} \mathbf{W}_\phi(\mu, \nu), \quad t \geq 0, \mu, \nu \in \mathcal{P}_\phi.$$

Consequently, if  $q > \theta_1 + \theta_2$ , then  $P_t^*$  has a unique invariant probability measure  $\bar{\mu} \in \mathcal{P}_\phi$  such that

$$(4.5) \quad \mathbf{W}_\phi(P_t^* \mu, \bar{\mu}) \leq e^{-(q-\theta_1-\theta_2)t} \mathbf{W}_\phi(\mu, \bar{\mu}), \quad t \geq 0, \mu \in \mathcal{P}_\phi.$$

To illustrate this result, we consider below a simple example which includes the one-dimensional diffusion process generated by

$$L := \Delta - \nabla H$$

with  $H(x) := -c|x|$  for some constant  $c > 0$  and larger  $|\cdot|$ , since for  $\sigma = \sqrt{2}$  and  $\phi(r) = \text{sgn}(r)e^{\varepsilon|r|}$  for  $|r| \geq 1$  we have

$$\bar{b}(r) := -\frac{q\phi + \frac{\sigma^2}{2}\phi'}{\phi'}(r) = -\frac{q + \varepsilon^2}{\varepsilon}, \quad |r| \geq 1.$$

This is a critical situation for the exponential ergodicity as explained in Introduction.

**Example 4.1.** For each  $1 \leq i \leq d$ , let  $\sigma_i \in C^1(\mathbb{R})$  and  $\phi_i \in C^3(\mathbb{R})$  with  $\phi'_i > 0$  and  $\phi_i(r) = \text{sgn}(r)e^{\varepsilon|r|}$  for some  $\varepsilon > 0$  and  $|r| \geq 1$ . For a constant  $q > 0$  we take

$$(4.6) \quad \bar{b}_i = -\frac{q\phi_i + \frac{\sigma_i^2}{2}\phi''_i}{\phi'_i}, \quad 1 \leq i \leq d.$$

Moreover, for a constant  $\alpha > 0$  and functions  $G_i \in C^1(\mathbb{R}^d \times \mathbb{R}^d)$  increasing in  $(x_j)_{j \neq i}$  and  $y$  with

$$(4.7) \quad \sum_{i=1}^d |G_i(x, y) - G_i(\bar{x}, \bar{y})| \leq d_\phi(x, \bar{x}) + d_\phi(y, \bar{y}), \quad x, y, \bar{x}, \bar{y} \in \mathbb{R}^d,$$

we take

$$(4.8) \quad Z_i(x, y) = \frac{\alpha G_i(x, y)}{\phi'_i(x_i)}, \quad x, y \in \mathbb{R}.$$

Then **(A)** holds for  $\theta_1 = \theta_2 = \alpha$ . Consequently, if  $\alpha < \frac{q}{2}$  then  $P_t^*$  has a unique invariant probability measure  $\bar{\mu} \in \mathcal{P}_\phi$  such that

$$\mathbf{W}_\phi(P_t^* \mu, \bar{\mu}) \leq e^{-(q-2\alpha)t} \mathbf{W}_\phi(\mu, \bar{\mu}), \quad t \geq 0, \mu \in \mathcal{P}_\phi.$$

*Proof.* Obviously, each  $Z_i$  is locally Lipschitz continuous with  $Z_i(x, y)$  increasing in  $(x_j)_{j \neq i}$  and  $y$ , and (4.6) implies

$$L_i \phi_i(r) = \frac{\sigma_i(r)^2}{2} \phi''_i(r) + \bar{b}_i(r) \phi'_i(r) = -q \phi_i(r), \quad r \in \mathbb{R}.$$

Then  $(A_2)$  holds. Next, (4.7) and (4.8) yield

$$\sum_{i=1}^d |Z_i(x, y) \phi'_i(x_i) - Z_i(\bar{x}, \bar{y}) \phi'_i(\bar{x}_i)| \leq \sum_{i=1}^d |G_i(x, y) - G_i(\bar{x}, \bar{y})| \leq \alpha \{d_\phi(x, \bar{x}) + d_\phi(y, \bar{y})\},$$

so that  $(A_3)$  holds for  $\theta_1 = \theta_2 = \alpha$ . Then the desired assertion follows from Theorem 4.1.  $\square$

## 4.2 Proof of Theorem 4.1

(1) We first prove the well-posedness by using the fixed-point theorem in measures as in the proof of Lemma 2.3. Let  $X_0$  be  $\mathcal{F}_0$ -measurable with  $\mathcal{L}_{X_0} \in \mathcal{P}_\phi$ , and let  $T > 0$ . For any  $\mu \in C_w([0, T]; \mathcal{P}_\phi)$ , consider the SDE

$$dX_t^\mu = b_t(X_t^\mu, \mu_t) + \sigma(X_t^\mu) dW_t, \quad t \in [0, T], X_0^\mu = X_0,$$

where  $b$  and  $\sigma$  are given in (4.1) and (4.2). By  $(A_1)$ , the coefficients of this SDE are locally Lipschitz continuous, so the SDE is well-posed up to the life time  $\tau := \lim_{n \rightarrow \infty} \tau_n$ , where

$$\tau_n := \inf\{t \geq 0 : |X_t^\mu| \geq n\}, \quad n \geq 1.$$

By  $(A_3)$  with  $\bar{x} = \bar{y} = 0$  we obtain

$$\sum_{i=1}^d |Z_i(t, x, y) \phi'_i(x)| \leq c_1(1 + |\phi(x)| + |\phi(y)|),$$

which together with  $(A_1)$  yields

$$\sum_{i=1}^d L_i \phi_i^2 + 2 \sum_{i=1}^d (\phi_i \phi'_i)(x_i) \int_{\mathbb{R}^d} Z_i(x, y) \mu(dy) \leq c_2(1 + |\phi(x)|^2 + |\phi(x)| \mu(|\phi|))$$

for some constant  $c_2 > 0$ . Then by Itô's formula, we obtain

$$d|\phi|^2(X_t^\mu) \leq c_2(1 + |\phi(X_t^\mu)|^2 + \mu_t(|\phi|)^2) dt + 2 \sum_{i=1}^d (\phi_i \phi'_i \sigma_i)((X_t^\mu)_i) dW_t^i.$$

So, letting  $\xi_t := \sqrt{1 + |\phi(X_t^\mu)|^2}$ , we derive

$$d\xi_t \leq c_3 \xi_t dt + \frac{1}{\xi_t} \sum_{i=1}^d (\phi_i \phi'_i \sigma_i)((X_t^\mu)_i) dW_t^i$$

for some constant  $c_3 > 0$  depending on  $\mu$ . Thus,

$$\sup_{t \in [0, T]} (P_t^* \mu)(|\phi|) \leq \sup_{t \in [0, T]} \mathbb{E} \xi_t < \infty.$$

This together with the continuity of  $X^\mu$  yields  $H(\mu) := \mathcal{L}_{X^\mu} \in C_w([0, T]; \mathcal{P}_\phi)$ . So, as explained in the proof of Lemma 2.3, it remains to show that  $H$  is contractive under the metric

$$\mathbf{W}_{\phi, \lambda}(\mu, \nu) := \sup_{t \in [0, T]} e^{-\lambda t} \mathbf{W}_\phi(\mu_t, \nu_t), \quad \mu, \nu \in C_w([0, T]; \mathcal{P}_\phi)$$

for large  $\lambda > 0$ .

For  $\mu^1, \mu^2 \in C_w([0, T]; \mathcal{P}_\phi)$ , we choose random variables  $\eta^1, \eta^2$  on  $C([0, T]; \mathbb{R}^d)$  such that  $\mathcal{L}_{\eta^i} = \mu^i, i = 1, 2$ . Let

$$\bar{\mu}_t := \mathcal{L}_{\eta_t^1 \vee \eta_t^2}, \quad \hat{\mu}_t := \mathcal{L}_{\eta_t^1 \wedge \eta_t^2}, \quad t \in [0, T].$$

Then  $\bar{\mu}_t \geq \mu_t^i \geq \hat{\mu}_t$  in the sense

$$\bar{\mu}_t(f) \geq \mu_t^i(f) \geq \hat{\mu}_t(f), \quad t \in [0, T], f \in \mathcal{M}_b(\mathbb{R}^d), i = 1, 2$$

where  $\mathcal{M}_b(\mathbb{R}^d)$  is the class of all bounded increasing functions on  $\mathbb{R}^d$ . Combining this with  $(A_3)$ , we conclude that

$$(4.9) \quad b_i(t, x, \hat{\mu}_t) \leq b_i(t, y, \mu_t^1), b_i(t, y, \mu_t^2) \leq b_i(t, z, \bar{\mu}_t), \quad t \geq 0$$

holds for  $1 \leq i \leq d$  and  $x, y, z \in \mathbb{R}^d$  with  $x_i = y_i = z_i$  and  $x_j \leq y_j \leq z_j$  for  $j \neq i$ . By the order-preservation, this implies

$$(4.10) \quad \mathbb{P}(X_t^{\bar{\mu}} \geq X_t^{\mu^i} \geq X_t^{\hat{\mu}}, t \in [0, T], i = 1, 2) = 1.$$

Indeed, when  $b$  and  $\sigma$  is Lipschitz continuous, by for instance [14, Theorem 1.1] with  $\gamma = \bar{\gamma} = 0$  and  $r_0 = 0$ , (4.10) follows from (4.9) and (4.1). Since  $b$  and  $\sigma$  are locally Lipschitz continuous and  $X_t^\mu$  is non-explosive for any  $\mu \in C_w([0, T]; \mathcal{P}_\phi)$ , we prove (4.10) by a truncation argument. Obviously, (4.10) and  $\phi'_i > 0$  for  $1 \leq i \leq d$  imply

$$(4.11) \quad \sum_{i=1}^d \mathbb{E}[\phi_i(\{X_t^{\bar{\mu}}\}_i) - \phi_i(\{X_t^{\hat{\mu}}\}_i)] \geq \sum_{i=1}^d \mathbb{E}|\phi_i(\{X_t^{\mu^1}\}_i) - \phi_i(\{X_t^{\mu^2}\}_i)|.$$

Moreover, by (4.10) and  $(A_2)$  we see that

$$\xi_t := d_\phi(X_t^{\bar{\mu}}, X_t^{\hat{\mu}}) = \sum_{i=1}^d [\phi_i(\{X_t^{\bar{\mu}}\}_i) - \phi_i(\{X_t^{\hat{\mu}}\}_i)]$$

satisfies  $\xi_t \geq 0$  and

$$d\xi_t \leq (\theta_1 + \theta_2 - q)\xi_t dt + dM_t$$

for some local martingale  $M_t$ . As shown in the proof of Lemma 2.3 that for  $\lambda > 2(\theta_1 + \theta_2 - q)^+$ , this implies the contraction of  $H$  under the metric  $\mathbf{W}_{\phi, \lambda}$ .

(2) Next, since  $Z_i(x, y)$  is increasing in  $(x_j)_{j \neq i}$  and  $y$ , it is easy to see that conditions (1) and (2) in [14, Theorem 1.1] holds for  $b = \bar{b}$ , and its proof applies also with  $\mathbf{W}_\phi$  replacing  $\mathbf{W}_2$  therein, so that the order-preserving property holds. We omit the details to save space. Moreover, since  $\mathcal{P}_\phi$  is complete under  $\mathbf{W}_\phi$ , according to the proof of [23, Theorem 3.1(2)], when  $q > \theta_1 + \theta_2$  the inequality (4.4) implies that  $P_t^*$  has a unique invariant probability measure  $\bar{\mu} \in \mathcal{P}_\phi$  and (4.5) holds. Therefore, below we only prove (4.4).

(3) To prove (4.4), let  $\xi_0, \eta_0$  be  $\mathcal{F}_0$ -measurable random variable with  $\mathcal{L}_{\xi_0} = \mu, \mathcal{L}_{\eta_0} = \nu$  and

$$(4.12) \quad \mathbb{E}d_\phi(\xi_0, \eta_0) = \mathbf{W}_\phi(\mu, \nu).$$

For  $x, y \in \mathbb{R}^d$ , let  $x \vee y = (x_i \vee y_i)_{1 \leq i \leq d}$  and  $x \wedge y = (x_i \wedge y_i)_{1 \leq i \leq d}$ . Take

$$(4.13) \quad X_0 = \xi_0 \vee \eta_0, \quad Y_0 = \xi_0 \wedge \eta_0.$$

Let  $X_t, Y_t, \xi_t, \eta_t$  solve (4.3) with initial values  $X_0, Y_0, \xi_0, \eta_0$  respectively. By the order-preservation, we have

$$(4.14) \quad Y_t \leq \xi_t \wedge \eta_t \leq \xi_t \vee \eta_t \leq X_t, \quad t \geq 0.$$

Consequently,

$$d_\phi(X_t, Y_t) = \sum_{i=1}^d \{\phi_i(X_t^i) - \phi_i(Y_t^i)\}, \quad t \geq 0.$$

By Itô's formula and applying  $(A_1), (A_2)$ , we obtain

$$dd_\phi(X_t, Y_t) \leq \{-qd_\phi(X_t, Y_t) + \theta_1 d_\phi(X_t, Y_t) + \theta_2 \mathbb{E}d_\phi(X_t, Y_t)\}dt + dM_t$$

for a local martingale  $M_t$ . By a standard argument with Gronwall's lemma, this implies

$$\mathbb{E}d_\phi(X_t, Y_t) \leq e^{-(q-\theta_1-\theta_2)t} \mathbb{E}d_\phi(X_0, Y_0) = e^{-(q-\theta_1-\theta_2)t} \mathbb{E}d_\phi(\xi_0, \eta_0), \quad t \geq 0.$$

Combining this with (4.12) and (4.14), we arrive at

$$\mathbf{W}_\phi(P_t^* \mu, P_t^* \nu) \leq \mathbb{E}d_\phi(\xi_t, \eta_t) \leq \mathbb{E}d_\phi(X_t, Y_t) \leq e^{-(q-\theta_1-\theta_2)t} \mathbf{W}_\phi(\mu, \nu), \quad t \geq 0.$$

Then the proof is finished.

**Acknowledgement.** The author would like to thank Professor Jian Wang for helpful comments and corrections.

## References

- [1] V. Barbu, M. Röckner, *From nonlinear Fokker-Planck equations of distribution dependent SDE*, Ann. Probab. 48(2020), 1902–1920.
- [2] O. Butkovsky, *On ergodic properties of nonlinear Markov chains and stochastic McKean-Vlasov equations*, Theory Probab. Appl. 58(2014), 661–674.
- [3] J. A. Carrillo, R. J. McCann, C. Villani, *Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates*, Rev. Mat. Iberoam. 19(2003), 971–1018.
- [4] M.-F. Chen, *From Markov Chains to Non-Equilibrium Particle Systems*, World Scientific, 1992.
- [5] M.-F. Chen, F.-Y. Wang, *Application of coupling method to the first eigenvalue on manifold*, Sci. Sin. (A) 37(1994), 1–14.
- [6] M.-F. Chen, F.-Y. Wang, *General formula for lower bound of the first eigenvalue*, Sci. Sin. (A) 40(1997), 384–394.
- [7] M.-F. Chen, F.-Y. Wang, *Estimation of spectral gap for elliptic operators*, Trans. Amer. Math. Soc. 349(1997), 1239–1267.

- [8] A. Eberle, *Reflection couplings and contraction rates for diffusions*, Probab. Theo. Relat. Fields 166(2016), 851–886.
- [9] A. Eberle, A. Guillin, R. Zimmer, *Quantitative Harris-type theorems for diffusions and McKean-Vlasov processes*, Trans. Amer. Math. Soc. 371 (2019), 7135–7173.
- [10] A. Eberle, R. Zimmer *Sticky couplings of multidimensional diffusions with different drifts*, Ann. Inst. H. Poinc. Probab. Stat. 55(2019), 2370–2394.
- [11] A. Guillin, W. Liu, L. Wu, *Uniform Poincaré and logarithmic Sobolev inequalities for mean field particle systems*, arXiv:1909.07051v1
- [12] M. Hairer, J. C. Mattingly, M. Scheutzow, *Asymptotic coupling and a general form of Harris theorem with applications to stochastic delay equations*, Probab. Theo. Relat. Fields 149(2011), 223–259.
- [13] W. Hammersley, D. Šiška, L. Szpruch, *McKean-Vlasov SDEs under measure dependent Lyapunov conditions*, to appear in Ann. Inst. H. Poinc. Probab. Stat. arXiv:1802.03974v3.
- [14] X. Huang, C. Liu, F.-Y. Wang, *Order Preservation for path-distribution dependent SDEs*, Comm. Pure Appl. Anal. 17(2018), 2125–2133.
- [15] X. Huang, P. Ren, F.-Y. Wang, *Distribution dependent stochastic differential equation*, Front. Math. China 16(2021), 257–301.
- [16] M. Liang, M. B. Majka, J. Wang, *Exponential ergodicity for SDEs and McKean-Vlasov processes with Lévy noise*, to appear in Ann. Inst. H. Poinc. Probab. Stat. arXiv: 1901.11125,
- [17] W. Liu, L. Wu, Ch. Zhang, *Long-time behaviors of mean-field interacting particle systems related to McKean-Vlasov equations*, arXiv:2007.09462.
- [18] F. Malrieu, *Convergence to equilibrium for granular media equations and their Euler schemes*, Ann. Appl. Probab. 13(2003), 540–560.
- [19] E. Priola, F.-Y. Wang, *Gradient estimates for diffusion semigroups with singular coefficients*, J. Funct. Anal. 236(2006), 244–264.
- [20] P. Ren, F.-Y. Wang, *Exponential convergence in entropy and Wasserstein distance for McKean-Vlasov SDEs*, Nonlinear Analysis 206(2021), 112259.
- [21] A. Y. Veretennikov, *On ergodic measures for McKean-Vlasov stochastic equations*, In: Monte Carlo and Quasi-Monte Carlo Methods 2004 (pp. 471–486). Springer, Berlin, Heidelberg, 2006.
- [22] F.-Y. Wang, *Existence of the spectral gap for elliptic operators*, Arkiv för Math. 37(1999), 395–407.
- [23] F.-Y. Wang, *Distribution-dependent SDEs for Landau type equations*, Stoch. Proc. Appl. 128(2018), 595–621.