Electronic Journal of Probability

Electron. J. Probab. **28** (2023), article no. 93, 1–34.
ISSN: 1083-6489 https://doi.org/10.1214/23-EJP981

Existence and non-uniqueness of stationary distributions for distribution dependent SDEs*

Shao-Qin Zhang†

Abstract

The existence and non-uniqueness of stationary distributions for distribution dependent stochastic differential equations with regular coefficients and singular coefficients are investigated. The existence of several stationary distributions is referred to as the phase transition. Our criterion on the existence and the non-uniqueness allow the drift to be in the non-gradient case and the noise to be multiplicative and depend on the law of the solution. By using our criterion, McKean-Vlasov stochastic equations in double-wells landscape with the quadratic interaction and the non-quadratic interaction driven by distribution dependent multiplicative noise are investigated.

Keywords: distribution dependent SDEs; invariant probabilities; phase transition; Zvonkin's transform; McKean-Vlasov stochastic differential equations.
MSC2020 subject classifications: 60H10; 60G10.
Submitted to EJP on August 3, 2021, final version accepted on June 22, 2023.

1 Introduction

When investigating the propagation of chaos for interacting diffusions, McKean [20] introduced a nonlinear stochastic differential equation (SDE) whose coefficients depend on the own law of the solution. This SDE is referred as the McKean(-Vlasov) SDE, established from systems of interacting diffusions by passing to the mean field limit. The associated empirical measure converges in the weak sense to a probability measure with density, and the density satisfies a nonlinear parabolic partial differential equation (PDE) called McKean-Vlasov equation in the literature, see e.g. [22]. Let $\mathscr{P}$ be the space of probability measures on $\mathbb{R}^d$ equipped with the weak topology, $\{W_t\}_{t\geq 0}$ be a d-dimensional Brownian motion on a complete filtration probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$, and let $\mathscr{L}_\eta$ be the law of the random variable η. We consider the SDE on $\mathbb{R}^d$ of the following form

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

*Supported by the National Natural Science Foundation of China (Grant No. 11901604, 11771326), and Program for Innovation Research in Central University of Finance and Economics.

†School of Statistics and Mathematics, Central University of Finance and Economics, Beijing 100081, China. E-mail: zhangsq@cufe.edu.cn

where the coefficients

$$b:\ \mathbb{R}^d\times\mathscr{P}\to\mathbb{R}^d,\qquad \sigma:\ \mathbb{R}^d\times\mathscr{P}\to\mathbb{R}^d\otimes\mathbb{R}^d$$

are measurable. By setting

$$b(x,\mu)=-\nabla V(x)-\nabla F*\mu(x), \tag{1.2}$$

we get the McKean-Vlasov SDE with differentiable confining potential V and interaction potential F, where $*$ stands for the convolution on $\mathbb{R}^d$:

$$f*\mu(x)=\int_{\mathbb{R}^d} f(x-y)\mu(\mathrm{d}y),\ f\in\mathscr{B}(\mathbb{R}^d).$$

Solutions to (1.1) in strong and weak sense are extensively studied by many works, see e.g. [16, 17, 21, 28, 33] and references therein. In this paper, we aim to find stationary distributions to (1.1), i.e. $\mu\in\mathscr{P}$ so that for $\mathscr{L}_{X_0}=\mu$, there is a solution X_t such that $\mathscr{L}_{X_t}\equiv\mu$ for all $t\geq 0$. This solution is called the stationary solution. When b and σ are independent of the law of the solution, the stationary probability measure is the invariant probability measure of the associated classical SDE.

The existence of several stationary distributions to McKean-Vlasov SDEs can occur for the non-convex confining potential. This phenomenon is referred to as phase transition. [6] established for the first time the phase transition for the equation with a particular double-well confinement and Curie-Weiss interaction on the line. Let $d=1$, and let $b(x,\mu)$ satisfy (1.2) with

$$V(x)=\frac{x^4}{4}-\frac{1}{2}x^2,\qquad F(x)=\frac{\alpha}{2}x^2, \tag{1.3}$$

and let $\sigma(x)\equiv\sigma_0$ for some $\sigma_0>0$. Then, Dawson proved in [6] that for fixed $\alpha>0$, there exists $\sigma_c>0$ so that (1.1) has a unique stationary distribution if $\sigma_0>\sigma_c$ and has three stationary distributions if $0<\sigma_0<\sigma_c$. For a classical SDE, if the drift is dissipative at infinity and the noise is non-degenerate, then the associated Markov process is ergodic, see [11] for example. As we see from [6], $-\nabla V$ is dissipative at infinity, but the interaction potential F leads to the non-uniqueness of stationary distributions. Despite [6], there are many works studying the phase transition for McKean-Vlasov equations. [23] provided a criteria for McKean-Vlasov equations with an unphysical assumption that the interaction potential is an odd function. Equations with multi-wells confinement on the whole space were investigated extensively by Tugaut et al. in [10, 14, 15, 24, 25, 26, 27]. In [2, 5], quantitative analyses for continuous and discontinuous phase transition were provided for McKean-Vlasov equations on the torus with $V\equiv 0$, and in [8], the diffusive-mean field limit was investigated for a mean field system with periodic potentials when the associated constrained system on the torus undergoes a phase transition. The phase transition is also studied for nonlinear Markov jump processes, see e.g. [3, 12].

According to the papers mentioned above and references therein, the phase transition is investigated under the assumption that b is of the gradient form like (1.2) and the noise is additive. However, in this paper, we give sufficient conditions for the existence and non-uniqueness of stationary distributions to (1.1), which allow the drift to be in the non-gradient case and the noise to be multiplicative and depend on the law of the solution, see Theorem 2.2, Theorem 2.5 and Theorem 3.1 for details. Moreover, our conditions on the non-uniqueness of stationary distributions can also deal with equations in double-wells landscape considered in [6, 26]. If the drift term is of the gradient form and $\sigma=\sigma_0 I$ for some $\sigma_0\in\mathbb{R}$, the stationary distributions are of an explicit formulation:

$$\mu(\mathrm{d}x)=\frac{\exp\left\{-\frac{2}{\sigma_0^2}\left(V(x)+F*\mu(x)\right)\right\}}{\int_{\mathbb{R}^d}\exp\left\{-\frac{2}{\sigma_0^2}\left(V(x)+F*\mu(x)\right)\right\}\mathrm{d}x}\mathrm{d}x. \tag{1.4}$$

Then the phase transition can be investigated by this explicit formulation and the fixed point theorem for multi-well confinement V and even polynomial interaction F (see e.g. [6, 10, 26]), or combining it with the free energy functional associated with the McKean-Vlasov equation (see [2, 5, 9, 25, 27] for instance):

$$\mathscr{E}^{V,F}(\mu) := \frac{\sigma_0^2}{2} Ent(\mu|\mu_{V_{\sigma_0}}) + \frac{1}{2}\mu(F * \mu),$$

where $\mu_{V_{\sigma_0}}(\mathrm{d}x) = Z_0 \exp\{-2\sigma_0^{-2}V(x)\}\mathrm{d}x$ is a probability measure with the normalizing constant Z_0 and $\mathrm{Ent}(\mu|\mu_{V_{\sigma_0}})$ is the classical relative entropy. However, for (1.1), the explicit formulation as (1.4) for stationary measures is usually not available, and the free energy functional $\mathscr{E}^{V,F}$ may be not available or less explicit even it exists. Hence, we used different methods to establish the existence and non-uniqueness of stationary distributions for (1.1).

Instead of (1.4), we establish a mapping $\mathcal{T}$ on $\mathscr{P}$ whose fixed points are stationary distributions of (1.1). By freezing $\mathscr{L}_{X_t} \equiv \mu \in \mathscr{P}$, we get from (1.1) that

$$\mathrm{d}X_t^\mu = b(X_t^\mu, \mu)\mathrm{d}t + \sigma(X_t^\mu, \mu)\mathrm{d}W_t. \tag{1.5}$$

Denote by $X_t^{\mu,x}$ the solution to (1.5) with $X_0^{\mu,x} = x$, and define $P_t^\mu f(x) = \mathbb{E}f(X_t^{\mu,x})$ the associated linear Markov semigroup to (1.5). Then we take advantage of the ergodic theory for linear Markov semigroups, see e.g. [3, 13, 31], to establish $\mathcal{T}$. Indeed, if P_t^μ has a unique invariant probability measure, then we let the value of $\mathcal{T}$ at μ (denoted by $\mathcal{T}_\mu$) be the invariant probability measure of P_t^μ. The existence of stationary measures to (1.1) is then reformulated as the existence of fixed points of $\mathcal{T}$. Here, we apply the Schauder fixed point theorem to $\mathcal{T}$, instead of the contractive-mapping principle. In fact, the existence of stationary distributions for (1.1) has been investigated by establishing exponential contraction of the transition probability measure in the Wasserstein (quasi-)distance, see e.g. [28, 29]. The exponential contraction can lead to a contractive mapping which implies the existence and uniqueness of stationary distributions. However, the contracting-mapping principle implies the existence and uniqueness at once time. This excludes equations with several invariant measures.

To investigate the non-uniqueness, we establish a general condition on the drift (see Theorem 3.1, Corollary 3.2 and Corollary 3.3 below) which implies that (1.1) has a stationary distribution concentrated around a point $a \in \mathbb{R}^d$. If our condition is satisfied for different $a_1, a_2 \in \mathbb{R}^d$ far enough from each other, then (1.1) has at least two stationary distributions. We use the Schauder fixed point theorem again. By using our condition, we can investigate equations considered in [6, 26] with additive noise replaced by multiplicative noise, see Example 3.5–Example 3.7.

This paper is structured as follows. Section 2 is devoted to the existence of stationary distributions for regular SDEs and singular SDEs. Results on the non-uniqueness and concrete examples are presented in Section 3.

2 Existence of stationary distributions

2.1 Main results

For $\mu \in \mathscr{P}$ and a measurable function f, we denote by $\mu(f)$ the integral $\int_{\mathbb{R}^d} f(x)\mu(\mathrm{d}x)$. Let

$$\mathscr{P}^r = \{\mu \in \mathscr{P} \mid \|\mu\|_r := (\mu(|\cdot|^r))^{\frac{1}{r}} < \infty\},$$
$$\mathscr{P}_M^r = \{\mu \in \mathscr{P} \mid \|\mu\|_r \le M\},\ r > 0, M > 0.$$

Denote by $\|\cdot\|$ the operator norm of a matrix and $\|\cdot\|_{HS}$ the Hilbert-Schmidt norm. For any measurable matrix-value function f on $\mathbb{R}^d \times \mathscr{P}$, we denote

$$\|f\|_\infty = \sup_{x\in\mathbb{R}^d,\mu\in\mathscr{P}} \|f(x,\mu)\|_{HS}.$$

Regular SDEs We first consider (1.1) with the coefficients b and σ are regular. Assume that b and σ satisfy following hypothesis.

(**H1**) There exist constants $r_1 \geq 0, r_2 \geq r_1, r_3 > 0$, $C_1 > 0$, and nonnegative C_2, C_3 such that for any $\mu \in \mathscr{P}^{1+r_2}$

$$\begin{aligned} 2\langle b(x,\mu), x\rangle + (1 + r_2 - r_1)\|\sigma(x,\mu)\|_{HS}^2 \\ \leq -C_1|x|^{1+r_1} + C_2 + C_3\|\mu\|_{1+r_2}^{r_3}. \end{aligned} \tag{2.1}$$

(**H2**) For every $n \in \mathbb{N}$ and $\mu \in \mathscr{P}^{1+r_2}$, there exists $K_n > 0$ such that

$$|b(x,\mu) - b(y,\mu)| + \|\sigma(x,\mu) - \sigma(y,\mu)\|_{HS} \leq K_n|x - y|,\ |x| \vee |y| \leq n. \tag{2.2}$$

There exists a locally bounded function $C_4 :\ [0, +\infty) \to [0, +\infty)$ such that

$$|b(x,\mu)| \leq C_4(\|\mu\|_{1+r_2})(1 + |x|^{r_1}),\ x \in \mathbb{R}^d,\ \mu \in \mathscr{P}^{1+r_2}. \tag{2.3}$$

When $r_1 < 1$, we also assume that for any $\mu \in \mathscr{P}^{1+r_2}$

$$\sup_{x\in\mathbb{R}^d} \frac{\|\sigma(x,\mu)\|_{HS}^2}{1 + |x|^{2r_1}} < +\infty. \tag{2.4}$$

(**H3**) For each $n \geq 1$ and $M > 0$, and $\mu_m, \mu \in \mathscr{P}_M^{1+r_2}$ with $\mu_m \xrightarrow{w} \mu$, there is

$$\lim_{m\to+\infty} \sup_{|x|\leq n} \left(|b(x,\mu) - b(x,\mu_m)| + \|\sigma(x,\mu) - \sigma(x,\mu_m)\|_{HS}\right) = 0. \tag{2.5}$$

Remark 2.1. Fix $\mu \in \mathscr{P}^{1+r_2}$. The condition (H1) and inequality (2.2) yield (1.5) has a unique non-explosive strong solution, see e.g. [19, Theorem 3.1].

For a linear Markov semigroup, it is a general approach that verifying a Lyapunov condition to establish the existence of the invariant probability measures, see e.g. [18]. The condition (H1) comes from the Lyapunov function $|x|^{2+r_2-r_1}$ indeed, see the proof of Lemma 2.8 below.

The following condition was used in [28] to prove the existence of stationary distributions

$$2\langle b(x,\mu) - b(y,\nu), x - y\rangle + \|\sigma(x,\mu) - \sigma(y,\nu)\|_{HS}^2 \leq C_1 W_2(\mu,\nu)^2 - C_2|x - y|^2.$$

When $C_2 > C_1$, the existence and uniqueness and Wasserstein contraction for (1.1) was established. This condition is stronger than (H1) and excludes (1.3).

In (2.6) below or more generally for a symmetric matrix $A \in \mathbb{R}^d \otimes \mathbb{R}^d$ and $\lambda \in \mathbb{R}$, the inequality $A \geq \lambda$ means that

$$\langle Av, v\rangle \geq \lambda|v|^2,\ v \in \mathbb{R}^d.$$

The inequality $A > \lambda$ is defined similarly. Then we have the following theorem.

Theorem 2.2. *Assume (H1)–(H3) and that σ is non-degenerate on $\mathbb{R}^d \times \mathscr{P}^{1+r_2}$:*

$$\sigma(x,\mu)\sigma^*(x,\mu) > 0,\ x \in \mathbb{R}^d, \mu \in \mathscr{P}^{1+r_2}. \tag{2.6}$$

If $r_2 > 0$, $r_3 \leq 1 + r_1$, and $C_1 > C_3$ when $r_3 = 1 + r_1$, then (1.1) *has a stationary distribution.*

Remark 2.3. The non-uniqueness of the steady solution to the aggregation equation with a degenerate and nonlinear second order term was considered in [9]. The aggregation equation in [9] is also associated with a distribution dependent SDE whose the noise is multiplicative and depends on the law of the solution, but it can not be covered by our conditions.

Set $b(x,\mu) = -\nabla V(x) - \nabla F * \mu$ for some twice continuous differential functions V and F. Then we have the following corollary which will be used in discussing the non-uniqueness of stationary distributions.

Corollary 2.4. *Assume that σ is non-degenerate, bounded on $\mathbb{R}^d \times \mathscr{P}$ and satisfies (H2) and (H3), and that V, F are twice continuous differential functions with non-negative constants $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ and $\gamma_0 \in [0,3)$ such that*

$$|\nabla V(x)| \le \alpha_0(1+|x|^3), \tag{2.7}$$

$$|\nabla F(x)| \le \alpha_1 + \alpha_2|x|^3, \tag{2.8}$$

$$\|\nabla^2 F(x)\| \le \alpha_3(1+|x|^{\gamma_0}),\ x \in \mathbb{R}^d. \tag{2.9}$$

Suppose that there exist constants β_0, β_1, and $\beta_2 > 0$, $\beta_3 \ge 0$ such that

$$\nabla^2 V(x) + \nabla^2 F(x-y) \ge \beta_0 - 2\beta_1|x| + 3\beta_2|x|^2 - \beta_3|y|^2,\ x,y \in \mathbb{R}^d, \tag{2.10}$$

and $\alpha_2 + \beta_3 < \beta_2$. Then (1.1) *has a stationary distribution in $\mathscr{P}^4$.*

Proofs of Theorem 2.2 and Corollary 2.4 are given in Subsection 2.2. Concrete examples can be found in Section 3.

Singular SDEs We investigate the existence of stationary distributions for distribution dependent SDEs with singular coefficients by using the Zvonkin transformation introduced in [34]. Well-posedness results for (1.1) have been established by [16] recently. While only stationary distributions are consider in this paper, we use weaker conditions on the coefficients, see (H4) and (H5) below.

We denote by L^p the usual L^p-space on $\mathbb{R}^d$ and by $\|\cdot\|_p$ the L^p-norm. For $(\theta,p) \in [0,2] \times (1,+\infty)$, we define $H^{\theta,p} = (\mathbb{1}-\Delta)^{-\frac{\theta}{2}}(L^p)$ to be the usual Bessel potential space with norm

$$\|f\|_{\theta,p} = \|(\mathbb{1}-\Delta)^{\frac{\theta}{2}} f\|_p,$$

where Δ is the Laplace operator on $\mathbb{R}^d$ and $\mathbb{1}$ is the identity operator. Let $\chi \in C_c^\infty(\mathbb{R}^d)$ with $\mathbb{1}_{[|x|\le 1]} \le \chi \le \mathbb{1}_{[|x|\le 2]}$. We define

$$\chi_r(x) = \chi\left(\frac{x}{r}\right), \qquad \chi_r^z(x) = \chi\left(\frac{x-z}{r}\right),\ r > 0, x, z \in \mathbb{R}^d.$$

Denote by $\tilde{H}^{\theta,p}$ the localized $H^{\theta,p}$-space introduced in [30]:

$$\tilde{H}^{\theta,p} := \left\{ f \in H^{\theta,p}_{loc}(\mathbb{R}^d) \mid \|f\|_{\tilde{H}^{\theta,p}} := \sup_z \|\chi_r^z f\|_{\theta,p} < \infty \right\}.$$

In particular, we denote $\tilde{L}^p = \tilde{H}^{0,p}$. Fixing a probability measure μ, we consider (1.5) of the following form

$$\mathrm{d}X_t^\mu = b_0(X_t^\mu,\mu)\mathrm{d}t + b_1(X_t^\mu,\mu)\mathrm{d}t + \sigma(X_t^\mu,\mu)\mathrm{d}W_t. \tag{2.11}$$

The drift term b_0 is regular and satisfies (H1)–(H3), and b_1 is singular satisfying the following hypothesis, where the constant r_2 is the constant from (H1).

(H4) There exists $p > d$ such that

$$\kappa_0 := \sup_{\mu \in \mathscr{P}^{1+r_2}} \|b_1(\cdot,\mu)\|_{\tilde{L}^p} < \infty.$$

For every $n \geq 1$ and $M \geq 1$,

$$\lim_{\nu \xrightarrow{w} \mu \text{ in } \mathscr{P}_M^{1+r_2}} \|(b_1(\cdot,\mu) - b_1(\cdot,\nu))\mathbb{1}_{[|\cdot|\leq n]}\|_{L^p} = 0. \tag{2.12}$$

We assume that σ satisfies

(H5) For every $\mu \in \mathscr{P}^{1+r_2}$, $\sigma(\cdot,\mu)$ is uniformly continuous on $\mathbb{R}^d$ and $\nabla\sigma(\cdot,\mu) \in \tilde{L}^p$, and there are positive constants λ_1, λ_2 such that

$$\lambda_1 \leq (\sigma\sigma^*)(x,\mu) \leq \lambda_2,\ x \in \mathbb{R}^d, \mu \in \mathscr{P}^{1+r_2}. \tag{2.13}$$

Then we have the following theorem, whose proof is presented in Subsection 2.3.

Theorem 2.5. *Assume that b_0 satisfies (H1)–(H3) (set $\sigma \equiv 0$ there) and satisfies a condition stronger than* (2.3)*: there are positive constants C_5, C_6 such that*

$$|b_0(x,\mu)| \leq C_5(1 + |x|^{r_1}) + C_6\|\mu\|_{1+r_2}^{\frac{r_3 r_1}{1+r_1}},\ x \in \mathbb{R}^d, \mu \in \mathscr{P}^{1+r_2}, \tag{2.14}$$

where r_1, r_2, r_3 are constants from (H1). Assume that b_1 satisfies (H4), and σ satisfies (H3) (set $b \equiv 0$ there) and (H5). If $r_2 > 0$, $r_3 \leq 1 + r_1$, and $C_1 > C_3$ when $r_3 = 1 + r_1$, then (2.11) *has a stationary distribution.*

Throughout the following proofs, notations r_1, r_2, r_3, C_1, C_2, C_3, C_5, C_6 and C_4 are always used to denote the constants and the function in (H1), (H2) and (2.14).

2.2 Proof of Theorem 2.2 and Corollary 2.4

To prove Theorem 2.2, we prove firstly that P_t^μ has a unique invariant probability measure for every $\mu \in \mathscr{P}^{1+r_2}$, see Lemma 2.7 below. We denote by $\mathcal{T}_\mu$ the invariant probability measure of P_t^μ. Then there is a well-defined mapping on $\mathscr{P}^{1+r_2}$:

$$\mathcal{T}:\ \mu \in \mathscr{P}^{1+r_2} \mapsto \mathcal{T}_\mu.$$

Secondly, we prove that $\mathcal{T}$ has an invariant set in $\mathscr{P}^{1+r_2}$, which is nonempty, convex and compact in the topology of weak convergence, see Lemma 2.8. Finally, we apply the Schauder fixed point theorem to $\mathcal{T}$ and the existence of stationary distributions for (1.1) is established. We remark here that the Schauder fixed point theorem is available although $\mathscr{P}^1$ is not a Banach space. The space $\mathscr{P}^1$ equipped with the Kantorovich-Rubinstein-Wasserstein distance (W-distance for short) is a complete metric space (see e.g. [3, Theorem 5.4]):

$$W(\mu,\nu) := \inf_{\pi \in \mathscr{C}(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|\pi(\mathrm{d}x, \mathrm{d}y),\ \mu,\nu \in \mathscr{P}^1,$$

where $\mathscr{C}(\mu,\nu)$ consists of all the couplings of μ and ν. Let $\mathcal{M}^1$ be the set of all finite signed measures on $\mathbb{R}^d$ with $|\mu|(|\cdot|) < \infty$, and let

$$\|\mu\|_{KR} = |\mu(\mathbb{R}^d)| + \sup_{h \in Lip(\mathbb{R}^d), h(0)=0} \int_{\mathbb{R}^d} h(x)\mu(\mathrm{d}x),\ \mu \in \mathcal{M}^1.$$

Then $(\mathcal{M}^1, \|\cdot\|_{KR})$ is a normed space. Moreover,

$$\|\mu - \nu\|_{KR} = W(\mu,\nu),\ \mu,\nu \in \mathscr{P}^1,$$

see [1, Corollary 5.4]. This, together with the fact that $(\mathscr{P}^1, W)$ is a complete metric space, yields that the Schauder fixed point theorem (see e.g. [7, Theorem 8.8]) is available on a nonempty, convex and compact subset of $(\mathscr{P}^1, W)$.

Lemma 2.7 indicates that $\mathcal{T}$ is well-defined. Moreover, we also establish the so called V-uniformly exponential ergodicity (see [13, 31]) for P_t^μ:

Definition 2.6. *Let $V : \mathbb{R}^d \to [1, +\infty)$ be measurable and ν be an invariant probability measure of P_t^μ. The Markov semigroup P_t^μ is V-uniformly exponential ergodic, if there exist $C > 0$ and $\gamma > 0$ such that*

$$\sup_{\|f\|_V \le 1} |P_t^\mu f(x) - \nu(f)| \le CV(x)e^{-\gamma t},\ x \in \mathbb{R}^d,$$

where $\|f\|_V := \sup_{x\in\mathbb{R}^d} |f(x)/V(x)|$.

Lemma 2.7. *Assume that (H1) holds. Fix $\mu \in \mathscr{P}^{1+r_2}$. Let X_t^μ be the solution of (1.5) and P_t^μ the associated Markov semigroup. Then*

(1) *P_t^μ has an invariant probability measure.*
(2) *If (2.2) and (2.3) hold and $\sigma(\cdot,\mu)$ satisfies (2.6), then P_t^μ has a unique invariant probability measure.*
(3) *If (H2) holds and $\sigma(\cdot,\mu)$ satisfies (2.6), then P_t^μ is V-uniformly exponential ergodic with*

$$V(x) = \begin{cases} e^{\delta(1+|X_t|^2)^{\frac{1-r_1}{2}}}, & \text{if } r_1 < 1,\\ 1+|x|, & \text{if } r_1 = 1,\\ 1, & \text{if } r_1 > 1,\end{cases}$$

where the constant $\delta > 0$ depends on μ and $C_1, C_2, C_3, r_1, r_2, C_4$.

Proof. (1) By (H1) and the Itô formula, we have that

$$\begin{aligned} \mathrm{d}|X_t^{\mu,x}|^2 &= 2\langle b(X_t^{\mu,x},\mu), X_t\rangle \mathrm{d}t + 2\langle X_t^{\mu,x}, \sigma(X_t^{\mu,x},\mu)\mathrm{d}W_t\rangle + \|\sigma(X_t^{\mu,x},\mu)\|_{HS}^2 \mathrm{d}t \\ &\le \left(-C_1|X_t^{\mu,x}|^{1+r_1} + C_2 + C_3\|\mu\|_{1+r_2}^{r_3}\right)\mathrm{d}t + 2\langle X_t^{\mu,x}, \sigma(X_t^{\mu,x},\mu)\mathrm{d}W_t\rangle. \end{aligned} \tag{2.15}$$

Writing (2.15) in integral form and taking expectation, we find that there are positive constants $\tilde{C}_1, \tilde{C}_2$ such that for all $s \ge 0$

$$\mathbb{E}\int_0^s |X_t^{\mu,x}|^{1+r_1}\mathrm{d}t + \mathbb{E}|X_s^{\mu,x}|^2 \le \tilde{C}_1|x|^2 + \tilde{C}_2\left(1 + \|\mu\|_{1+r_2}^{r_3}\right)s.$$

Thus

$$\begin{aligned} \sup_{s\ge 1}\left(\frac{1}{s}\int_0^s P_t^{\mu,*}\delta_x\left(|\cdot|^{1+r_1}\right)\mathrm{d}t\right) &= \sup_{s\ge 1}\left(\frac{1}{s}\int_0^s \mathbb{E}|X_t^{\mu,x}|^{1+r_1}\mathrm{d}t\right) \\ &\le C_1|x|^2 + \tilde{C}_2\left(1+\|\mu\|_{1+r_2}^{r_3}\right) \\ &< +\infty. \end{aligned} \tag{2.16}$$

Since $|\cdot|^{1+r_1}$ has is a compact function, i.e. the level set $\{x \mid |x|^{1+r_1} \le c\}$ is compact for every $c > 0$, (2.16) implies the tightness of $\{s^{-1}\int_0^s P_t^{\mu,*}\delta_x \mathrm{d}t\}_{s\ge 1}$. This concludes the proof of the first statement.

(2) Since (2.2), $\sigma(\cdot,\mu)$ is continuous. Let $\chi \in C_c^2(\mathbb{R}^d)$ with $\mathbb{1}_{[|x|\le 1]} \le \chi(x) \le \mathbb{1}_{[|x|\le 2]}$. Then for each $m \ge 1$,

$$\sigma_m(x) := \sigma(\chi(x/m)x, \mu),\ x \in \mathbb{R}^d$$

is Lipschitz, bounded and non-degenerate on $\mathbb{R}^d$. Due to (2.2) again, $b_m(\cdot) := b(\cdot,\mu)\chi(\cdot/m)$ is bounded and belongs to $L^p(\mathbb{R}^d)$ for any $p > 0$. Combining these with (2.1) and (2.3), one has that assumptions of [31, Lemma 7.3] hold. Thus, P_t^μ has the strong Feller property and irreducibility due to [31, Lemma 7.3]. Hence, P_t^μ has a unique invariant probability measure.

(3) One can see that (H1), (2.2) and (2.3) also implies that assumptions of [31, Theorem 7.4] hold. Thus, for $r_1 \geq 1$, P_t^μ is V-uniformly exponential ergodic with V defined as above. For $r_1 \in [0,1)$, it follows from (H1) and (2.4) that there exist positive constants $\bar{C}_1, \bar{C}_2, \bar{C}_3$ depending on $\mu(|\cdot|^{1+r_2})$ such that

$$\begin{aligned} 2\langle x, b(x,\mu)\rangle + \|\sigma(x,\mu)\|_{HS}^2 &\leq -\bar{C}_1(1+|x|^2)^{\frac{1+r_1}{2}} + \bar{C}_2, \\ \|\sigma(x,\mu)\|_{HS}^2 &\leq \bar{C}_3(1+|x|^2)^{r_1},\ x \in \mathbb{R}^d. \end{aligned}$$

Then we have by the Itô formula that

$$\begin{aligned} \mathrm{d}\left(1+|X_t^{\mu,x}|^2\right)^{\frac{1-r_1}{2}} \leq &-\left(\frac{(1-r_1)\bar{C}_1}{2} - \frac{\bar{C}_2(1-r_1)}{2(1+|X_t^{\mu,x}|^2)^{\frac{1+r_1}{2}}} + \frac{(1-r_1^2)|\sigma^*(X_t^{\mu,x})X_t|^2}{2(1+|X_t^{\mu,x}|^2)^{\frac{3+r_1}{2}}}\right)\mathrm{d}t \\ &+ \frac{(1-r_1)\langle X_t^{\mu,x}, \sigma(X_t^{\mu,x})\mathrm{d}W_t\rangle}{(1+|X_t^{\mu,x}|^2)^{\frac{1+r_1}{2}}},\ x \in \mathbb{R}^d. \end{aligned}$$

Let

$$\delta = \frac{\bar{C}_1}{2\bar{C}_3(1+r_1)},\ \hat{C}_2 = \frac{\delta \bar{C}_2(1-r_1)}{2} \sup_{1 \leq u \leq (\frac{4\bar{C}_2}{\bar{C}_1})^{\frac{1}{1+r_1}}} u^{-(1+r_1)} e^{\delta u^{1-r_1}},$$

and let $V(x) = e^{\delta(1+|x|^2)^{\frac{1-r_1}{2}}}$. Then

$$\begin{aligned} &\mathrm{d}V(X_t^{\mu,x}) - \frac{\delta(1-r_1)V(X_t^{\mu,x})\langle X_t^\mu, \sigma(X_t^{\mu,x})\mathrm{d}W_t\rangle}{(1+|X_t^{\mu,x}|^2)^{\frac{1+r_1}{2}}} \\ &\leq -\delta V(X_t^\mu)\left(\frac{\bar{C}_1(1-r_1)}{2} - \frac{\bar{C}_2(1-r_1)}{2(1+|X_t^{\mu,x}|^2)^{\frac{1+r_1}{2}}} - \frac{\delta(1-r_1)^2|\sigma^*(X_t^\mu)X_t^{\mu,x}|^2}{2(1+|X_t^{\mu,x}|^2)^{1+r_1}}\right)\mathrm{d}t \\ &\leq -\delta V(X_t^\mu)\left(\frac{\bar{C}_1(1-r_1)}{2} - \frac{\delta(1-r_1)^2\bar{C}_3}{2} - \frac{\bar{C}_2(1-r_1)}{2(1+|X_t^{\mu,x}|^2)^{\frac{1+r_1}{2}}}\right)\mathrm{d}t \\ &\leq \delta V(X_t^{\mu,x})\left(-\frac{\bar{C}_1(1-r_1)}{4} + \frac{\bar{C}_2(1-r_1)}{2(1+|X_t^{\mu,x}|^2)^{\frac{1+r_1}{2}}}\mathbb{1}_{[\frac{4\bar{C}_2}{\bar{C}_1} \leq (1+|X_t^{\mu,x}|^2)^{\frac{1+r_1}{2}}]}\right)\mathrm{d}t \\ &\quad + \sup_{\frac{4\bar{C}_2}{\bar{C}_1} \geq (1+|x|^2)^{\frac{1+r_1}{2}}}\left(\frac{\bar{C}_2(1-r_1)V(x)}{2(1+|x|^2)^{\frac{1+r_1}{2}}}\right)\mathrm{d}t \\ &\leq -\frac{\bar{C}_1(1-r_1)\delta}{8}V(X_t^{\mu,x})\mathrm{d}t + \hat{C}_2\mathrm{d}t. \end{aligned}$$

Hence,

$$\mathbb{E}e^{\delta(1+|X_t^{\mu,x}|^2)^{\frac{1-r_1}{2}}} \leq e^{-\frac{\bar{C}_1(1-r_1)\delta t}{8}} e^{\delta(1+|x|^2)^{\frac{1-r_1}{2}}} + \frac{8\hat{C}_2}{\bar{C}_1(1-r_1)\delta},\ t \geq 0,\ x \in \mathbb{R}^d.$$

Combining this with the strong Feller property and irreducibility of P_t^μ, we can prove that P_t^μ is V-uniformly exponential ergodic by following line by line of the proof of [13, Theorem 2.5], and we omit the details here. □

Next, we prove that $\mathscr{P}_M^{1+r_2}$ is an invariant subset of $\mathcal{T}$ for large M when $\mathcal{T}$ is well-defined on $\mathscr{P}^{1+r_2}$.

Lemma 2.8. *Assume that (H1) holds and for each $\mu \in \mathscr{P}^{1+r_2}$, P_t^μ has a unique invariant probability measure $\mathcal{T}_\mu$. Then there exists $M_0 > 0$ such that for all $M \ge M_0$, $\mathcal{T}$ maps $\mathscr{P}_M^{1+r_2}$ into $\mathscr{P}_M^{1+r_2}$.*

Proof. Let $q = 1 + \frac{r_2 - r_1}{2}$. Then

$$q \ge 1, \qquad q - 1 = \frac{r_2 - r_1}{2}, \qquad 2q + r_1 - 1 = 1 + r_2. \tag{2.17}$$

It follows from the Itô formula, (2.1) and $2q - 1 = 1 + r_2 - r_1$ that

$$\begin{aligned} \mathrm{d}|X_t^{\mu,x}|^{2q} &\le q|X_t^{\mu,x}|^{2(q-1)} \left(2\langle b(X_t^{\mu,x},\mu), X_t^{\mu,x}\rangle + (2q-1)\|\sigma(X_t^{\mu,x},\mu)\|_{HS}^2\right)\mathrm{d}t \\ &\quad + 2q|X_t^{\mu,x}|^{2(q-1)}\langle X_t^{\mu,x}, \sigma(X_t^{\mu,x},\mu)\mathrm{d}W_t\rangle \\ &\le q|X_t^{\mu,x}|^{r_2-r_1}\left(-C_1|X_t^{\mu,x}|^{1+r_1} + C_2 + C_3\|\mu\|_{1+r_2}^{r_3}\right)\mathrm{d}t \\ &\quad + 2q|X_t^{\mu,x}|^{r_2-r_1}\langle X_t^{\mu,x}, \sigma(X_t^{\mu,x},\mu)\mathrm{d}W_t\rangle \\ &\le -qC_1|X_t^{\mu,x}|^{1+r_2} + q\left(C_2 + C_3\|\mu\|_{1+r_2}^{r_3}\right)|X_t^{\mu,x}|^{r_2-r_1}\mathrm{d}t \\ &\quad + 2q|X_t^{\mu,x}|^{r_2-r_1}\langle X_t^{\mu,x}, \sigma(X_t^{\mu,x},\mu)\mathrm{d}W_t\rangle. \end{aligned} \tag{2.18}$$

Since $0 \le r_2 - r_1 < 1 + r_2$, it follows from the Hölder inequality that

$$|X_t^{\mu,x}|^{r_2-r_1} \le \frac{1+r_1}{1+r_2} + \frac{r_2-r_1}{1+r_2}|X_t^{\mu,x}|^{1+r_2}.$$

Then writing (2.18) in integral form, putting this into (2.18), and taking the expectation, we find that there are positive constants $\tilde{C}_1, \tilde{C}_2$ such that

$$\mathbb{E}\left(|X_s^{\mu,x}|^{2+r_2-r_1} + \int_0^s |X_t^{\mu,x}|^{1+r_2}\mathrm{d}t\right) \le \tilde{C}_1|x|^{2+r_2-r_1} + \tilde{C}_2 s,\ s \ge 0. \tag{2.19}$$

Since that P_t^μ has a unique invariant probability measure, it follows from (2.16) that there exists a sequence $t_n \uparrow +\infty$ such that as $n \to +\infty$,

$$\nu_n := \frac{1}{t_n}\int_0^{t_n} P_s^{\mu,*}\delta_0\mathrm{d}s \xrightarrow{w} \mathcal{T}_\mu. \tag{2.20}$$

Thus

$$\lim_{n\to+\infty} \nu_n(|\cdot|^{1+r_2} \wedge N) = \mathcal{T}_\mu(|\cdot|^{1+r_2} \wedge N),\ N \ge 1,$$

which, together with (2.19), yields that

$$\mathcal{T}_\mu(|\cdot|^{1+r_2} \wedge N) \le \tilde{C}_2,\ N \ge 1.$$

By Fatou's lemma, $\mathcal{T}_\mu(|\cdot|^{1+r_2}) < \infty$. It follows from (2.19) that $\sup_{n\ge1}\nu_n(|\cdot|^{1+r_2}) < +\infty$. Combining this with the fact that $r_2 - r_1 < 1 + r_2$ and (2.20), we have that

$$\lim_{n\to+\infty} \nu_n(|\cdot|^{r_2-r_1}) = \mathcal{T}_\mu(|\cdot|^{r_2-r_1}).$$

Writing (2.18) in integral form, setting $x = 0$ and taking expectations, we find that

$$C_1\nu_n(|\cdot|^{1+r_2} \wedge N) \le C_1\nu_n(|\cdot|^{1+r_2}) \le C_2\nu_n(|\cdot|^{r_2-r_1}) + C_3\|\mu\|_{1+r_2}^{r_3}\nu_n(|\cdot|^{r_2-r_1}).$$

Taking $n \to +\infty$ and $N \to +\infty$, we arrive at

$$\mathcal{T}_\mu(|\cdot|^{1+r_2}) \le \frac{C_2}{C_1}\mathcal{T}_\mu(|\cdot|^{r_2-r_1}) + \frac{C_3}{C_1}\|\mu\|_{1+r_2}^{r_3}\mathcal{T}_\mu(|\cdot|^{r_2-r_1}).$$

Then the Jensen inequality yields that

$$\mathcal{T}_\mu(|\cdot|^{1+r_2}) \le \frac{C_2}{C_1}\|\mathcal{T}_\mu\|_{1+r_2}^{r_2-r_1} + \frac{C_3}{C_1}\|\mu\|_{1+r_2}^{r_3}\|\mathcal{T}_\mu\|_{1+r_2}^{r_2-r_1}. \tag{2.21}$$

Since $1+r_1>0$, we derive from (2.21) that (even $\mathcal{T}_\mu(|\cdot|^{1+r_2})=0$)

$$\|\mathcal{T}_\mu\|_{1+r_2}^{1+r_1} \le \frac{C_2}{C_1} + \frac{C_3}{C_1}\|\mu\|_{1+r_2}^{r_3}, \tag{2.22}$$

If $r_3 = 1+r_1$, then for

$$M \ge \left(\frac{C_2}{C_1-C_3}\right)^{\frac{1}{r_3}} =: M_0,$$

we have that

$$\frac{C_2}{C_1} + \frac{C_3}{C_1}M^{r_3} \le M^{1+r_1}. \tag{2.23}$$

Consequently, $\|\mathcal{T}_\mu\|_{1+r_2} \le M$ for all $\|\mu\|_{1+r_2}\le M$.

If $r_3 < 1+r_1$, then it is easy to see that there exists M_0 depending on C_i ($i=1,2,3$) and r_1, r_3 such that (2.23) holds for each $M \ge M_0$. Hence, for every $\mu \in \mathscr{P}_M^{1+r_2}$, we derive from (2.22) that

$$\|\mathcal{T}_\mu\|_{1+r_2}^{1+r_1} \le \frac{C_2}{C_1} + \frac{C_3}{C_1}M^{r_3} \le M^{1+r_1}.$$

Therefore, we prove that $\mathcal{T}_\mu \in \mathscr{P}_M^{1+r_2}$ for every $M \ge M_0$. □

Combining Lemma 2.7 with Lemma 2.8, we arrive at that under the assumption of Theorem 2.2, $\mathcal{T}$ is well-defined and maps $\mathscr{P}_M^{1+r_2}$ into itself. It is clear that $\mathscr{P}_M^{1+r_2}$ is a convex and weak compact set in $\mathscr{P}^1$ if $r_2>0$. Next, we prove that $\mathcal{T}$ is weak continuous in $\mathscr{P}_M^{1+r_2}$.

Lemma 2.9. *Under the assumption of Theorem 2.2, $\mathcal{T}$ is weak continuous in $\mathscr{P}_M^{1+r_2}$.*

Proof. For all $f \in \mathscr{B}_b(\mathbb{R}^d)$ and $t>0$, we have that

$$\begin{aligned}|\mathcal{T}_\mu(f) - \mathcal{T}_\nu(f)| &= |\mathcal{T}_\mu(f) - \mathcal{T}_\nu(P_t^\nu f)| \\ &\le |\mathcal{T}_\nu(\mathcal{T}_\mu(f)) - \mathcal{T}_\nu(P_t^\mu f)| + |\mathcal{T}_\nu(P_t^\mu f) - \mathcal{T}_\nu(P_t^\nu f)| \\ &\le \mathcal{T}_\nu\left(|\mathcal{T}_\mu(f) - P_t^\mu f|\right) + \mathcal{T}_\nu\left(|P_t^\mu f - P_t^\nu f|\right).\end{aligned} \tag{2.24}$$

By (3) of Lemma 2.7, there are positive constants γ and C depending on μ such that

$$|\mathcal{T}_\mu(f) - P_t^\mu f(x)| \le CV(x)e^{-\gamma t},\ x\in\mathbb{R}^d.$$

Then for any $\mu,\nu \in \mathscr{P}_M^{1+r_2}$ and each $m>0$, we have by Lemma 2.7 that

$$\begin{aligned}\mathcal{T}_\nu\left(|\mathcal{T}_\mu(f) - P_t^\mu f|\right) &\le \mathcal{T}_\nu\left(CV(\cdot)e^{-\gamma t}\mathbb{1}_{[|\cdot|\le m]} + \|f\|_\infty \mathbb{1}_{[|\cdot|>m]}\right) \\ &\le C\left(\sup_{|x|\le m} V(x)\right)e^{-\gamma t} + \|f\|_\infty \mathcal{T}_\nu(|\cdot|>m) \\ &\le C\left(\sup_{|x|\le m} V(x)\right)e^{-\gamma t} + \frac{\|f\|_\infty}{m}\|\mathcal{T}_\nu\|_{1+r_2} \\ &\le C\left(\sup_{|x|\le m} V(x)\right)e^{-\gamma t} + \frac{\|f\|_\infty M}{m}.\end{aligned}$$

Consequently,

$$\overline{\lim_{t\to+\infty}} \sup_{\nu\in\mathscr{P}_M^{1+r_2}} \mathcal{T}_\nu\left(|\mathcal{T}_\mu(f)-P_t^\mu f|\right)=0.$$

Hence, for each $\epsilon>0$, we choose $t_\epsilon>0$ such that

$$\sup_{\nu\in\mathscr{P}_M^{1+r_2}} \mathcal{T}_\nu\left(\left|\mathcal{T}_\mu(f)-P_{t_\epsilon}^\mu f\right|\right)<\epsilon. \tag{2.25}$$

We remark here that t_ϵ depends on $\mu, M, \|f\|_\infty$ and is independent of $\nu\in\mathscr{P}_M^{1+r_2}$ and $x\in\mathbb{R}^d$.

Let

$$\tau_n(x)=\inf\{\, t>0 \mid |X_t^{\mu,x}|\vee|X_t^{\nu,x}|\geq n\}.$$

For every $f\in C_b(\mathbb{R}^d)\cap Lip(\mathbb{R}^d)$, we have that

$$\begin{aligned}
\mathcal{T}_\nu\left(\left|P_{t_\epsilon}^\mu f-P_{t_\epsilon}^\nu f\right|\right) &\leq \int_{\mathbb{R}^d}\mathbb{E}\left|f(X_{t_\epsilon}^{\mu,x})-f(X_{t_\epsilon}^{\nu,x})\right|\mathcal{T}_\nu(\mathrm{d}x)\\
&\leq \int_{\mathbb{R}^d}\mathbb{E}\left(\left|f(X_{t_\epsilon}^{\mu,x})-f(X_{t_\epsilon}^{\nu,x})\right|\mathbb{1}_{[t_\epsilon<\tau_n(x)]}\right)\mathcal{T}_\nu(\mathrm{d}x)\\
&\quad+\int_{\mathbb{R}^d}\mathbb{E}\left(\left|f(X_{t_\epsilon}^{\mu,x})-f(X_{t_\epsilon}^{\nu,x})\right|\mathbb{1}_{[t_\epsilon\geq\tau_n(x)]}\right)\mathcal{T}_\nu(\mathrm{d}x)\\
&\leq \|\nabla f\|_\infty\int_{|x|\leq\frac{n}{2}}\mathbb{E}\left(\left|X_{t_\epsilon}^{\mu,x}-X_{t_\epsilon}^{\nu,x}\right|\mathbb{1}_{[t_\epsilon<\tau_n(x)]}\right)\mathcal{T}_\nu(\mathrm{d}x)\\
&\quad+2\|f\|_\infty\mathcal{T}_\nu(|\cdot|\geq\frac{n}{2})+2\|f\|_\infty\mathcal{T}_\nu\left(\mathbb{P}(t_\epsilon\geq\tau_n(\cdot))\right)\\
&\leq \|\nabla f\|_\infty\int_{|x|\leq\frac{n}{2}}\mathbb{E}\left(\left|X_{t_\epsilon\wedge\tau_n(x)}^{\mu,x}-X_{t_\epsilon\wedge\tau_n(x)}^{\nu,x}\right|\right)\mathcal{T}_\nu(\mathrm{d}x)\\
&\quad+\frac{4}{n}\|f\|_\infty\|\mathcal{T}_\nu\|_{1+r_2}\\
&\quad+2\|f\|_\infty\int_{\mathbb{R}^d}\mathbb{P}\left(\sup_{s\in[0,t_\epsilon]}|X_s^{\mu,x}|\vee|X_s^{\nu,x}|\geq n\right)\mathcal{T}_\nu(\mathrm{d}x)\\
&=: I_{1,n}+I_{2,n}+I_{3,n}.
\end{aligned}\tag{2.26}$$

For $I_{2,n}$. Due to Lemma 2.7 and $\nu\in\mathscr{P}_M^{1+r_2}$, we have that $\|\mathcal{T}_\nu\|_{1+r_2}\leq M$. Then

$$\lim_{n\to+\infty}\sup_{\nu\in\mathscr{P}_M^{1+r_2}} I_{2,n}=0. \tag{2.27}$$

For $I_{3,n}$. For every $\eta\in\mathscr{P}^{1+r_2}$ and $s\geq 0$, we derive from the Itô formula and (H1) that

$$\begin{aligned}
|X_s^{\eta,x}|^2+C_1\int_0^s|X_t^{\eta,x}|^{1+r_1}\mathrm{d}t &\leq |x|^2+\left(C_2+C_3\|\mu\|_{1+r_2}^{r_3}\right)s\\
&\quad+2\int_0^s\langle X_t^{\eta,x},\sigma(X_t^{\eta,x},\eta)\mathrm{d}W_t\rangle.
\end{aligned}\tag{2.28}$$

It follows from (H1), (2.3) of (H2) and the Hölder inequality that

$$\begin{aligned}
\|\sigma(x,\eta)\|_{HS}^2 &\leq C_4(\|\eta\|_{1+r_2})(1+|x|^{r_1})|x|-C_1|x|^{1+r_1}+C_2+C_3\|\eta\|_{1+r_2}\\
&\leq \left(C_4(\|\eta\|_{1+r_2})+(1+r_1)^{-1}-C_1\right)|x|^{1+r_1}\\
&\quad+\frac{r_1}{1+r_1}C_4(\|\eta\|_{1+r_2})+C_2+C_3\|\eta\|_{1+r_2}\\
&\equiv \bar{C}_5(\|\eta\|_{1+r_2})|x|^{1+r_1}+\bar{C}_6(\|\eta\|_{1+r_2}),\ \eta\in\mathscr{P}^{1+r_2}.
\end{aligned}$$

It is clear $\bar{C}_5, \bar{C}_6$ are locally bounded functions. Then, by the B-D-G inequality and the Hölder inequality, we derive from (2.28) that for every $T > 0$

$$\begin{aligned}\mathbb{E}\sup_{s\in[0,T]}|X_s^{\eta,x}|^2 &\leq |x|^2 + \left(C_2 + C_3\|\eta\|_{1+r_2}^{r_3}\right)T + \mathbb{E}\left(\int_0^T |X_t^{\eta,x}|^2\|\sigma^*(X_t^{\eta,x},\eta)\|^2\mathrm{d}t\right)^{\frac{1}{2}}\\ &\leq |x|^2 + \left(C_2 + C_3\|\eta\|_{1+r_2}^{r_3}\right)T + \frac{1}{2}\mathbb{E}\sup_{s\in[0,T]}|X_s^{\eta,x}|^2\\ &\quad + \frac{1}{2}\mathbb{E}\int_0^T \left(\bar{C}_5(\|\eta\|_{1+r_2})|X_t^{\eta,x}|^{1+r_1} + \bar{C}_6(\|\eta\|_{1+r_2})\right)\mathrm{d}t.\end{aligned}$$

This, together with the expectation of (2.28), implies that

$$\begin{aligned}\mathbb{E}\sup_{s\in[0,T]}|X_s^{\eta,x}|^2 &\leq 2\left(|x|^2 + \left(C_2 + C_3\|\eta\|_{1+r_2}^{r_3}\right)T\right)\\ &\quad + \bar{C}_6(\|\eta\|_{1+r_2})T + \bar{C}_5(\|\eta\|_{1+r_2})\mathbb{E}\int_0^T |X_t^{\eta,x}|^{1+r_1}\mathrm{d}t\\ &\leq \left(2 + \frac{\bar{C}_5(\|\eta\|_{1+r_2})}{C_1}\right)\left(|x|^2 + \left(C_2 + C_3\|\eta\|_{1+r_2}^{r_3}\right)T\right) + \bar{C}_6(\|\eta\|_{1+r_2})T.\end{aligned}$$

Combining this with the Chebyshev inequality and that $\nu, \mu \in \mathscr{P}_M^{1+r_2}$, there exist $C > 0$ depending on C_i, $i = 1, 2, 3$, r_1, r_2, r_3, M and t_ϵ such that

$$\begin{aligned}&\mathbb{P}\left(\sup_{s\in[0,t_\epsilon]}|X_s^{\mu,x}| \vee |X_s^{\nu,x}| \geq n\right)\\ &\quad\leq \mathbb{P}\left(\sup_{s\in[0,t_\epsilon]}|X_s^{\mu,x}| \geq n\right) + \mathbb{P}\left(\sup_{s\in[0,t_\epsilon]}|X_s^{\nu,x}| \geq n\right)\\ &\quad\leq n^{-(2\wedge(1+r_2))}\mathbb{E}\left(\sup_{s\in[0,t_\epsilon]}|X_s^{\mu,x}|^{2\wedge(1+r_2)} + \sup_{s\in[0,t_\epsilon]}|X_s^{\nu,x}|^{2\wedge(1+r_2)}\right)\\ &\quad\leq C(1 + |x|^{2\wedge(1+r_2)})n^{-(2\wedge(1+r_2))}.\end{aligned} \tag{2.29}$$

Recalling that t_ϵ is independent of $\nu \in \mathscr{P}_M^{1+r_2}$ and $x \in \mathbb{R}^d$, we derive from (2.29) and $\mathcal{T}_\nu \in \mathscr{P}_M^{1+r_2}$ that

$$\begin{aligned}\varlimsup_{n\to+\infty}\sup_{\nu\in\mathscr{P}_M^{1+r_2}} I_{3,n} &= \varlimsup_{n\to+\infty}\sup_{\nu\in\mathscr{P}_M^{1+r_2}}\int_{\mathbb{R}^d}\mathbb{P}\left(\sup_{s\in[0,t_\epsilon]}|X_s^{\mu,x}| \vee |X_s^{\nu,x}| \geq n\right)\mathcal{T}_\nu(\mathrm{d}x)\\ &\leq \left(\lim_{n\to+\infty}\frac{C}{n^{2\wedge(1+r_2)}}\right)\sup_{\nu\in\mathscr{P}_M^{1+r_2}}\mathcal{T}_\nu(1 + |\cdot|^{2\wedge(1+r_2)})\\ &= 0.\end{aligned} \tag{2.30}$$

For $I_{1,n}$. It is easy to see that for $s < \tau_n(x)$

$$\begin{aligned}|X_s^{\mu,x} - X_s^{\nu,x}| &\leq \int_0^s |b(X_t^{\mu,x},\mu) - b(X_t^{\nu,x},\nu)|\,\mathrm{d}t + \left|\int_0^s (\sigma(X_t^{\mu,x},\mu) - \sigma(X_t^{\nu,x},\nu))\,\mathrm{d}W_t\right|\\ &\leq \int_0^s |b(X_t^{\mu,x},\mu) - b(X_t^{\nu,x},\mu)|\,\mathrm{d}t + \int_0^s |b(X_t^{\nu,x},\mu) - b(X_t^{\nu,x},\nu)|\,\mathrm{d}t\\ &\quad + \left|\int_0^s (\sigma(X_t^{\mu,x},\mu) - \sigma(X_t^{\nu,x},\mu))\,\mathrm{d}W_t\right|\\ &\quad + \left|\int_0^s (\sigma(X_t^{\nu,x},\mu) - \sigma(X_t^{\nu,x},\nu))\,\mathrm{d}W_t\right|.\end{aligned} \tag{2.31}$$

It follows from (H2) that

$$\int_0^{s\wedge\tau_n(x)}|b(X_t^{\mu,x},\mu)-b(X_t^{\nu,x},\mu)|\,\mathrm{d}t\le K_n\int_0^{s\wedge\tau_n(x)}|X_t^{\mu,x}-X_t^{\nu,x}|\,\mathrm{d}t$$

and

$$\begin{aligned}
&\mathbb{E}\sup_{r\in[0,s]}\left|\int_0^{r\wedge\tau_n(x)}(\sigma(X_t^{\mu,x},\mu)-\sigma(X_t^{\nu,x},\mu))\,\mathrm{d}W_t\right|\\
&\quad\le\mathbb{E}\left(\int_0^{s\wedge\tau_n(x)}\|\sigma(X_t^{\mu,x},\mu)-\sigma(X_t^{\nu,x},\mu)\|_{HS}^2\,\mathrm{d}t\right)^{\frac12}\\
&\quad\le K_n\mathbb{E}\left(\int_0^{s\wedge\tau_n(x)}|X_t^{\mu,x}-X_t^{\nu,x}|^2\,\mathrm{d}t\right)^{\frac12}\\
&\quad\le K_n\mathbb{E}\left(\sup_{t\in[0,s\wedge\tau_n(x)]}|X_t^{\mu,x}-X_t^{\nu,x}|^{\frac12}\right)\left(\int_0^{s\wedge\tau_n(x)}|X_t^{\mu,x}-X_t^{\nu,x}|\,\mathrm{d}t\right)^{\frac12}\\
&\quad\le\frac12\mathbb{E}\sup_{t\in[0,s]}\left|X_{t\wedge\tau_n(x)}^{\mu,x}-X_{t\wedge\tau_n(x)}^{\nu,x}\right|+\frac12K_n^2\mathbb{E}\int_0^{s\wedge\tau_n(x)}|X_t^{\mu,x}-X_t^{\nu,x}|\,\mathrm{d}t.
\end{aligned}$$

We remark that K_n here is independent of ν. Putting these into (2.31), we derive by the Gronwall inequality that for every $s>0$

$$\begin{aligned}
&\mathbb{E}\sup_{t\in[0,s]}\left|X_{t\wedge\tau_n(x)}^{\mu,x}-X_{t\wedge\tau_n(x)}^{\nu,x}\right|\\
&\quad\le 2e^{(2K_n+K_n^2)s}\Big(\mathbb{E}\sup_{r\in[0,s]}\int_0^{r\wedge\tau_n(x)}|b(X_t^{\nu,x},\mu)-b(X_t^{\nu,x},\nu)|\,\mathrm{d}t\\
&\qquad+\mathbb{E}\sup_{r\in[0,s]}\left|\int_0^{r\wedge\tau_n(x)}(\sigma(X_t^{\nu,x},\mu)-\sigma(X_t^{\nu,x},\nu))\,\mathrm{d}W_t\right|\Big).
\end{aligned}\tag{2.32}$$

Because

$$\begin{aligned}
\mathbb{E}\sup_{r\in[0,s]}\int_0^{r\wedge\tau_n(x)}|b(X_t^{\nu,x},\mu)-b(X_t^{\nu,x},\nu)|\,\mathrm{d}t&\le\mathbb{E}\int_0^s\left|b(X_{t\wedge\tau_n(x)}^{\nu,x},\mu)-b(X_{t\wedge\tau_n(x)}^{\nu,x},\nu)\right|\mathrm{d}t\\
&\le s\sup_{|x|\le n}|b(x,\mu)-b(x,\nu)|
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\sup_{r\in[0,s]}\left|\int_0^{r\wedge\tau_n(x)}(\sigma(X_t^{\nu,x},\mu)-\sigma(X_t^{\nu,x},\nu))\,\mathrm{d}W_t\right|\\
&\quad\le\mathbb{E}\left(\int_0^{s\wedge\tau_n(x)}\|\sigma(X_t^{\nu,x},\mu)-\sigma(X_t^{\nu,x},\nu)\|_{HS}^2\mathrm{d}t\right)^{\frac12}\\
&\quad\le\mathbb{E}\left(\int_0^{s}\|\sigma(X_{t\wedge\tau_n(x)}^{\nu,x},\mu)-\sigma(X_{t\wedge\tau_n(x)}^{\nu,x},\nu)\|_{HS}^2\mathrm{d}t\right)^{\frac12}\\
&\quad\le\sqrt{s}\sup_{|x|\le n}\|\sigma(x,\mu)-\sigma(x,\nu)\|_{HS},
\end{aligned}$$

we have that

$$I_{1,n}\le\|\nabla f\|_\infty\int_{|x|\le\frac n2}\mathbb{E}\sup_{t\in[0,t_\epsilon]}\left|X_{t\wedge\tau_n(x)}^{\mu,x}-X_{t\wedge\tau_n(x)}^{\nu,x}\right|\mathcal{T}_\nu(\mathrm{d}x)$$

$$\leq 2e^{(2K_n+K_n^2)t_\epsilon}\|\nabla f\|_\infty \times\left(t_\epsilon \sup_{|x|\leq n}|b(x,\mu)-b(x,\nu)|+\sqrt{t_\epsilon}\sup_{|x|\leq n}\|\sigma(x,\mu)-\sigma(x,\nu)\|_{HS}\right).$$

Recalling that t_ϵ, K_n are independent of $\nu\in\mathscr{P}_M^{1+r_2}$. Then this, together with (2.5), implies that $\lim_{\nu\xrightarrow{w}\mu} I_{1,n}=0$ as $\nu\xrightarrow{w}\mu$ in $\mathscr{P}_M^{1+r_2}$.

Hence, letting $\nu\xrightarrow{w}\mu$ in $\mathscr{P}_M^{1+r_2}$ and $n\to+\infty$, we find that

$$\lim_{\nu\xrightarrow{w}\mu}\mathcal{T}_\nu\left(\left|P_{t_\epsilon}^\mu f-P_{t_\epsilon}^\nu f\right|\right)=0.$$

This, together with (2.25) and (2.24), yields that for all $\epsilon>0$

$$\lim_{\nu\xrightarrow{w}\mu}|\mathcal{T}_\mu(f)-\mathcal{T}_\nu(f)|\leq\epsilon.$$

Therefore, the proof is complete. □

Proof of Theorem 2.2. By Lemma 2.7, there is some $M_0>0$ such that for each $M\geq M_0$, $\mathscr{P}_M^{1+r_2}$ is invariant under $\mathcal{T}$. Since $r_2>0$, by [3, Theorem 5.5] or [1, Theorem 5.5], $\mathscr{P}_M^{1+r_2}$ is a convex and compact subset of $(\mathscr{P}^1, W)$. By Lemma 2.9, $\mathcal{T}$ is weak continuous in $\mathscr{P}_M^{1+r_2}$. Since $\mathscr{P}_M^{1+r_2}$ is compact in $\mathscr{P}_1$, $\mathcal{T}$ is continuous in $\mathscr{P}_M^{1+r_2}$ w.r.t. the Kantorovich-Rubinstein norm. Hence, the Schauder fixed point theorem yields that $\mathcal{T}$ has a fixed in $\mathscr{P}_M^{1+r_2}$. □

Proof of Corollary 2.4. Because σ is bounded and satisfies (H2) and (H3), we focus on the drift term below when verifying (H1)–(H3). It follows from (2.10) that

$$\begin{aligned}&\langle\nabla V(x)-\nabla V(0),x\rangle+\mu(\langle\nabla F(x-\cdot)-\nabla F(0-\cdot),x\rangle)\\&=\int_0^1\int_{\mathbb{R}^d}\left(\langle\nabla^2V(\theta x)x,x\rangle+\langle\nabla^2F(\theta x-y)x,x\rangle\right)\mu(\mathrm{d}y)\mathrm{d}\theta\\&\geq\left(\beta_2|x|^2-\beta_1|x|+\beta_0-\beta_3\mu(|\cdot|^2)\right)|x|^2.\end{aligned}\tag{2.33}$$

Let $\delta_1=\beta_3/(2\beta_2)$. It follows by the Hölder inequality that

$$\beta_3\mu(|\cdot|^2)|x|^2\leq\delta_1\beta_2|x|^4+\frac{\beta_3^2}{4\delta_1\beta_2}\|\mu\|_4^4.\tag{2.34}$$

By (2.8) and the Hölder inequality, for $\delta_2=\alpha_2/(4\beta_2)$, we have that

$$\begin{aligned}|\mu(\langle\nabla F(0-\cdot),x\rangle)|&\leq\alpha_1|x|+\alpha_2\|\mu\|_3^3|x|\leq\alpha_1|x|+\alpha_2\|\mu\|_4^3|x|\\&\leq\alpha_1|x|+\frac{3\alpha_2^{\frac{4}{3}}}{4(4\delta_2\beta_2)^{\frac{1}{3}}}\|\mu\|_4^4+\delta_2\beta_2|x|^4.\end{aligned}\tag{2.35}$$

Putting (2.35) and (2.34) into (2.33), we have that

$$\begin{aligned}\langle\nabla V(x)+\nabla F*\mu(x),x\rangle&\geq(1-\delta_1-\delta_2)\beta_2|x|^4-\beta_1|x|^3+\beta_0|x|^2-(|\nabla V(0)|+\alpha_1)|x|\\&\quad-\left(\frac{3\alpha_2^{\frac{4}{3}}}{4(4\delta_2\beta_2)^{\frac{1}{3}}}+\frac{\beta_3^2}{4\delta_1\beta_2}\right)\|\mu\|_4^4\\&=\left(\beta_2-\frac{\beta_3}{2}-\frac{\alpha_2}{4}\right)|x|^4-\beta_1|x|^3+\beta_0|x|^2\\&\quad-(|\nabla V(0)|+\alpha_1)|x|-\left(\frac{3\alpha_2}{4}+\frac{\beta_3}{2}\right)\|\mu\|_4^4.\end{aligned}\tag{2.36}$$

Since $\beta_2 > \alpha_2 + \beta_3$, it holds that

$$\beta_2 - \frac{\alpha_2}{4} - \frac{\beta_3}{2} > \frac{3\alpha_2}{4} + \frac{\beta_3}{2}.$$

This, together with (2.36) and the Hölder inequality, yields that there are positive constants C_1, C_2, C_3 with $C_1 > C_3$ such that

$$2\langle b(x,\mu), x\rangle \le -C_1|x|^4 + C_2 + C_3\|\mu\|_4^4.$$

Combining this with that σ is bounded on $\mathbb{R}^d \times \mathscr{P}$, we have that (H1) holds with $r_1 = r_2 = 3, r_3 = 4 = 1 + r_2$ and $C_1 > C_3$. Hence, to verify the assumptions of Theorem 2.2, we just need to check (H2) and (H3).

Because V is twice continuous differentiable, ∇V is local Lipschitz. By (2.9) and the mean value theorem, we have that

$$\begin{aligned}
&\mu\left(|\nabla F(x-\cdot) - \nabla F(y-\cdot)|\right)\\
&\quad\le |x-y|\int_0^1 \mu\left(\|\nabla^2 F(x+\theta(y-x)-\cdot)\|\right)\mathrm{d}\theta\\
&\quad\le \alpha_3|x-y|\left(1 + 2^{(\gamma_0-1)^+}\left(\int_0^1((1-\theta)|x|+\theta|y|)^{\gamma_0}\mathrm{d}\theta + \|\mu\|_{\gamma_0}^{\gamma_0}\right)\right)\\
&\quad\le C_{\alpha_3,\gamma_0}|x-y|\left(1+|x|^{\gamma_0}+|y|^{\gamma_0}+\|\mu\|_4^{\gamma_0}\right).
\end{aligned}$$

Thus (2.2) holds. The inequality (2.3) follows from (2.7) and (2.8) directly. Hence, (H2) holds.

For any $\pi \in \mathscr{C}(\mu,\nu)$, we have by (2.9) that

$$\begin{aligned}
|b(x,\mu) - b(x,\nu)| &= |\mu(\nabla F(x-\cdot)) - \nu(\nabla F(x-\cdot))|\\
&= \left|\int_{\mathbb{R}^d\times\mathbb{R}^d}(\nabla F(x-y_1) - \nabla F(x-y_2))\,\pi(\mathrm{d}y_1,\mathrm{d}y_2)\right|\\
&\le C_{\alpha_3,\gamma_0}\int_{\mathbb{R}^d\times\mathbb{R}^d}(1+|x|^{\gamma_0}+|y_1|^{\gamma_0}+|y_2|^{\gamma_0})\,|y_1-y_2|\pi(\mathrm{d}y_1,\mathrm{d}y_2)\\
&\le C_{\alpha_3,\gamma_0}(1+|x|^{\gamma_0})\int_{\mathbb{R}^d\times\mathbb{R}^d}|y_1-y_2|\pi(\mathrm{d}y_1,\mathrm{d}y_2)\\
&\quad + C_{\alpha_3,\gamma_0}\int_{\mathbb{R}^d\times\mathbb{R}^d}(|y_1|^{\gamma_0}+|y_2|^{\gamma_0})|y_1-y_2|\pi(\mathrm{d}y_1,\mathrm{d}y_2)\\
&\le C_{\alpha_3,\gamma_0}(1+|x|^{\gamma_0})\int_{\mathbb{R}^d\times\mathbb{R}^d}|y_1-y_2|\pi(\mathrm{d}y_1,\mathrm{d}y_2)\\
&\quad + \tilde{C}_{\alpha_3,\gamma_0}(\|\mu\|_4^{\gamma_0}+\|\nu\|_4^{\gamma_0})\left(\int_{\mathbb{R}^d\times\mathbb{R}^d}|y_1-y_2|^{\frac{4}{4-\gamma_0}}\pi(\mathrm{d}y_1,\mathrm{d}y_2)\right)^{\frac{4-\gamma_0}{4}}.
\end{aligned}$$

Thus

$$|b(x,\mu)-b(x,\nu)| \le C_{\alpha_3,\gamma_0}(1+|x|^{\gamma_0})W(\mu,\nu) + \tilde{C}_{\alpha_3,\gamma_0}(\|\mu\|_4^{\gamma_0}+\|\nu\|_4^{\gamma_0})W_{\frac{4}{4-\gamma_0}}(\mu,\nu),$$

where $W_{\frac{4}{4-\gamma_0}}(\mu,\nu)$ is the $\frac{4}{4-\gamma_0}$-Wasserstein distance. Note that $\frac{4}{4-\gamma_0} < 1 + r_2$ since $\gamma_0 < 3$ and $r_2 = 3$. For $\mu_m, \mu \in \mathscr{P}_M^{1+r_2}$ with $\mu_m \xrightarrow{w} \mu$, we have by [3, Theorem 5.6] or [1, Theorem 5.5] that

$$\lim_{m\to+\infty}\Big(W(\mu_m,\mu) + W_{\frac{4}{4-\gamma_0}}(\mu,\mu_m)\Big) = 0.$$

Hence, (H3) holds. □

2.3 Proof of Theorem 2.5

For given $\mu \in \mathscr{P}^{1+r_2}$, under the assumptions of Theorem 2.5, it has been proved by [31, Theorem 2.10] that (2.11) has a unique strong solution and the associated semigroup P_t^μ has a unique invariant probability measure. Though σ is assumed to be Hölder continuous in [31, (H^σ)], we remark here that this assumption can be replaced by that σ is uniformly continuous according to [30, Theorem 3.2]. We also denote by $\mathcal{T}_\mu$ the associated invariant probability measure of P_t^μ and by $\mathcal{T}$ the mapping on $\mathscr{P}^{1+r_2}$.

Let $a(x,\mu) = (\sigma\sigma^*)(x,\mu)$, and let u_μ be the solution of the following equation

$$\frac{1}{2}\operatorname{tr}\left(a(\cdot,\mu)\nabla^2 u_\mu\right)(x) + (\nabla_{b_1(x,\mu)}u_\mu)(x) = \lambda u_\mu(x) - b_1(x,\mu),\ x \in \mathbb{R}^d. \tag{2.37}$$

According to [32, Theorem 2.1] or [31, Theorem 7.5], it follows from (H4) and (H5) that $u_\mu \in \tilde{H}^{2,p} \cap C^{1+\epsilon}(\mathbb{R}^d)$ for some $\epsilon > 0$ and

$$\lim_{\lambda\to+\infty} \sup_{\mu\in\mathscr{P}^{1+r_2}} \left(\|u_\mu\|_\infty + \|\nabla u_\mu\|_\infty\right) = 0. \tag{2.38}$$

Let $U_\mu(x) = x + u_\mu(x)$. Then by the Itô formula (see e.g. [30, Lemma 4.1 (iii)] or [32, Lemma 3.3]), we get that

$$\mathrm{d}U_\mu(X_t^\mu) = \left(\nabla U_\mu b_0(\cdot,\mu) + \lambda u_\mu\right)(X_t^\mu)\mathrm{d}t + (\nabla U_\mu \sigma(\cdot,\mu))(X_t^\mu)\mathrm{d}W_t. \tag{2.39}$$

By (2.38), we choose $\lambda_0 > 0$ such that for $\lambda > \lambda_0$, $\sup_{\mu\in\mathscr{P}^{1+r_2}} \|\nabla u_\mu\|_\infty < \frac{1}{2}$. Then U_μ is a diffeomorphism on $\mathbb{R}^d$ and

$$\sup_{\mu\in\mathscr{P}^{1+r_2}} \left(\|\nabla (U_\mu)^{-1}\|_\infty \vee \|\nabla U_\mu\|_\infty\right) \le \sup_{\mu\in\mathscr{P}^{1+r_2}} \frac{1}{1-\|\nabla u_\mu\|_\infty} \le 2. \tag{2.40}$$

Lemma 2.10. *Under the assumptions of Theorem 2.5, there exists $\lambda_0 > 0$ and for each $\lambda \ge \lambda_0$, there is $M_0 > 0$ such that for any $M \ge M_0$, $\mathscr{P}_M^{1+r_2}$ is invariant under the mapping $\mathcal{T}$.*

Proof. Let $Y_t^\mu = U_\mu(X_t^\mu)$. Then $Y_0^\mu = U_\mu(X_0^\mu)$, and by (2.39),

$$\mathrm{d}Y_t^\mu = \left(\nabla U_\mu b_0(\cdot,\mu) + \lambda u_\mu\right)(U_\mu^{-1}(Y_t^\mu))\mathrm{d}t + (\nabla U_\mu \sigma(\cdot,\mu))(U_\mu^{-1}(Y_t^\mu))\mathrm{d}W_t. \tag{2.41}$$

We denote by $\hat{P}_t^\mu$ the Markov semigroup associated with (2.41). Define a mapping $\mathcal{T}\circ U^{-1}$ on $\mathscr{P}^{1+r_2}$ as follows

$$(\mathcal{T}\circ U^{-1})_\mu = \mathcal{T}_\mu \circ U_\mu^{-1},\ \mu \in \mathscr{P}^{1+r_2}.$$

Because

$$\mathcal{T}_\mu \circ U_\mu^{-1}(\hat{P}_t^\mu f) = \mathcal{T}_\mu((\hat{P}_t^\mu f)\circ U_\mu) = \mathcal{T}_\mu(P_t^\mu(f\circ U_\mu)) = \mathcal{T}_\mu(f\circ U_\mu) = \mathcal{T}_\mu \circ U_\mu^{-1}(f).$$

The probability measure $(\mathcal{T}\circ U^{-1})_\mu$ is the invariant probability measure of $\hat{P}_t^\mu$.

Next, we check the coefficients of (2.41) are subject to (H1) and (H2) except (2.2).

We first verify (H2) except (2.2). By (2.13) and (2.40), it is clear that the diffusion term $\nabla U_\mu \sigma(\cdot,\mu)$ of (2.41) satisfies (2.4):

$$\sup_{\mu\in\mathscr{P}^{1+r_2}, y\in\mathbb{R}^d} \|(\nabla U_\mu \sigma(y,\mu))(U_\mu^{-1}(y))\|_{HS}^2 \le 4\lambda_2^2. \tag{2.42}$$

For every $x \in \mathbb{R}^d$,

$$|x| - \|u_\mu\|_\infty \le |(U_\mu)^{-1}(x)| = |x - u_\mu((U_\mu)^{-1}(x))| \le |x| + \|u_\mu\|_\infty. \tag{2.43}$$

Then the Jensen inequality and the C-r inequality yield that

$$|U_\mu^{-1}(x)|^{1+r_1} \geq (1-\epsilon_0)^{r_1}|x|^{1+r_1} - \frac{(1-\epsilon_0)^{r_1}}{\epsilon_0^{r_1}}\|u_\mu\|_\infty^{1+r_1},\ \epsilon_0 \in (0,1), \tag{2.44}$$

$$|U_\mu^{-1}(x)|^{r_1} \leq 2^{(r_1-1)^+}|x|^{r_1} + 2^{(r_1-1)^+}\|u_\mu\|_\infty^{r_1}. \tag{2.45}$$

Inequality (2.45), together with (2.14), yields that

$$|b_0(U_\mu^{-1}(x),\mu)| \leq C_5\left[1 + 2^{(r_1-1)^+}(|x|^{r_1} + \|u_\mu\|_\infty^{r_1})\right] + C_6\|\mu\|_{1+r_2}^{\frac{r_3 r_1}{1+r_1}}. \tag{2.46}$$

Noting that $\sup_{\mu\in\mathscr{P}^{1+r_2}}\|u_\mu\|_\infty < \infty$, (2.3) holds.

We then verify (H1) for (2.41). Since (2.42), we focus on the drift term. By (2.1), we have that

$$\begin{aligned}
&\langle(\nabla U_\mu b_0(\cdot,\mu))(U_\mu^{-1}(x)), x\rangle + \lambda\langle u_\mu(U_\mu^{-1}(x)), x\rangle \\
&\quad\leq \langle b_0(U_\mu^{-1}(x),\mu), x\rangle + \|\nabla u_\mu\|_\infty |b_0(U_\mu^{-1}(x),\mu)||x| + \lambda\|u_\mu\|_\infty|x| \\
&\quad\leq \langle b_0(U_\mu^{-1}(x),\mu), U_\mu^{-1}(x)\rangle + \langle b_0(U_\mu^{-1}(x),\mu), u_\mu(U_\mu^{-1}(x))\rangle \\
&\qquad + \|\nabla u_\mu\|_\infty|b_0(U_\mu^{-1}(x),\mu)||x| + \lambda\|u_\mu\|_\infty|x| \\
&\quad\leq -\frac{C_1}{2}|U_\mu^{-1}(x)|^{1+r_1} + \frac{C_2}{2} + \frac{C_3}{2}\|\mu\|_{1+r_2}^{r_3} \\
&\qquad + \|u_\mu\|_\infty|b_0(U_\mu^{-1}(x),\mu)| + \|\nabla u_\mu\|_\infty|b_0(U_\mu^{-1}(x),\mu)||x| + \lambda\|u_\mu\|_\infty|x|.
\end{aligned} \tag{2.47}$$

The Hölder inequality, (2.46) and (2.14) imply that for any positive $\epsilon_1, \epsilon_2, \epsilon_3$,

$$\begin{aligned}
&\|u_\mu\|_\infty|b_0(U_\mu^{-1}(x),\mu)| + \|\nabla u_\mu\|_\infty|b_0(U_\mu^{-1}(x),\mu)||x| \\
&\quad\leq \left(\frac{r_1\|u_\mu\|_\infty}{1+r_1}\epsilon_1^{\frac{1+r_1}{r_1}} + \|\nabla u_\mu\|_\infty\left(C_5 2^{(r_1-1)^+} + \frac{\epsilon_2^{1+r_1}}{1+r_1} + \frac{C_6}{1+r_2}\right)\right)|x|^{1+r_1} \\
&\qquad + C_5\|u_\mu\|_\infty^{1+r_1}2^{(r_1-1)^+} + (1+r_1)^{-1}\|u_\mu\|_\infty\left(\frac{C_5 2^{(r_1-1)^+}}{\epsilon_1}\right)^{1+r_1} \\
&\qquad + \frac{r_1\|\nabla u_\mu\|_\infty}{1+r_1}\left(\frac{C_5(1+2^{(r_1-1)^+}\|u_\mu\|^{r_1})}{\epsilon_2}\right)^{\frac{1+r_1}{r_1}} + \frac{\|u_\mu\|_\infty C_6}{(1+r_1)\epsilon_3^{1+r_1}} \\
&\qquad + \left(\frac{r_1 C_6\|u_\mu\|_\infty\epsilon_3^{\frac{1+r_1}{r_1}}}{1+r_1} + \frac{C_6\|\nabla u_\mu\|_\infty r_1}{1+r_1}\right)\|\mu\|_{1+r_2}^{r_3} \\
&\quad =: \bar{C}_1(\epsilon_1,\epsilon_2,\mu,\lambda)|x|^{1+r_1} + \bar{C}_2(\epsilon_1,\epsilon_2,\epsilon_3,\mu,\lambda) + \bar{C}_3(\epsilon_3,\mu,\lambda)\|\mu\|_{1+r_2}^{r_3}.
\end{aligned} \tag{2.48}$$

Putting this and (2.44) into (2.47), and by using the Hölder inequality, we have that

$$\begin{aligned}
&\langle(\nabla U_\mu b_0(\cdot,\mu))(U_\mu^{-1}(x)), x\rangle + \lambda\langle u_\mu(U_\mu^{-1}(x)), x\rangle \\
&\quad\leq -\left(\frac{C_1(1-\epsilon_0)^{r_1}}{2} - \bar{C}_1(\epsilon_1,\epsilon_2,\mu,\lambda) - \frac{\epsilon_4^{1+r_1}}{1+r_1}\right)|x|^{1+r_1} \\
&\qquad + \bar{C}_2(\epsilon_1,\epsilon_2,\epsilon_3,\mu,\lambda) + \frac{C_1\|u_\mu\|_\infty^{1+r_1}(1-\epsilon_0)^{r_1}}{2\epsilon_0^{r_1}} + \frac{r_1(\lambda\|u_\mu\|_\infty)^{\frac{r_1}{1+r_1}}}{(1+r_1)\epsilon_4^{\frac{r_1}{1+r_1}}} \\
&\qquad + \left(\frac{C_3}{2} + \bar{C}_3(\epsilon_3,\mu,\lambda)\right)\|\mu\|_{1+r_2}^{r_3},
\end{aligned}$$

where ϵ_4 is any positive constant. Due to (2.38), we have that

$$\lim_{\lambda\to+\infty}\sup_{\mu\in\mathscr{P}^{1+r_2}}\bar{C}_1(\epsilon_1,\epsilon_2,\mu,\lambda) = 0.$$

Taking into account that ϵ_4 can be arbitrary small, there are $\lambda_0 > 0$ and $\epsilon_4' > 0$ such that for $\lambda > \lambda_0$ and $0 < \epsilon_4 < \epsilon_4'$, it holds that

$$\frac{C_1(1-\epsilon_0)^{r_1}}{2} - \bar{C}_1(\epsilon_1, \epsilon_2, \mu, \lambda) - \frac{\epsilon_4^{1+r_1}}{1+r_1} \geq \tilde{C}_1 > 0$$

with some constant $\tilde{C}_1$ independent of μ. Moreover, for given $\lambda > \lambda_0$ and $\epsilon_0, \epsilon_1, \epsilon_3, \epsilon_4$

$$\sup_{\mu \in \mathscr{P}^{1+r_2}} \left\{ (\bar{C}_1 + \bar{C}_2 + \bar{C}_3)(\epsilon_0, \epsilon_1, \epsilon_2, \epsilon_3, \mu, \lambda) + \frac{C_1 \|u_\mu\|_\infty^{1+r_1}}{2\epsilon_0^{r_1}} + \frac{r_1(\lambda \|u_\mu\|_\infty)^{\frac{r_1}{1+r_1}}}{(1+r_1)\epsilon_4^{\frac{r_1}{1+r_1}}} \right\} < \infty.$$

Hence, we find that there are some positive constants $\tilde{C}_i,\ i = 1, 2, 3$ independent of μ such that

$$\begin{aligned} \langle (\nabla U_\mu b_0(\cdot, \mu))(U_\mu^{-1}(x)), x \rangle &+ \lambda \langle u_\mu(U_\mu^{-1}(x)), x \rangle \\ &\leq -\tilde{C}_1 |x|^{1+r_1} + \tilde{C}_2 + \tilde{C}_3 \|\mu\|_{1+r_2}^{r_3}. \end{aligned} \tag{2.49}$$

In the case that $C_1 > C_3$, due to (2.38), it is clear that

$$\lim_{\lambda \to +\infty} \sup_{\mu \in \mathscr{P}^{1+r_2}} (\bar{C}_1(\epsilon_1, \epsilon_2, \mu, \lambda) + \bar{C}_3(\epsilon_3, \mu, \lambda)) = 0.$$

We can choose $\epsilon_0, \epsilon_1, \epsilon_3, \epsilon_4$ small and a larger $\lambda_0 > 0$ such that for $\lambda > \lambda_0$

$$\sup_{\mu \in \mathscr{P}^{1+r_2}} \frac{\frac{C_3}{2} + \bar{C}_3(\epsilon_3, \mu, \lambda)}{\frac{C_1(1-\epsilon_0)^{r_1}}{2} - \bar{C}_1(\epsilon_1, \epsilon_2, \mu, \lambda) - \frac{\epsilon_4^{1+r_1}}{1+r_1}} < \frac{C_3 + 1}{C_1 + 1} < 1. \tag{2.50}$$

Combining (2.42) with (2.49), (2.50) and (2.46), we have proven that there exists $\lambda_0 > 0$ and for any $\lambda \geq \lambda_0$, coefficients of (2.41) satisfy (H1) with some constants $\hat{C}_i,\ i = 1, 2, 3$ independent of μ, (H2) except (2.2), and

$$\frac{\hat{C}_3}{\hat{C}_1} < \frac{C_3 + 1}{C_1 + 1} < 1, \text{ when } r_3 = 1 + r_1 \text{ and } C_1 > C_3. \tag{2.51}$$

Though we can find an invariant subset for $\mathcal{T} \circ U^{-1}$ in $\mathscr{P}^{1+r_2}$ by Lemma 2.8, we need to prove that $\mathscr{P}_M^{1+r_2}$ with large enough M is an invariant subset of $\mathcal{T}$ instead of $\mathcal{T} \circ U^{-1}$. Following the proof of Lemma 2.8, we have by (2.22) that for $\mathcal{T} \circ U^{-1}$

$$\|(\mathcal{T} \circ U^{-1})_\mu\|_{1+r_2}^{1+r_1} \leq \frac{\hat{C}_2}{\hat{C}_1} + \frac{\hat{C}_3}{\hat{C}_1} \|\mu\|_{1+r_2}^{r_3}.$$

In the case that $r_3 = 1 + r_1$, by (2.51), there exists $\tilde{M}_0 > 0$ such that

$$\frac{\hat{C}_2}{\hat{C}_1} + \frac{\hat{C}_3}{\hat{C}_1} M^{r_3} \leq \frac{C_3 + 1}{C_1 + 1} M^{1+r_1},\ M \geq \tilde{M}_0.$$

In the case that $r_3 < 1 + r_1$, it is clear that there exists $\tilde{\tilde{M}}_0 > 0$ such that

$$\frac{\hat{C}_2}{\hat{C}_1} + \frac{\hat{C}_3}{\hat{C}_1} M^{r_3} \leq \left(\frac{C_3 + 1}{C_1 + 1} \wedge \frac{1}{2} \right) M^{1+r_1},\ M \geq \tilde{\tilde{M}}_0.$$

Consequently, there exists $c_0 \in (0, 1)$ such that for $M \geq \tilde{M}_0 \vee \tilde{\tilde{M}}_0$,

$$\|(\mathcal{T} \circ U^{-1})_\mu\|_{1+r_2}^{1+r_1} \leq \frac{\hat{C}_2}{\hat{C}_1} + \frac{\hat{C}_3}{\hat{C}_1} M^{r_3} \leq c_0 M^{1+r_1},\ \mu \in \mathscr{P}_M^{1+r_2}.$$

This, together with (2.43), implies that

$$\begin{aligned}\|\mathcal{T}_\mu\|_{1+r_2} &= \left(\mathcal{T}_\mu \circ U_\mu^{-1}(|U_\mu^{-1}(\cdot)|^{1+r_2})\right)^{\frac{1}{1+r_2}}\\ &\le \left(\mathcal{T}_\mu \circ U_\mu^{-1}(|\cdot|^{1+r_2})\right)^{\frac{1}{1+r_2}} + \|u_\mu\|_\infty\\ &\le c_0^{\frac{1}{1+r_1}} M + \sup_{\mu\in\mathscr{P}^{1+r_2}} \|u_\mu\|_\infty,\ \mu\in\mathscr{P}_M^{1+r_2}.\end{aligned}$$

Thus, for

$$M \ge M_0 := \tilde{M}_0 \vee \tilde{\tilde{M}}_0 \vee \frac{\sup_{\mu\in\mathscr{P}^{1+r_2}}\|u_\mu\|_\infty}{(1-c_0^{\frac{1}{1+r_1}})},$$

we have that

$$\|\mathcal{T}_\mu\|_{1+r_2} \le c_0^{\frac{1}{1+r_1}} M + (1-c_0^{\frac{1}{1+r_1}})M = M,\ \mu\in\mathscr{P}_M^{1+r_2}.$$

Therefore, $\mathscr{P}_M^{1+r_2}$ is invariant under $\mathcal{T}$. □

According to the proof of Lemma 2.10, (H1) holds for (2.41). It follows from (2.40), (H5) and the continuity of ∇U_μ that $(\nabla U_\mu \sigma(\cdot,\mu))(U_\mu^{-1}(\cdot))$ is continuous, non-degenerate and bounded. Combining this with (H4) and (H5), by [31, Theorem 2.1], [30, Theorem 3.1] and the proof of Lemma 2.7, we can prove that P_t^μ is V-uniformly exponential ergodic with some locally bounded V.

We next prove that $\mathcal{T}$ is weakly sequentially continuous in $\mathscr{P}_M^{1+r_2}$. According to the proof of Lemma 2.9, especially (2.24), (2.26), (2.27), (2.29) and (2.30), the weak sequential continuity of $\mathcal{T}$ follows from the lemma below. For $\mu,\nu\in\mathscr{P}$, we let

$$\tau_n^{\mu,\nu}(x) = \inf\left\{t>0 \mid |X_t^{\nu,x}| + |X_t^{\mu,x}| > n\right\}.$$

Lemma 2.11. *Under the assumptions of Theorem 2.5, and let $\mu,\mu_m\in\mathscr{P}_M^{1+r_2}$ such that $\mu_m \xrightarrow{w} \mu$. Then for every $t>0$*

$$\lim_{m\to+\infty}\int_{|x|\le\frac{n}{2}} \mathbb{E}\left(\left|X^{\mu,x}_{t\wedge\tau_n^{\mu,\mu_m}(x)} - X^{\mu_m,x}_{t\wedge\tau_n^{\mu,\mu_m}(x)}\right|\right)\mathcal{T}_{\mu_m}(\mathrm{d}x) = 0. \tag{2.52}$$

Proof. For any $\nu\in\mathscr{P}_M^{1+r_2}$, let $X_t^{n,\nu}$ be the solution of the following SDE

$$\mathrm{d}X_t^{n,\nu} = (b_0\eta_n)(X_t^{n,\nu},\nu)\mathrm{d}t + b_1(X_t^{n,\nu},\nu)\mathrm{d}t + \sigma(X_t^{n,\nu},\nu)\mathrm{d}W_t,\ X_0^{n,\nu}=x, \tag{2.53}$$

where $n\in\mathbb{N}$ with $x>n$, and $\eta_n\in C_b^1(\mathbb{R}^d)$ is a cutoff function with $\mathbb{1}_{[|y|\le n]}\le\eta_n(y)\le\mathbb{1}_{[|y|\le n+1]}$. By the pathwise uniqueness, for $t<\tau_n^{\mu,\nu}(x)$

$$X_t^{n,\mu} = X_t^{\mu,x},\ \ X_t^{n,\nu} = X_t^{\nu,x}.$$

Denote $b_{0n} = b_0\eta_n$. Then

$$\begin{aligned}\mathrm{d}u_\mu(X_t^{n,\nu}) &= \left(\frac{1}{2}\mathrm{tr}\left(a(\cdot,\nu)\nabla^2 u_\mu\right)(X_t^{n,\nu}) + (\nabla_{b_1(\cdot,\nu)}u_\mu + \nabla_{b_{0n}(\cdot,\nu)}u_\mu)(X_t^{n,\nu})\right)\mathrm{d}t\\ &\quad + (\nabla_{\sigma(\cdot,\nu)\mathrm{d}W_t}u_\mu)(X_t^{n,\nu}).\end{aligned}$$

Then by using (2.37) and the Itô formula (see, e.g. [30, Lemma 4.1 (iii)] or [32, Lemma 3.3]), we have that

$$\begin{aligned}\mathrm{d}U_\mu(X_t^{n,\nu}) &= \mathrm{d}X_t^{n,\nu} + \mathrm{d}u_\mu(X_t^{n,\nu})\\ &= (b_{0n}+b_1)(X_t^{n,\nu},\nu)\mathrm{d}t + \sigma(X_t^{n,\nu},\nu)\mathrm{d}W_t + \nabla_{\sigma(X_t^{n,\nu},\nu)\mathrm{d}W_t}u_\mu(X_t^{n,\nu})\end{aligned}$$

$$
\begin{aligned}
&+ \left(\frac{1}{2}\operatorname{tr}\left(a(\cdot,\nu)\nabla^2 u_\mu\right)\right)(X_t^{n,\nu})\mathrm{d}t + \left(\nabla_{b_1(\cdot,\nu)}u_\mu + \nabla_{b_{0n}(\cdot,\nu)}u_\mu\right)(X_t^{n,\nu})\mathrm{d}t\\
&= (b_{0n}+b_1)(X_t^{n,\nu},\nu)\mathrm{d}t + \sigma(X_t^{n,\nu},\nu)\mathrm{d}W_t + \nabla_{\sigma(X_t^{n,\nu},\nu)\mathrm{d}W_t}u_\mu(X_t^\nu)\\
&+ \left(\frac{1}{2}\operatorname{tr}\left(a(\cdot,\mu)\nabla^2 u_\mu\right)(X_t^{n,\nu}) + (\nabla_{b_1(X_t^{n,\nu},\mu)}u_\mu)(X_t^{n,\nu})\right)\mathrm{d}t\\
&+ \frac{1}{2}\operatorname{tr}\left((a(\cdot,\nu)-a(\cdot,\mu))\nabla^2 u_\mu\right)(X_t^{n,\nu})\mathrm{d}t + (\nabla_{b_1(\cdot,\nu)-b_1(\cdot,\mu)}u_\mu)(X_t^{n,\nu})\mathrm{d}t\\
&+ (\nabla_{b_{0n}(\cdot,\nu)}u_\mu)(X_t^{n,\nu})\mathrm{d}t\\
&= \left(b_{0n}(\cdot,\nu) + \nabla_{b_{0n}(\cdot,\nu)}u_\mu + b_1(\cdot,\nu) - b_1(\cdot,\mu) + \lambda u_\mu\right)(X_t^{n,\nu})\mathrm{d}t\\
&+ \frac{1}{2}\operatorname{tr}\left((a(\cdot,\nu)-a(\cdot,\mu))\nabla^2 u_\mu\right)(X_t^{n,\nu})\mathrm{d}t + (\nabla_{b_1(\cdot,\nu)-b_1(\cdot,\mu)}u_\mu)(X_t^{n,\nu})\mathrm{d}t\\
&+ \sigma(X_t^{n,\nu},\nu)\mathrm{d}W_t + \nabla_{\sigma(X_t^{n,\nu},\nu)\mathrm{d}W_t}u_\mu(X_t^{n,\nu})\\
&= \left(\nabla U_\mu b_{0n}(\cdot,\nu) + \nabla U_\mu\left(b_1(\cdot,\nu)-b_1(\cdot,\mu)\right) + \lambda u_\mu\right)(X_t^{n,\nu})\mathrm{d}t\\
&+ \frac{1}{2}\operatorname{tr}\left((a(\cdot,\nu)-a(\cdot,\mu))\nabla^2 u_\mu\right)(X_t^{n,\nu})\mathrm{d}t + \nabla U_\mu(X_t^{n,\nu})\sigma(X_t^{n,\nu},\nu)\mathrm{d}W_t.
\end{aligned}
$$

In particular,

$$
\mathrm{d}U_\mu(X_t^{n,\mu}) = \left(\nabla U_\mu b_{0n}(\cdot,\mu) + \lambda u_\mu\right)(X_t^{n,\mu})\mathrm{d}t + \nabla U_\mu(X_t^{n,\mu})\sigma(X_t^{n,\mu},\mu)\mathrm{d}W_t.
$$

Then

$$
\begin{aligned}
\mathrm{d}\left(U_\mu(X_t^{n,\mu}) - U_\mu(X_t^{n,\nu})\right) &= \left(\nabla U_\mu(X_t^\mu)b_{0n}(X_t^\mu,\mu) - \nabla U_\mu(X_t^\nu)b_{0n}(X_t^\nu,\nu)\right)\mathrm{d}t\\
&+ \nabla U_\mu(X_t^{n,\nu})\left(b_1(X_t^{n,\nu},\mu) - b_1(X_t^{n,\nu},\nu)\right)\mathrm{d}t\\
&+ \lambda\left(u_\mu(X_t^{n,\mu}) - u_\mu(X_t^{n,\nu})\right)\mathrm{d}t\\
&+ \frac{1}{2}\operatorname{tr}\left((a(\cdot,\mu)-a(\cdot,\nu))\nabla^2 u_\mu\right)(X_t^{n,\nu})\mathrm{d}t\\
&+ \left((\nabla U_\mu\sigma(\cdot,\mu))(X_t^{n,\mu}) - (\nabla U_\mu\sigma(\cdot,\nu))(X_t^{n,\nu})\right)\mathrm{d}W_t.
\end{aligned}
$$

Then

$$
\begin{aligned}
&\mathrm{d}|U_\mu(X_t^{n,\mu}) - U_\mu(X_t^{n,\nu})|^2\\
&= 2\left\langle(\nabla U_\mu b_{0n}(\cdot,\mu))(X_t^{n,\mu}) - (\nabla U_\mu b_{0n}(\cdot,\nu))(X_t^{n,\nu}), U_\mu(X_t^{n,\mu}) - U_\mu(X_t^{n,\nu})\right\rangle\mathrm{d}t\\
&+ 2\left\langle\left(\nabla U_\mu\left(b_1(\cdot,\mu) - b_1(\cdot,\nu)\right)\right)(X_t^{n,\nu}), U_\mu(X_t^{n,\mu}) - U_\mu(X_t^{n,\nu})\right\rangle\mathrm{d}t\\
&+ 2\lambda\left\langle u_\mu(X_t^{n,\mu}) - u_\mu(X_t^{n,\nu}), U_\mu(X_t^{n,\mu}) - U_\mu(X_t^{n,\nu})\right\rangle\mathrm{d}t\\
&+ \left\langle\operatorname{tr}\left((a(\cdot,\mu)-a(\cdot,\nu))\nabla^2 u_\mu\right)(X_t^{n,\nu}), U_\mu(X_t^{n,\mu}) - U_\mu(X_t^{n,\nu})\right\rangle\mathrm{d}t\\
&+ \|\nabla U_\mu(X_t^{n,\mu})\sigma(X_t^{n,\mu},\mu) - \nabla U_\mu(X_t^{n,\nu})\sigma(X_t^{n,\nu},\nu)\|_{HS}^2\mathrm{d}t\\
&+ 2\left\langle U_\mu(X_t^{n,\mu}) - U_\mu(X_t^{n,\nu}), \left((\nabla U_\mu\sigma(\cdot,\mu))(X_t^{n,\mu}) - (\nabla U_\mu\sigma(\cdot,\nu))(X_t^{n,\nu})\right)\mathrm{d}W_t\right\rangle\\
&= (I_1+I_2+I_3+I_4+I_5)\mathrm{d}t + \mathrm{d}M_t,
\end{aligned}
$$

where

$$
M_t = 2\int_0^t\left\langle U_\mu(X_s^{n,\mu}) - U_\mu(X_s^{n,\nu}), \left((\nabla U_\mu\sigma(\cdot,\mu))(X_s^{n,\mu}) - (\nabla U_\mu\sigma(\cdot,\nu))(X_s^{n,\nu})\right)\mathrm{d}W_t\right\rangle.
$$

We choose $\lambda > 0$ large enough such that $\frac{1}{2} \le \|\nabla U_\mu\|_\infty \le \frac{3}{2}$. By the Krylov estimate, see e.g. [31, Theorem 5.6] or [32, Theorem 3.1], the distributions of $X_t^{n,\mu}$ and $X_t^{n,\nu}$ are absolutely w.r.t. the Lebesgue measure for almost every $t > 0$. Combining this with (2.2) and $\nabla U_\mu \in C_b(\mathbb{R}^d;\mathbb{R}^d\otimes\mathbb{R}^d)$, we have by [30, Lemma 2.1] that for each $n > 0$, there is $\hat{C}_n > 0$ independent of x such that for any $t \le \tau_n^{\mu,\nu}(x)$

$$
I_1 = 2\left\langle\left(\nabla U_\mu(X_t^{n,\mu}) - \nabla U_\mu(X_t^{n,\nu})\right)b_{0n}(X_t^{n,\mu},\mu), U_\mu(X_t^{n,\mu}) - U_\mu(X^{n,\nu})\right\rangle
$$

$$
\begin{aligned}
&+2\left\langle \nabla U_\mu(X_t^{n,\nu})\left(b_{0n}(X^{n,\mu},\mu)-b_{0n}(X^{n,\nu},\mu)\right), U_\mu(X_t^{n,\mu})-U_\mu(X_t^{n,\nu})\right\rangle\\
&+2\left\langle \nabla U_\mu(X_t^{n,\nu})\left(b_{0n}(X^{n,\nu},\mu)-b_{0n}(X^{n,\nu},\nu)\right), U_\mu(X_t^{n,\mu})-U_\mu(X_t^{n,\nu})\right\rangle\\
\le&\ \frac{3C}{2}\left(\sup_{|x|\le n}|b_0(x)|\right)\left(|\mathcal{M}_1(\nabla^2 u_\mu)(X_t^{n,\mu})|+|\mathcal{M}_1(\nabla^2 u_\mu)(X_t^{n,\nu})|+\frac{3}{2}\right)|X_t^{n,\mu}-X_t^{n,\nu}|^2\\
&+\frac{9}{2}K_n|X_t^{n,\mu}-X_t^{n,\nu}|^2+\frac{9}{2}|X_t^{n,\mu}-X_t^{n,\nu}|\sup_{|x|\le n}|b_0(x,\mu)-b_0(x,\mu)|\\
\le&\ \hat{C}_n\left(|\mathcal{M}_1(\nabla^2 u_\mu)(X_t^{n,\mu})|+|\mathcal{M}_1(\nabla^2 u_\mu)(X_t^{n,\nu})|+1\right)|X_t^{n,\mu}-X_t^{n,\nu}|^2\\
&+\sup_{|x|\le n}|b_0(x,\mu)-b_0(x,\mu)|^2.
\end{aligned}
$$

For I_2,

$$
\begin{aligned}
I_2 &\le 9\left|b_1(X_t^{n,\nu},\mu)-b_1(X_t^{n,\nu},\nu)\right|\cdot|X_t^{n,\mu}-X_t^{n,\nu}|\\
&\le \frac{9}{2}\left|b_1(X_t^{n,\nu},\mu)-b_1(X_t^{n,\nu},\nu)\right|^2+\frac{9}{2}|X_t^{n,\mu}-X_t^{n,\nu}|^2
\end{aligned}
$$

For I_3,

$$
I_3\le \frac{3\lambda}{2}|X_t^{n,\mu}-X_t^{n,\nu}|^2.
$$

For I_4,

$$
\begin{aligned}
I_4 &\le \frac{3}{2}\|a(X_t^{n,\nu},\mu)-a(X_t^{n,\nu},\nu)\|_{HS}\|\nabla^2 u_\mu(X_t^{n,\nu})\|_{HS}|X_t^{n,\mu}-X_t^{n,\nu}|\\
&\le 3\|\sigma\|_\infty\|\nabla^2 u_\mu(X_t^{n,\nu})\|_{HS}\|\sigma(X_t^{\nu},\mu)-\sigma(X_t^{\nu},\nu)\|_{HS}|X_t^{n,\mu}-X_t^{n,\nu}|\\
&\le \frac{3}{2}\|\nabla^2 u_\mu(X_t^{n,\nu})\|_{HS}^2|X_t^{n,\mu}-X_t^{n,\nu}|^2+\frac{3}{2}\|\sigma\|_\infty^2\|\sigma(X_t^{n,\nu},\mu)-\sigma(X_t^{n,\nu},\nu)\|_{HS}^2.
\end{aligned}
$$

For I_5,

$$
\begin{aligned}
I_5 &\le \|\sigma\|_\infty^2\|\nabla U_\mu(X_t^{n,\mu})-\nabla U_\mu(X_t^{n,\nu})\|_{HS}^2+\frac{9}{4}\|\sigma(X_t^{n,\mu},\mu)-\sigma(X_t^{n,\nu},\nu)\|_{HS}^2\\
&\le C\|\sigma\|_\infty^2\left(\|\mathcal{M}_1(\nabla^2 u_\mu)(X_t^{\mu})\|_{HS}+\|\mathcal{M}_1(\nabla u_\mu)(X_t^{\nu})\|_{HS}^2+\frac{3}{2}\right)|X_t^{\mu}-X_t^{\nu}|^2\\
&\quad+C\left(\|\mathcal{M}_1(\nabla\sigma(\cdot,\mu))(X_t^{n,\mu})\|_{HS}^2+\|\mathcal{M}_1(\nabla\sigma(\cdot,\mu))(X_t^{n,\nu})\|_{HS}^2+\|\sigma\|_\infty^2\right)|X_t^{n,\mu}-X_t^{n,\nu}|^2\\
&\quad+\frac{9}{2}\|\sigma(X_t^{n,\nu},\mu)-\sigma(X_t^{n,\nu},\nu)\|_{HS}^2.
\end{aligned}
$$

Hence, there exist $C_1'>0$ such that

$$
\begin{aligned}
|U_\mu(X^{\mu}_{t\wedge\tau_n^{\mu,\nu}(x)})-U_\mu(X^{\nu}_{t\wedge\tau_n^{\mu,\nu}(x)})|^2 &= |U_\mu(X^{n,\mu}_{t\wedge\tau_n^{\mu,\nu}(x)})-U_\mu(X^{n,\nu}_{t\wedge\tau_n^{\mu,\nu}(x)})|^2\\
&\le \int_0^{t\wedge\tau_n^{\mu,\nu}(x)}|X_s^{n,\mu}-X_s^{n,\nu}|^2\mathrm{d}A^{\mu,\nu}_{n,s}+t\sup_{|x|\le n}|b_0(x,\mu)-b_0(x,\mu)|^2\\
&\quad+\frac{9}{2}\int_0^{t\wedge\tau_n^{\mu,\nu}(x)}|b_1(X_s^{n,\nu},\mu)-b_1(X_s^{n,\nu},\nu)|^2\mathrm{d}s\\
&\quad+C_1't\sup_{|x|\le n}\|\sigma(x,\mu)-\sigma(x,\nu)\|_{HS}^2+M_{t\wedge\tau_n^{\mu,\nu}(x)},
\end{aligned}
$$

where

$$
A^{\mu,\nu}_{n,t}=\int_0^t\Big\{\tilde{C}_n\left(1+\|\mathcal{M}_1(\nabla^2 u_\mu)(X_s^{n,\mu})\|_{HS}^2+\|\mathcal{M}_1(\nabla^2 u_\mu)(X_s^{n,\nu})\|_{HS}^2\right)
$$

$$+ C_2' \left(1 + \|\mathcal{M}_1(\nabla\sigma(\cdot,\mu))(X_s^{n,\mu})\|_{HS}^2 + (\|\nabla^2 u_\mu\|_{HS}^2 + \|\mathcal{M}_1(\nabla\sigma(\cdot,\mu))\|_{HS}^2)(X_s^{n,\nu})\right) \Big\}\mathrm{d}s$$

for some positive constants C_2' and $\hat{C}_n$ which depends on n and is independent of x. By the stochastic Gronwall lemma, we have that

$$\begin{aligned}
&\left(\mathbb{E}|X^\mu_{t\wedge\tau_n^{\mu,\nu}(x)} - X^\nu_{t\wedge\tau_n^{\mu,\nu}(x)}|\right)^2 \le 4\left(\mathbb{E}|U_\mu(X^\mu_{t\wedge\tau_n^{\mu,\nu}(x)}) - U_\mu(X^\nu_{t\wedge\tau_n^{\mu,\nu}(x)})|\right)^2 \\
&\le 2c_p C_2' \left(\mathbb{E}e^{\frac{p}{p-1}A^{\mu,\nu}_{n,t\wedge\tau_n^{\mu,\nu}(x)}}\right)^{\frac{1-p}{p}} \Big\{ t\sup_{|x|\le n}\|\sigma(x,\mu)-\sigma(x,\nu)\|_{HS}^2 \\
&\quad + t\sup_{|x|\le n}|b_0(x,\mu)-b_0(x,\mu)|^2 + \mathbb{E}\int_0^{t\wedge\tau_n^{\mu,\nu}(x)} |b_1(X_s^{n,\nu},\mu) - b_1(X_s^{n,\nu},\nu)|^2\mathrm{d}s \Big\}.
\end{aligned}$$

It is clear that

$$\begin{aligned}
&\int_0^{t\wedge\tau_n^{\mu,\nu}(x)} |b_1(X_s^{n,\nu},\mu) - b_1(X_s^{n,\nu},\nu)|^2\mathrm{d}s \\
&\quad = \int_0^{t\wedge\tau_n^{\mu,\nu}(x)} |b_1(X_s^{n,\nu},\mu) - b_1(X_s^{n,\nu},\nu)|^2 \mathbb{1}_{[|X_s^{n,\nu}|\le n]}\mathrm{d}s \\
&\quad = \int_0^{t\wedge\tau_n^{\mu,\nu}(x)} |b_{1n}(X_s^{n,\nu},\mu) - b_{1,n}(X_s^{n,\nu},\nu)|^2\mathrm{d}s,
\end{aligned}$$

where $b_{1n}(x,\nu) = b_1(x,\nu)\mathbb{1}_{[|x|\le n]}$. For $q > \frac{2p}{p-d}$, it follows from the Krylov estimate (see e.g. [30, Lemma 4.1 (i)] that

$$\begin{aligned}
\mathbb{E}\int_0^{t\wedge\tau_n^{\mu,\nu}(x)} |b_{1n}(X_s^{n,\nu},\mu) - b_{1n}(X_s^{n,\nu},\nu)|^2\mathrm{d}s &\le Ct^{\frac{2}{q}}\|b_{1n}(\cdot,\mu) - b_{1n}(\cdot,\nu)\|_{\tilde{L}^p}^2 \\
&\le Ct^{\frac{2}{q}}\|b_{1n}(\cdot,\mu) - b_{1n}(\cdot,\nu)\|_{L^p}^2,
\end{aligned}$$

where the constant C depends on $p, q, t, n, \|b_{1n}(\cdot,\nu)\|_{L^p}$. For fixed $\mu \in \mathscr{P}_M^{1+r_2}$ and a sequence $\mu_m \in \mathscr{P}_M^{1+r_2}$ such that $\mu_m \xrightarrow{w} \mu$, we have that

$$\lim_{m\to+\infty} \|b_{1n}(\cdot,\mu_m) - b_{1n}(\cdot,\mu)\|_{L^p} = 0.$$

Then the sequence $\{\|b_{1n}(\cdot,\mu_m)\|_{L^p}\}_{m\ge1}$ is bounded for any $n \ge 1$. Moreover, combining this with Khasminskii's estimate (see e.g. [30, Lemma 4.1 (ii)]) for X_t^{n,μ_m} and $X_t^{n,\mu}$, we find that

$$\sup_{x\in\mathbb{R}^d}\sup_m \mathbb{E}e^{\frac{p}{p-1}A^{\mu,\mu_m}_{n,t\wedge\tau_n^{\mu,\mu_m}(x)}} \le \sup_{x\in\mathbb{R}^d}\sup_m \mathbb{E}e^{\frac{p}{p-1}A^{\mu,\mu_m}_{n,t}} < \infty.$$

Hence,

$$\begin{aligned}
&\overline{\lim_{m\to+\infty}} \int_{|x|\le\frac{n}{2}} \mathbb{E}\left(\left|X^{\mu,x}_{t\wedge\tau_n^{\mu,\mu_m}(x)} - X^{\mu_m,x}_{t\wedge\tau_n^{\mu,\mu_m}(x)}\right|\right)\mathcal{T}_{\mu_m}(\mathrm{d}x) \\
&\le C_n \overline{\lim_{m\to+\infty}} \Big[\sqrt{t}\sup_{|x|\le n}\left(\|\sigma(\cdot,\mu)-\sigma(\cdot,\mu_m)\|_{HS} + |b_0(\cdot,\mu) - b_0(\cdot,\mu_m)|\right)(x) \\
&\quad + t^{\frac{1}{q}}\|b_{1n}(\cdot,\mu) - b_{1n}(\cdot,\mu_m)\|_{L^p}\Big] \\
&= 0.
\end{aligned}$$

Therefore the proof is complete. □

Proof of Theorem 2.5. So far, we have proved that there exists $M_0 > 0$ such that for any $M \ge M_0$, $\mathcal{T}$ maps $\mathscr{P}_M^{1+r_2}$ into $\mathscr{P}_M^{1+r_2}$ and $\mathcal{T}$ is continuous in $\mathscr{P}_M^{1+r_2}$ w.r.t. to the Kantorovich-Rubinstein distance. Therefore, the assertion of this theorem follows from the Schauder fixed point theorem as proving in Theorem 2.2. □

3 Non-uniqueness

For $a \in \mathbb{R}^d$, denote by μ_a the shifted probability of μ by a:

$$\mu_a(f) = \mu(f(\cdot - a)).$$

Next theorem is devoted to a sufficient condition to find a stationary distribution concentrated around $a \in \mathbb{R}^d$. For $\kappa_2 \geq \kappa_1 > 0$ and $0 < \gamma_1 \leq \gamma_2 < 1 + r_2$, we denote

$$\mathscr{P}^{\gamma_1,\gamma_2}_{a,\kappa_1,\kappa_2} = \{\mu \in \mathscr{P}^{1+r_2} \mid \|\mu_a\|_{\gamma_1} \leq \kappa_1, \|\mu_a\|_{\gamma_2} \leq \kappa_2\},$$
$$\mathscr{D}_{\kappa_1,\kappa_2} = \{(w_1, w_2) \mid 0 \leq w_1 \leq w_2 \leq \kappa_2, 0 \leq w_1 \leq \kappa_1\}.$$

Then non-uniqueness can be established by using this criteria.

Theorem 3.1. *Suppose that the coefficients b, σ satisfy assumptions of Theorem 2.2 or $b = b_0 + b_1, \sigma$ satisfy assumptions of Theorem 2.5. Assume that there are $a \in \mathbb{R}^d$, $\gamma_1, \gamma_2 \in (0, 1 + r_2)$ with $\gamma_1 \leq \gamma_2$, and measurable functions g_1, g_2 on $[0, +\infty)^3$, which satisfy that $g_1(\cdot, w_1, w_2)$ and $g_2(\cdot, w_1, w_2)$ are continuous and convex on $[0, +\infty)$ for any $w_1, w_2 \in [0, +\infty)$, such that for $\mu \in \mathscr{P}^{1+r_2}$*

$$\begin{aligned} &2\langle b(x + a, \mu), x\rangle + \|\sigma(x + a, \mu)\|^2_{HS} \\ &\quad \leq -g_1(|x|^{\gamma_1}, \|\mu_a\|_{\gamma_1}, \|\mu_a\|_{\gamma_2}) - g_2(|x|^{\gamma_2}, \|\mu_a\|_{\gamma_1}, \|\mu_a\|_{\gamma_2}),\ x \in \mathbb{R}^d. \end{aligned} \tag{3.1}$$

Let

$$g(w, w_1, w_2) = g_1(w, w_1, w_2) + g_2(w^{\frac{\gamma_2}{\gamma_1}}, w_1, w_2),\ w_1, w_2 \geq 0.$$

Suppose that $g(\cdot, w_1, w_2)$ is convex on $[0, +\infty)$ for any $w_1, w_2 \in [0, +\infty)$, and that there exist $\kappa_2 \geq \kappa_1 > 0$ such that for all $(w_1, w_2) \in \mathscr{D}_{\kappa_1,\kappa_2}$,

$$g(w^{\gamma_1}, w_1, w_2) > 0, \quad w \in [\kappa_1, +\infty), \tag{3.2}$$
$$g_2(w^{\gamma_2}, w_1, w_2) + \inf_{w\in[0,\kappa_1]} g_1(w^{\gamma_1}, w_1, w_2) > 0, \quad w \in [\kappa_2, +\infty). \tag{3.3}$$

Then there is $\mu \in \mathscr{P}^{\gamma_1,\gamma_2}_{a,\kappa_1,\kappa_2}$ being a stationary probability measure of (1.1).

Consequently, if there exist $a_1, a_2 \in \mathbb{R}^d$ and $\kappa_1 < \frac{|a_1-a_2|}{4}$ such that the above assumptions hold, then (1.1) has two distinct stationary probabilities $\mu_1 \in \mathscr{P}^{\gamma_1,\gamma_2}_{a_1,\kappa_1,\kappa_2}, \mu_2 \in \mathscr{P}^{\gamma_1,\gamma_2}_{a_2,\kappa_1,\kappa_2}$.

If $\gamma_1 = \gamma_2 =: \gamma$, we denote $\mathscr{P}^{\gamma}_{a,\kappa} = \mathscr{P}^{\gamma,\gamma}_{a,\kappa,\kappa}$ and have a simplified criteria as follows.

Corollary 3.2. *Suppose that the coefficients b, σ satisfy assumptions of Theorem 2.2 or $b = b_0 + b_1, \sigma$ satisfy assumptions of Theorem 2.5. Assume that there are $a \in \mathbb{R}^d$, $\gamma \in (0, 1 + r_2)$ and a measurable function g on $[0, +\infty)^2$, which satisfies that $g(\cdot, w_1)$ is continuous and convex for each $w_1 \geq 0$, such that for any $\mu \in \mathscr{P}^{1+r_2}$*

$$2\langle b(x + a, \mu), x\rangle + \|\sigma(x + a, \mu)\|^2_{HS} \leq -g(|x|^{\gamma}, \|\mu_a\|_{\gamma}). \tag{3.4}$$

If there exist $\kappa > 0$ such that

$$g(w^{\gamma}, w_1) > 0,\ w \geq \kappa, 0 \leq w_1 \leq \kappa, \tag{3.5}$$

then there is $\mu \in \mathscr{P}^{\gamma}_{a,\kappa}$ being a stationary probability measure of (1.1). Consequently, if there exist $a_1, a_2 \in \mathbb{R}^d$ and $\kappa < \frac{|a_1-a_2|}{4}$ such that the above assumptions hold, then (1.1) has two different stationary probabilities $\mu_1 \in \mathscr{P}^{\gamma}_{a_1,\kappa}, \mu_2 \in \mathscr{P}^{\gamma}_{a_2,\kappa}$.

For McKean-Vlasov SDEs, we have the following corollary.

Corollary 3.3. *Set* $b(x,\mu) = -\nabla V(x) - \nabla F * \mu(x)$ *for twice continuous differentiable functions* V *and* F*. Assume that* ∇V *has polynomial growth, and that* σ *is non-degenerate, bounded on* $\mathbb{R}^d \times \mathscr{P}$ *and satisfies (H2) and (H3). The point* $a \in \mathbb{R}^d$ *is a critical point of* V*, and there are positive constants* β_0, β_2 *and non-negative constants* β_1 *such that for every* $x, y \in \mathbb{R}^d$

$$\nabla^2 V(a+x) + \nabla^2 F(a-y+x) \geq \beta_0 - 2\beta_1|x| + 3\beta_2|x|^2. \tag{3.6}$$

Suppose that there are non-negative constants $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ *such that* (2.7) *and* (2.9) *with* $\gamma_0 = 1$ *hold, and*

$$|\nabla F(x)| \leq \alpha_1 + \alpha_2|x|,\ x \in \mathbb{R}^d. \tag{3.7}$$

Let

$$\begin{aligned}\kappa_0 &= \inf\{w > 0 \mid \beta_2 w^3 - \beta_1 w^2 + (\beta_0 - \alpha_2)w - \alpha_1 > 0\},\\ \kappa_1 &= \inf\{w > \kappa_0 \mid \beta_2 w^3 - \beta_1 w^2 + (\beta_0 - \alpha_2)w - \alpha_1 < 0\},\\ \kappa_2 &= \sup\{w > \kappa_0 \mid \beta_2 w^3 - \beta_1 w^2 + (\beta_0 - \alpha_2)w - \alpha_1 < 0\}.\end{aligned}$$

If $\beta_0 > \alpha_2$*,* $\beta_0\beta_2 > 3\beta_1^2/8$ *and there is* $\kappa \in (\kappa_0, \kappa_1) \cup (\kappa_2, +\infty)$ *such that*

$$\|\sigma\|_\infty^2 < 2\left(\beta_2\kappa^3 - \beta_1\kappa^2 + (\beta_0 - \alpha_2)\kappa - \alpha_1\right)\kappa, \tag{3.8}$$

then there is a stationary probability measure $\mu \in \mathscr{P}^1_{a,\kappa}$ *for* (1.1)*.*

Remark 3.4. Stationary distributions are found around the minimums of V, see also [26]. In following examples, the inequality (3.6) holds for the minimums of V.

Corollary 3.3 provides a sufficient condition to find the stationary probability measure around the critical point of V. According to Theorem 3.1, to prove that there are several stationary distributions, κ needs to be small (e.g. less than a quarter of the distance between two minimums). To get small κ such that (3.8) holds, a sufficient condition is that $\alpha_1 = 0$ and $\|\sigma\|_\infty$ is small, see examples below.

We present concrete examples to illustrate the non-uniqueness of stationary distributions. The first example has been investigated by many papers in the case of additive noise, see e.g. [6, 26].

Example 3.5. Let $d = 1$, $a_1, a_2 \in \mathbb{R}$ with $a_1 a_2 < 0$, $\beta > 0$ and $\alpha > 0$. Consider the following McKean-Vlasov SDE with quadratic interaction

$$\begin{aligned}\mathrm{d}X_t = &-\beta(X_t - a_1)X_t(X_t - a_2)\mathrm{d}t - \alpha\int_{\mathbb{R}}(X_t - y)\mathscr{L}_{X_t}(\mathrm{d}y)\mathrm{d}t\\ &+ \sigma(X_t, \mathscr{L}_{X_t})\mathrm{d}W_t.\end{aligned} \tag{3.9}$$

Assume that σ is positive and bounded on $\mathbb{R} \times \mathscr{P}$ and satisfies (2.2) and (H3). Then (3.9) has a stationary distribution. Furthermore, if

$$a_1^2 + a_2^2 + 2(a_1 - a_2)^2 + a_1^2 \vee a_2^2 < \frac{8\alpha}{\beta}, \tag{3.10}$$

and there is some $\kappa \in (0, (|a_1| \wedge |a_2|)/2)$ such that

$$\|\sigma\|_\infty < \kappa\sqrt{2\beta(\kappa - |a_1 - a_2|)(\kappa - |a_1| \wedge |a_2|)}, \tag{3.11}$$

there exist two distinct stationary distributions $\nu_1, \nu_2 \in \mathscr{P}^{1+r_2}$ such that

$$\nu_1(|\cdot - a_1|) \leq \kappa, \qquad \nu_2(|\cdot - a_2|) \leq \kappa. \tag{3.12}$$

Consequently, if σ is a positive constant, $a_1 = -a_2$ with (3.10) holds, and $\|\sigma\|_\infty < \kappa\sqrt{2\beta(\kappa - 2|a_1|)(\kappa - |a_1|)}$ for some $0 < \kappa < |a_1|/2$, then (3.9) has at least three stationary distributions.

For $d \geq 1$, we give the following example.

Example 3.6. Let $a_1, a_2 \in \mathbb{R}^d$, $\beta > 0$ and $\alpha > 0$. Consider the following McKean-Vlasov SDE with quadratic interaction

$$\begin{aligned} \mathrm{d}X_t = &-\frac{\beta}{2}\left((X_t - a_1)|X_t - a_2|^2 + (X_t - a_2)|X_t - a_1|^2\right)\mathrm{d}t \\ &+ \sigma(X_t, \mathscr{L}_{X_t})\mathrm{d}W_t - \alpha_1 \int_{\mathbb{R}} ((X_t - y))\mathscr{L}_{X_t}(\mathrm{d}y)\mathrm{d}t. \end{aligned}$$

Assume that σ is non-degenerate and bounded on $\mathbb{R} \times \mathscr{P}$ and satisfies (2.2) and (H3). Then (3.9) has a stationary distribution. Furthermore, if $\frac{11}{16}|a_1 - a_2|^2 < \frac{\alpha}{\beta}$, and there is some $\kappa \in (0, |a_1 - a_2|/4)$ satisfying

$$\|\sigma\|_\infty < \kappa \sqrt{2\beta\left(\kappa - |a_1 - a_2|\right)\left(\kappa - \frac{|a_1 - a_2|}{2}\right)},$$

then there exist two distinct stationary distributions $\nu_1, \nu_2 \in \mathscr{P}^{1+r_2}$ such that (3.12) hold.

The following example shows that our criteria can also deal with McKean-Vlasov SDEs with non-quadratic interaction.

Example 3.7. Let $a_1, a_2 \in \mathbb{R}^d$, and let β, α_1 and α_2 be positive constants. Consider the following McKean-Vlasov SDE with non-quadratic interaction

$$\begin{aligned} \mathrm{d}X_t = &-\frac{\beta}{2}\left((X_t - a_1)|X_t - a_2|^2 + (X_t - a_2)|X_t - a_1|^2\right)\mathrm{d}t + \sigma(X_t, \mathscr{L}_{X_t})\mathrm{d}W_t \\ &- \int_{\mathbb{R}^d} (\alpha_1|X_t - y|^2(X_t - y) + \alpha_2(X_t - y))\mathscr{L}_{X_t}(\mathrm{d}y)\mathrm{d}t. \end{aligned} \tag{3.13}$$

Assume that σ is non-degenerate and bounded on $\mathbb{R} \times \mathscr{P}$ and satisfies (2.2) and (H3). Let

$$\begin{aligned} \theta_0 &= \frac{\alpha_2}{\beta|a_1 - a_2|^2}, \qquad \theta_1 = \frac{(\beta + \alpha_1)^{\frac{3}{2}}\beta^{\frac{1}{2}}}{2(3(4 + \theta_0))^{\frac{3}{2}}\alpha_1^2}, \\ \kappa_1 &= \theta_1|a_1 - a_2|, \qquad \kappa_2 = \left(\frac{\beta\kappa_1}{4\alpha_1}\right)^{\frac{1}{3}}|a_1 - a_2|^{\frac{2}{3}}. \end{aligned}$$

Suppose that

$$\alpha_1 + \beta > 243(4 + \theta_0)\beta, \tag{3.14}$$

$$\frac{\|\sigma\|_\infty^2}{2\beta|a_1 - a_2|^4} < \left(\frac{\theta_1(4 + \theta_0)}{2} - \frac{3\alpha_1}{\beta}\right) \wedge \frac{\theta_1^2}{5}. \tag{3.15}$$

Then there are two distinct stationary distributions $\nu_1, \nu_2 \in \mathscr{P}^{1+r_2}$ such that

$$\nu_i(|\cdot - a_i|) \leq \kappa_1, \qquad \left(\nu_i(|\cdot - a_2|^3)\right)^{\frac{1}{3}} \leq \kappa_2,\ i = 1, 2.$$

We finally give an example on the non-uniqueness of stationary distributions for distribution depended SDEs with a measurable and bounded drift, which can be in non-gradient form.

Example 3.8. Consider the SDE in Example 3.6 perturbed by a bounded drift:

$$\begin{aligned} \mathrm{d}X_t = &-\beta(X_t - a_1)X_t(X_t - a_2)\mathrm{d}t - \alpha\int_{\mathbb{R}^d}(X_t - y)\mathscr{L}_{X_t}(\mathrm{d}y)\mathrm{d}t \\ &+ h(X_t, \mathscr{L}_{X_t})\mathrm{d}t + \sigma(X_t, \mathscr{L}_{X_t})\mathrm{d}W_t, \end{aligned} \tag{3.16}$$

where the constants α, β, a_1, a_2 satisfy all the conditions of Example 3.6, σ satisfies (H5) and (2.13), and h is a bounded measurable function on $\mathbb{R}^d \times \mathscr{P}$ satisfying (2.12). If there is $\kappa \in (0, |a_1 - a_2|/4)$ such that

$$\|\sigma\|_\infty^2 + \|h\|_\infty \kappa < 2\beta\kappa^2(\kappa - |a_1 - a_2|)(\kappa - |a_1 - a_2|/2), \tag{3.17}$$

then there exist two distinct stationary distributions $\nu_1, \nu_2 \in \mathscr{P}^{1+r_2}$ satisfying (3.12).

3.1 Proof of Theorem 3.1 and Corollary 3.3

Proof of Theorem 3.1. According to Theorem 2.2 or Theorem 2.5, we find a fixed point of $\mathcal{T}$ in $\mathscr{P}_M^{1+r_2}$ for every M larger than some $M_0 > 0$. Let $\mathscr{P}_{a,\kappa}^{M,\gamma} = \mathscr{P}_M^{1+r_2} \cap \mathscr{P}_{a,\kappa_1,\kappa_2}^{\gamma_1,\gamma_2}$. We are to find the stationary probability in $\mathscr{P}_{a,\kappa}^{M,\gamma}$.

By the Itô formula, we have that

$$\begin{aligned}
&|X_t^{\mu,a} - a|^2 - 2\int_0^t \langle X_s^{\mu,a} - a, \sigma(X_s^{\mu,a}, \mu)\mathrm{d}W_s\rangle \\
&\quad= \int_0^t \left(\|\sigma(X_s^{\mu,a}, \mu)\|_{HS}^2 + 2\langle X_s^{\mu,a} - a, b(X_s^{\mu,a}, \mu)\rangle\right)\mathrm{d}s \\
&\quad\le -\int_0^t g_1(|X_s^{\mu,a} - a|^{\gamma_1}, \|\mu_a\|_{\gamma_1}, \|\mu_a\|_{\gamma_2})\mathrm{d}s \\
&\qquad - \int_0^t g_2(|X_s^{\mu,a} - a|^{\gamma_2}, \|\mu_a\|_{\gamma_1}, \|\mu_a\|_{\gamma_2})\mathrm{d}s \\
&\quad= -\int_0^t g(|X_s^{\mu,a} - a|^{\gamma_1}, \|\mu_a\|_{\gamma_1}, \|\mu_a\|_{\gamma_2})\mathrm{d}s.
\end{aligned}$$

Then by the Jensen inequality, we have that

$$\frac{\mathbb{E}|X_t^{\mu,a} - a|^2}{t} \le -g\left(\frac{1}{t}\int_0^t \mathbb{E}|X_s^{\mu,a} - a|^{\gamma_1}\mathrm{d}s, \|\mu_a\|_{\gamma_1}, \|\mu_a\|_{\gamma_2}\right) \tag{3.18}$$

and

$$\begin{aligned}
\frac{\mathbb{E}|X_t^{\mu,a} - a|^2}{t} &\le -\frac{1}{t}\int_0^t \mathbb{E}g_1(|X_s^{\mu,a} - a|^{\gamma_1}, \|\mu_a\|_{\gamma_1}, \|\mu_a\|_{\gamma_2})\mathrm{d}s \\
&\quad - \frac{1}{t}\int_0^t \mathbb{E}g_2(|X_s^{\mu,a} - a|^{\gamma_2}, \|\mu_a\|_{\gamma_1}, \|\mu_a\|_{\gamma_2})\mathrm{d}s \\
&\le -g_1\left(\frac{1}{t}\int_0^t \mathbb{E}|X_s^{\mu,a} - a|^{\gamma_1}\mathrm{d}s, \|\mu_a\|_{\gamma_1}, \|\mu_a\|_{\gamma_2}\right) \\
&\quad - g_2\left(\frac{1}{t}\int_0^t \mathbb{E}|X_s^{\mu,a} - a|^{\gamma_2}\mathrm{d}s, \|\mu_a\|_{\gamma_1}, \|\mu_a\|_{\gamma_2}\right).
\end{aligned} \tag{3.19}$$

By (2.16) and $\gamma_1 \le \gamma_2 < 1 + r_2$, there is a sequence $t_n \uparrow +\infty$ such that

$$\lim_{n\to+\infty} \frac{1}{t_n}\int_0^{t_n} \mathbb{E}|X_s^{\mu,a} - a|^{\gamma_i}\mathrm{d}s = \mathcal{T}_\mu(|\cdot - a|^{\gamma_i}),\ i = 1, 2.$$

By the continuity of $g_i(\cdot, \|\mu_a\|_{\gamma_1}, \|\mu_a\|_{\gamma_2})$, $i = 1, 2$, we take $t = t_n$ in (3.19), and let $n \to +\infty$. Then

$$g((\|(\mathcal{T}_\mu)_a\|_{\gamma_1}^{\gamma_1}, \|\mu_a\|_{\gamma_1}, \|\mu_a\|_{\gamma_2}) \le 0, \tag{3.20}$$

$$g_1(\|(\mathcal{T}_\mu)_a\|_{\gamma_1}^{\gamma_1}, \|\mu_a\|_{\gamma_1}, \|\mu_a\|_{\gamma_2}) + g_2(\|(\mathcal{T}_\mu)_a\|_{\gamma_2}^{\gamma_2}, \|\mu_a\|_{\gamma_1}, \|\mu_a\|_{\gamma_2}) \le 0. \tag{3.21}$$

Combining (3.20) with (3.2), for every $\mu \in \mathscr{P}_{a,\kappa}^{M,\gamma}$, we have that $\|(\mathcal{T}_\mu)_a\|_{\gamma_1} \le \kappa_1$. This, together with (3.21) and (3.3), yields that $\|(\mathcal{T}_\mu)_a\|_{\gamma_2} \le \kappa_2$. Hence, $\mathscr{P}_{a,\kappa}^{M,\gamma}$ is an invariant set of the mapping $\mathcal{T}$. It is clear that $\mathscr{P}^{M,\gamma}_{a,\kappa}$ is also a compact, convex subset of $(\mathscr{P}^1, W)$. By the Schauder fixed point theorem, there exists $\mu \in \mathscr{P}^{M,\gamma}_{a,\kappa}$ such that $\mu = \mathcal{T}_\mu$.

Let $\mu_1 \in \mathscr{P}^{M,\gamma}_{a_1,\kappa}, \mu_2 \in \mathscr{P}^{M,\gamma}_{a_2,\kappa}$ be two fixed point of $\mathcal{T}$. Since $\kappa_1 < \frac{|a_1-a_2|}{4}$, the Chebyshev inequality yields that

$$\mu_1\left(|\cdot - a_1| \geq \frac{|a_1-a_2|}{2}\right) \leq \frac{2(\mu_1(|\cdot - a_1|^{\gamma_1}))^{\frac{1}{\gamma_1}}}{|a_1-a_2|} \leq \frac{2\kappa_1}{|a_1-a_2|} < \frac{1}{2},$$
$$\mu_2\left(|\cdot - a_2| \geq \frac{|a_1-a_2|}{2}\right) \leq \frac{2(\mu_2(|\cdot - a_2|^{\gamma_1}))^{\frac{1}{\gamma_1}}}{|a_1-a_2|} \leq \frac{2\kappa_1}{|a_1-a_2|} < \frac{1}{2}.$$

Therefore $\mu_1 \neq \mu_2$. □

Proof of Corollary 3.3. It follows from (3.6) that

$$\begin{aligned}\nabla^2 V(x) + \nabla^2 F(x-y) &\geq \beta_0 - 2\beta_1|x-a| + 3\beta_2|x-a|^2 \\ &\geq \frac{3}{2}\beta_2|x|^2 - 2\beta_1|x| - \left(3\beta_2|a|^2 + 2\beta_1|a| - \beta_0\right). \end{aligned} \tag{3.22}$$

By (3.7), for any $\epsilon > 0$,

$$|\nabla F(x)| \leq \left(\alpha_1 + \frac{2\alpha_2^{\frac{3}{2}}}{3(3\epsilon)^{\frac{1}{2}}}\right) + \epsilon|x|^3. \tag{3.23}$$

By (2.7), (3.22), (3.23) with $\epsilon < \frac{3\beta_2}{2}$ and (2.9) with $\gamma_0 = 1$, it follows from Corollary 2.4 that (1.1) has a stationary distribution in $\mathscr{P}^4$.

We verify (3.4) and (3.5) for $\kappa \in (\kappa_0, \kappa_1) \cup (\kappa_2, +\infty)$, $\gamma = 1$ and

$$g(w, w_1) = 2\left(\beta_2 w^4 - \beta_1 w^3 + \beta_0 w^2 - (\alpha_1 + \alpha_2 w_1)w\right) - \|\sigma\|_\infty, \ w, w_1 \geq 0.$$

Since a is the critical point of V, $\nabla V(a) = 0$. Then

$$\langle \nabla V(x+a), x\rangle = \langle \nabla V(x+a) - \nabla V(a), x\rangle = \int_0^1 \langle \nabla^2 V(a+\theta x)x, x\rangle \mathrm{d}\theta.$$

Thus

$$\begin{aligned}&\langle \nabla V(x+a), x\rangle + \mu(\langle \nabla F(x+a-\cdot), x\rangle) \\ &\quad = \int_0^1 \langle \nabla^2 V(a+\theta x)x, x\rangle \mathrm{d}\theta + \mu(\langle \nabla F(a-\cdot), x\rangle) \\ &\quad\quad + \mu(\langle \nabla F(x+a-\cdot) - \nabla F(a-\cdot), x\rangle) \\ &\quad = \int_0^1 \int_{\mathbb{R}^d} \langle \left(\nabla^2 V(a+\theta x) + \nabla^2 F(a-y+\theta x)\right)x, x\rangle \mu(\mathrm{d}y)\mathrm{d}\theta \\ &\quad\quad + \mu(\langle \nabla F(a-\cdot), x\rangle) \\ &\quad \geq \left(\beta_0 - \beta_1|x| + \beta_2|x|^2\right)|x|^2 - (\alpha_1 + \alpha_2\mu(|a-\cdot|))|x|.\end{aligned}$$

Then

$$2\langle b(x+a,\mu), x\rangle + \|\sigma(x+a,\mu)\|_{HS} \leq -g(|x|, \|\mu_a\|_1), \ x \in \mathbb{R}^d, \mu \in \mathscr{P}^1.$$

It is easy to see that

$$(\partial_w^2 g)(w, w_1) = 2\left(12\beta_2 w^2 - 6\beta_1 w + 2\beta_0\right).$$

Since $\beta_1^2 < \frac{8}{3}\beta_0\beta_2$, we have that $(\partial_w^2 g)(w, w_1) > 0$. Thus $g(\cdot, w_1)$ is convex on $[0, +\infty)$. Consider the set

$$A_{w_1} = \{w \geq 0 \mid g(w, w_1) \leq 0\}.$$

For any $\kappa \geq 0$, we have that

$$\begin{aligned} A_{w_1} \subset A_\kappa &= \{w \geq 0 \mid g(w, \kappa) \leq 0\} \\ &= \{w \geq 0 \mid 2\left(\beta_2 w^4 - \beta_1 w^3 + \beta_0 w^2\right) \leq 2(\alpha_1 + \alpha_2\kappa)w + \|\sigma\|_\infty^2\},\ 0 \leq w_1 \leq \kappa. \end{aligned}$$

Since the polynomial $\beta_2 w^4 - \beta_1 w^3 + \beta_0 w^2$ is convex, there is $w_0 \geq 0$ such that

$$\{w \geq 0 \mid 2\left(\beta_2 w^4 - \beta_1 w^3 + \beta_0 w^2\right) \leq 2(\alpha_1 + \alpha_2\kappa)w + \|\sigma\|_\infty^2\} = [0, w_0].$$

Since $\beta_0 > \alpha_2$, for $\kappa \in (\kappa_0, \kappa_1) \cup (\kappa_2, +\infty)$, we have that

$$\beta_2\kappa^3 - \beta_1\kappa^2 + (\beta_0 - \alpha_2)\kappa - \alpha_1 > 0.$$

By (3.8), we have that

$$\begin{aligned} g(\kappa, \kappa) &= 2\left(\beta_2\kappa^4 - \beta_1\kappa^3 + \beta_0\kappa^2 - (\alpha_1 + \alpha_2\kappa)\kappa\right) - \|\sigma\|_\infty^2 \\ &= 2\left(\beta_2\kappa^3 - \beta_1\kappa^2 + (\beta_0 - \alpha_2)\kappa - \alpha_1\right)\kappa - \|\sigma\|_\infty^2 \\ &> 0. \end{aligned}$$

Thus $w_0 < \kappa$. This yields that $\bigcup_{0 \leq w_1 \leq \kappa} A_{w_1} \subset [0, \kappa)$. Consequently, (3.5) holds. □

3.2 Proofs of examples

Proof of Example 3.5. Setting

$$b(x, \mu) = -\beta(x - a_1)x(x - a_2) - \alpha \int_{\mathbb{R}} (x - y)\mu(\mathrm{d}y),$$

it follows from Corollary 3.3 that (3.9) has stationary distributions.

Next, we use Corollary 3.2 to prove the existence of two distinct stationary distributions. Let $\gamma = 1$,

$$g(w, w_1) = 2\beta(w^4 - |2a_1 - a_2|w^3 + (a_1(a_1 - a_2) + \frac{\alpha}{\beta})w^2 - \frac{\alpha w}{\beta}w_1 - \|\sigma\|_\infty^2,\ w \geq 0.$$

Then for given $w_1 \geq 0$

$$(\partial_w^2 g)(w, w_1) = 4\beta\left(6w^2 - 3|2a_1 - a_2|w + (a_1(a_1 - a_2) + \frac{\alpha}{\beta})\right).$$

By (3.10), we have that

$$3(2a_1 - a_2)^2 - 8a_1(a_1 - a_2) \leq a_1^2 + a_2^2 + 2(a_1 - a_2)^2 + a_1^2 \vee a_2^2 < \frac{8\alpha}{\beta}.$$

This implies that $(\partial_w^2 g)(w, w_1) > 0$. Thus $g(\cdot, w_1)$ is convex. Moreover,

$$\begin{aligned} &2\langle b(x + a_1, \mu), x\rangle + \|\sigma(x, \mu)\|_{HS}^2 \\ &\quad= -2\beta x^2(x + a_1)(x + a_1 - a_2) - 2\alpha x \int_{\mathbb{R}} (x + a_1 - y)\,\mu(\mathrm{d}y) + \|\sigma\|_\infty^2 \\ &\quad= -2\beta\left(x^4 + (2a_1 - a_2)x^3 + (a_1(a_1 - a_2) + \frac{\alpha}{\beta})x^2 - \frac{\alpha x}{\beta}\int_{\mathbb{R}}(a_1 - y)\mu(\mathrm{d}y)\right) + \|\sigma\|_\infty^2 \\ &\quad\leq -g(|x|, \|\mu_{a_1}\|_1). \end{aligned} \tag{3.24}$$

Let

$$p(w) = 2\beta(w^4 - |2a_1 - a_2|w^3 + (a_1(a_1 - a_2) + \frac{\alpha}{\beta})w^2). \tag{3.25}$$

Then p is is convex on $[0, +\infty)$. Thus for $\kappa > 0$, there is $w_0 > 0$ such that

$$\left\{w \geq 0 \mid p(w) \leq 2\alpha\kappa w + \|\sigma\|_\infty^2\right\} = [0, w_0].$$

Hence, for any $M > 0$, $\mu \in \mathscr{P}_{a_1,\kappa}^{M,1}$ and $0 \leq w_1 \leq \kappa$, we have that

$$\begin{aligned} A_{w_1} &:= \{w \geq 0 \mid g(w, w_1) \leq 0\} \\ &\subset \left\{w \geq 0 \mid 2\beta(w^4 - |2a_1 - a_2|w^3 + (a_1(a_1 - a_2) + \frac{\alpha}{\beta})w^2 - \frac{\alpha\kappa w}{\beta}) - \|\sigma\|_\infty^2 \leq 0\right\} \\ &= \left\{w \geq 0 \mid p(w) \leq 2\alpha\kappa w + \|\sigma\|_\infty^2\right\} \\ &= [0, w_0]. \end{aligned} \tag{3.26}$$

Since $a_1 a_2 < 0$, we have that $a_1(a_1 - a_2) > 0$ and $|a_1 - a_2| > |a_1|$. Then for $0 < \kappa < |a_1| \wedge |a_2|$, we have that

$$\begin{aligned} \kappa^2 - |2a_1 - a_2|\kappa + a_1(a_1 - a_2) &= (\kappa - |a_1|)(\kappa - |a_1 - a_2|) \\ &\geq (\kappa - |a_1| \wedge |a_2|)(\kappa - |a_1 - a_2|). \end{aligned}$$

Note that

$$p(\kappa) - 2\alpha\kappa^2 - \|\sigma\|_\infty^2 = 2\beta(\kappa^2 - |2a_1 - a_2|\kappa + a_1(a_1 - a_2))\kappa^2 - \|\sigma\|_\infty^2.$$

Thus, by (3.11),

$$p(\kappa) - 2\alpha\kappa^2 - \|\sigma\|_\infty^2 > 0.$$

This, together with (3.26), yields that $w_0 < \kappa$. Taking into account (3.26), we get (3.5). Hence, (3.9) has a stationary probability measure $\nu_1 \in \mathscr{P}_{a_1,\kappa}^{M,1}$. It is clear that we can replace a_1 by a_2. Thus (3.9) has a stationary probability measure $\nu_2 \in \mathscr{P}_{M,\kappa}^{a_1,1}$. Finally, since

$$\kappa < \frac{|a_1| \wedge |a_2|}{2} \leq \frac{1}{2}\left(\frac{|a_1 - a_2|}{2}\right),$$

the two probability measures ν_1 and ν_2 are distinct by Theorem 3.1.

If $a_1 = -a_2$ and σ is a constant, it has been proved that (3.9) has a symmetric stationary probability measure (see e.g. [14]):

$$\mu_0(\mathrm{d}x) = c_0 \exp\left\{-\frac{2}{\sigma^2}\left(\beta\left(\frac{x^4}{4} - \frac{a_1^2 x^2}{2}\right) - \frac{\alpha x^2}{2}\right)\right\}\mathrm{d}x,$$

where c_0 is the normalization constant so that μ_0 is a probability measure. Assuming $a_1 > 0$, it follows from (3.12) and $\kappa < a_1/2$ that

$$\nu_1((-\infty, 0]) \leq \nu_1(|\cdot - a_1| \geq a_1) \leq \frac{\nu_1(|\cdot - a_1|)}{a_1} \leq \frac{\kappa}{a_1} < \frac{1}{2},$$

and, similarly, $\nu_2([0, +\infty)) < 1/2$. Thus, ν_1 and ν_2 are not symmetric measures. Hence, (3.9) has three stationary distributions. □

Proof of Example 3.6. Set

$$g(w, w_1) = 2\beta\left(w^4 - \frac{3}{2}|a_1 - a_2|w^3 + \left(\frac{|a_1 - a_2|^2}{2} + \frac{\alpha}{\beta}\right)w^2 - \frac{\alpha w}{\beta}w_1\right) - \|\sigma\|_\infty^2,$$

$$p(w) = 2\beta \left(w^4 - \frac{3}{2}|a_1 - a_2|w^3 + \left(\frac{|a_1 - a_2|^2}{2} + \frac{\alpha}{\beta} \right) w^2 \right).$$

One can follow line by line the proof of Example 3.5 to get the conclusion of this example, and we omit the details. □

Proof of Example 3.7. According to (3.13), we have that

$$\begin{aligned} b(x,\mu) = &-\frac{\beta}{2}\left\{(x-a_1)|x-a_2|^2 + (x-a_2)|x-a_1|^2\right\} \\ &- \int_{\mathbb{R}^d} \left(\alpha_1|x-y|^2(x-y) + \alpha_2(x-y)\right)\mu(\mathrm{d}y). \end{aligned}$$

It is a routine to check that all the conditions of Corollary 2.4, and we omit the details. The rest is prove (3.1), (3.2) and (3.3) for $\gamma_1 = 1, \gamma_2 = 3$ and

$$\begin{aligned} g_1(w,w_1,w_2) &= 2\beta\Big[\left(\frac{|a_1-a_2|^2}{2} + \frac{\alpha_1}{\beta}w_1^2 + \frac{\alpha_2}{\beta} \right) w^2 - \left(\frac{\alpha_1}{\beta}w_2^3 + \frac{\alpha_2}{\beta}w_1 \right) w \Big] - \|\sigma\|_\infty^2, \\ g_2(w,w_1) &= 2\beta\left[\left(1 + \frac{\alpha_1}{\beta}\right) w^{\frac{4}{3}} - 3\left(|a_1-a_2| + \frac{\alpha_1}{\beta}w_1 \right) w \right], \\ g(w,w_1,w_2) &= g_1(w,w_1,w_2) + g_2(w^3,w_1),\ w \geq 0. \end{aligned}$$

It is clear that

$$\begin{aligned} &-\frac{\beta}{2}\left(|x|^2|x+a_1-a_2|^2 + \langle x+a_1-a_2, x\rangle|x|^2\right) \\ &\quad= -\beta\left(|x|^4 + \frac{3}{2}\langle a_1-a_2, x\rangle|x|^2 + \frac{|a_1-a_2|^2}{2}|x|^2 \right) \\ &\quad\leq -\beta\left(|x|^4 - \frac{3}{2}|a_1-a_2||x|^3 + \frac{|a_1-a_2|^2}{2}|x|^2 \right), \end{aligned}$$

and

$$\begin{aligned} &-\alpha_1 \int_{\mathbb{R}^d} |x+a_1-y|^2\langle x+a_1-y, x\rangle\mu(\mathrm{d}y) - \alpha_2\int_{\mathbb{R}^d}\langle x+a_1-y, x\rangle\mu(\mathrm{d}y) \\ &= -\alpha_1\int_{\mathbb{R}^d}\left(|x|^2 + 2\langle x, a_1-y\rangle + |a_1-y|^2\right)\left(|x|^2 + \langle a_1-y, x\rangle\right)\mu(\mathrm{d}y) \\ &\quad - \alpha_2\int_{\mathbb{R}^d}\left(|x|^2 + \langle a_1-y, x\rangle\right)\mu(\mathrm{d}y) \\ &\leq -\alpha_1\left(|x|^4 + 3|x|^2\int_{\mathbb{R}^d}\langle x, a_1-y\rangle\mu(\mathrm{d}y) + |x|^2\int_{\mathbb{R}^d}|a_1-y|^2\mu(\mathrm{d}y) \right) \\ &\quad - \alpha_1\left(2\int_{\mathbb{R}^d}\langle a_1-y, x\rangle^2\mu(\mathrm{d}y) + \int_{\mathbb{R}^d}|a_1-y|^2\langle a_1-y, x\rangle\mu(\mathrm{d}y) \right) \\ &\quad - \alpha_2|x|^2 - \alpha_2\int_{\mathbb{R}^d}\langle a_1-y, x\rangle\mu(\mathrm{d}y) \\ &\leq -\alpha_1\left(|x|^4 - 3\mu(|\cdot - a_1|)|x|^3 + \mu(|\cdot - a_1|^2)|x|^2 - \mu(|\cdot - a_1|^3)|x|\right) \\ &\quad - \alpha_2|x|^2 + \alpha_2\mu(|\cdot - a_1|)|x|. \end{aligned}$$

Thus

$$\begin{aligned} 2\langle b(x+a_1,\mu), x\rangle = &-\beta\left(|x|^2|x+a_1-a_2|^2 + \langle x+a_1-a_2, x\rangle|x|^2\right) \\ &- 2\alpha_1\int_{\mathbb{R}^d}|x+a_1-y|^2\langle x+a_1-y, x\rangle\mu(\mathrm{d}y) \end{aligned}$$

$$
\begin{aligned}
&- 2\alpha_2 \int_{\mathbb{R}^d} \langle x + a_1 - y, x\rangle \mu(\mathrm{d}y) \\
&\le -2\beta\Big[\left(1 + \frac{\alpha_1}{\beta}\right) |x|^4 - 3\left(|a_1 - a_2| + \frac{\alpha_1}{\beta}\mu(|\cdot - a_1|)\right)|x|^3 \\
&\quad + \left(\frac{|a_1 - a_2|^2}{2} + \frac{\alpha_1}{\beta}(\mu(|\cdot - a_1|))^2 + \frac{\alpha_2}{\beta}\right)|x|^2 \\
&\quad - \left(\frac{\alpha_1}{\beta}\mu(|\cdot - a_1|^3) + \frac{\alpha_2}{\beta}\mu(|\cdot - a_1|)\right)|x|\Big].
\end{aligned}
$$

Then

$$
2\langle b(x + a_1, \mu), x\rangle + \|\sigma(x + a_1, \mu)\|_{HS}^2 \le -g(|x|, \|\mu_{a_1}\|_1, \|\mu_{a_1}\|_3), \ x \in \mathbb{R}^d.
$$

For μ with $\|\mu_{a_1}\|_1 < \kappa_1$ and $\|\mu_{a_1}\|_3 < \kappa_2$. Since $\frac{\alpha_1}{\beta} > 1$, we have that

$$
\theta_1 = \frac{(1 + \frac{\alpha_1}{\beta})^{\frac{3}{2}} \left(\frac{\beta}{\alpha_1}\right)^2}{2(3(4 + \theta_0))^{\frac{3}{2}}} \le \frac{(\frac{2\alpha_1}{\beta})^{\frac{3}{2}} \left(\frac{\beta}{\alpha_1}\right)^2}{2(3(4 + \theta_0))^{\frac{3}{2}}} < \frac{1}{2}\sqrt{\frac{\beta}{\alpha_1}}.
$$

Let $\theta_2 = \frac{\kappa_2}{|a_1 - a_2|}$. Then

$$
\theta_1 < \left(\frac{\beta\theta_1}{4\alpha_1}\right)^{\frac{1}{3}} = \theta_2.
$$

As a consequence, we have that $\kappa_1 < \kappa_2$. It is clear that

$$
\begin{aligned}
(\partial_w^2 g)(w, w_1, w_2) = 2\beta\Big[&12\left(1 + \frac{\alpha_1}{\beta}\right)w^2 - 18\left(|a_1 - a_2| + \frac{\alpha_1}{\beta}w_1\right)w \\
&+ 2\left(\frac{|a_1 - a_2|^2}{2} + \frac{\alpha_1}{\beta}w_1^2 + \frac{\alpha_2}{\beta}\right)\Big].
\end{aligned}
$$

By (3.14), we have that

$$
\theta_1 < \left(\frac{2}{\sqrt{27}}\sqrt{(1 + \frac{\alpha_1}{\beta})(1 + 2\theta_0)} - 1\right)\frac{\beta}{\alpha_1}.
$$

Then

$$
\left(1 + \frac{\alpha_1\theta_1}{\beta}\right)^2 < \frac{4}{27}\left(1 + \frac{\alpha_1}{\beta}\right)(1 + 2\theta_0).
$$

This implies that

$$
\left(18\left(|a_1 - a_2| + \frac{\alpha_1}{\beta}\mu(|\cdot - a_1|)\right)\right)^2 < 48\left(1 + \frac{\alpha_1}{\beta}\right)\left(|a_1 - a_2|^2 + \frac{2\alpha_2}{\beta}\right).
$$

Consequently, $g(\cdot, w_1, w_2)$ is convex. Thus there is $w_0 > 0$ such that

$$
\begin{aligned}
&\{w \ge 0 \mid g(w, w_1, w_2) \le 0\} \\
&\quad \subset \Big\{ w \ge 0 \ \Big|\ (1 + \frac{\alpha_1}{\beta})w^4 - 3(|a_1 - a_2| + \frac{\alpha_1}{\beta}\kappa_1)w^3 \\
&\qquad + \left(\frac{|a_1 - a_2|^2}{2} + \frac{\alpha_2}{\beta}\right)w^2 - \left(\frac{\alpha_1}{\beta}\kappa_2^3 + \frac{\alpha_2}{\beta}\kappa_1\right)w \le \frac{\|\sigma\|_\infty^2}{2\beta}\Big\} \\
&\quad = [0, w_0].
\end{aligned}
$$

Since (3.14), we have that

$$\frac{\theta_1}{2} - 3\theta_1^2 + (1 - \frac{2\alpha_1}{\beta})\theta_1^3 > \frac{9\theta_1}{20} > \frac{\theta_1}{4}.$$

This, together with $\frac{1}{4}\theta_1 = \frac{\alpha_1}{\beta}\theta_2^3$ and $\kappa_1 = \theta_1|a_1 - a_2|$, implies that

$$\frac{|a_1 - a_2|^2}{2}\kappa_1 - 3|a_1 - a_2|\kappa_1^2 + (1 - \frac{2\alpha_1}{\beta})\kappa_1^3 > \frac{\alpha_1}{\beta}\kappa_2^3.$$

Moreover, it follows from (3.15) that

$$\begin{aligned}
&(1 + \frac{\alpha_1}{\beta})\kappa_1^4 - 3(|a_1 - a_2| + \frac{\alpha_1}{\beta}\kappa_1)\kappa_1^3 + \left(\frac{|a_1 - a_2|^2}{2} + \frac{\alpha_2}{\beta}\right)\kappa_1^2 - \left(\frac{\alpha_1}{\beta}\kappa_2^3 + \frac{\alpha_2}{\beta}\kappa_1\right)\kappa_1 \\
&\geq (1 - \frac{2\alpha_1}{\beta})\kappa_1^4 - 3|a_1 - a_2|\kappa_1^3 + \frac{|a_1 - a_2|^2}{2}\kappa_1^2 - \frac{\alpha_1}{\beta}\kappa_2^3\kappa_1 \\
&= \left((1 - \frac{2\alpha_1}{\beta})\theta_1^3 - 3\theta_1^2 + \frac{\theta_1}{4}\right)\theta_1|a_1 - a_2|^4 \\
&> \frac{\theta_1^2}{5}|a_1 - a_2|^4 > \frac{\|\sigma\|_\infty^2}{2\beta}.
\end{aligned}$$

Hence, $w_0 \leq \kappa_1$. Hence, (3.2) holds.

For $(w_1, w_2) \in \mathscr{D}_{\kappa_1,\kappa_2}$, we have that

$$\begin{aligned}
&g_2(\kappa_2^3, w_1) + \inf_{\kappa_1 \geq w \geq 0} g_1(w, w_1, w_2) \\
&\geq 2\beta\left(\left(1 + \frac{\alpha_1}{\beta}\right)\kappa_2^4 - 3\left(|a_1 - a_2| + \frac{\alpha_1}{\beta}\kappa_1\right)\kappa_2^3\right) - \frac{\beta\left(\frac{\alpha_1}{\beta}\kappa_2^3 + \frac{\alpha_2}{\beta}\kappa_1\right)^2}{|a_1 - a_2|^2 + \frac{2\alpha_2}{\beta}} - \|\sigma\|_\infty^2 \\
&= 2\beta\left(\left(1 + \frac{\alpha_1}{\beta}\right)\theta_2^4 - 3\left(1 + \frac{\alpha_1}{\beta}\theta_1\right)\theta_2^3 - \frac{\left(\frac{\alpha_1}{\beta}\theta_2^3 + \theta_0\theta_1\right)^2}{2 + 4\theta_0}\right)|a_1 - a_2|^4 - \|\sigma\|_\infty^2.
\end{aligned}$$

Taking into account that $\frac{\beta\theta_1}{4\alpha_1} = \theta_2^3$, we have by (3.15) that

$$\begin{aligned}
&\left(1 + \frac{\alpha_1}{\beta}\right)\theta_2^4 - 3\left(1 + \frac{\alpha_1}{\beta}\theta_1\right)\theta_2^3 - \frac{(\frac{\alpha_1}{\beta}\theta_2^3 + \theta_0\theta_1)^2}{2 + 4\theta_0} \\
&\geq \left[\left(1 + \frac{\alpha_1}{\beta}\right)\left(\frac{\beta}{4\alpha_1}\right)^{\frac{4}{3}}\theta_1^{\frac{1}{3}} - \frac{3\beta}{4\alpha_1} - \left(\frac{3}{4} + \frac{(\frac{1}{4} + \theta_0)^2}{2 + 4\theta_0}\right)\theta_1\right]\theta_1 \\
&> \left[\left(1 + \frac{\alpha_1}{\beta}\right)\left(\frac{\beta}{4\alpha_1}\right)^{\frac{4}{3}}\theta_1^{\frac{1}{3}} - \frac{3\beta}{4\alpha_1} - \left(1 + \frac{\theta_0}{4}\right)\theta_1\right]\theta_1 \\
&= \frac{(1 + \frac{\alpha_1}{\beta})^{\frac{3}{2}}\left(\frac{\beta}{\alpha_1}\right)^2}{24\sqrt{3}(1 + \frac{\theta_0}{4})^{\frac{1}{2}}} - \frac{3\beta}{4\alpha_1} = \frac{(4 + \theta_0)\theta_1}{2} - \frac{3\alpha_1}{4\beta} \\
&> \frac{\|\sigma\|_\infty^2}{2\beta|a_1 - a_2|^4}.
\end{aligned}$$

Thus

$$g_2(\kappa_2^3, w_1) + g_1(w, w_1, w_2) > 0,\ (w_1, w_2) \in \mathscr{D}_{\kappa_1,\kappa_2}.$$

Note that $g_2(\cdot, w_1)$ decreases first and then increase. Thus,

$$\bigcup_{(w_1,w_2)\in\mathscr{D}_{\kappa_1,\kappa_2}}\left\{0 \leq w \;\middle|\; g_2(w^3, w_1) + \inf_{0\leq w\leq\kappa_1} g_1(w, w_1, w_2) \leq 0\right\} \subset [0, \kappa_2].$$

Hence, (3.3) holds. Therefore, the proof is complete. □

Proof of Example 3.8. With $g(w,w_1)$ in the proof of Example 3.6 replaced by $g(w,w_1) - \|h\|_\infty w$, as in the proof of Example 3.5, one has that $\bigcup_{0<w\le\kappa} A_w \subset [0,w_0]$ for some $w_0 \ge 0$. Let $p(w)$ be the polynomial defined in the proof of Example 3.6. Then it follows from (3.17) that

$$\begin{aligned} &p(\kappa) - 2\alpha\kappa^2 - \|h\|_\infty\kappa - \|\sigma\|_\infty^2 \\ &\quad \ge 2\beta\kappa^2(\kappa - |a_1 - a_2|/2)(\kappa - |a_1 - a_2|) - \|h\|_\infty\kappa - \|\sigma\|_\infty^2 \\ &\quad > 0. \end{aligned}$$

Hence $w_0 \le \kappa$. Replacing a_1 by a_2 we have the same consequence. Therefore, there exist two distinct stationary distributions $\nu_1 \in \mathscr{P}_{M,\kappa}^{a_1,1}$, $\nu_2 \in \mathscr{P}_{M,\kappa}^{a_2,1}$. □

References

[1] Carmona, R. and Delarue, F.: Probabilistic Theory of Mean Field Games with Applications I. *Springer*, Switzerland, 2018. MR3753660

[2] Carrillo, J. A., Gvalani, R. S., Pavliotis, G.A. and Schlichting, A.: Long-time behaviour and phase transitions for the Mckean-Vlasov equation on the torus, *Arch. Rational Mech. Anal.*, **235** (2020), 635–690. MR4062483

[3] Chen, M.-F.: From Markov Chains to Non-Equilibrium Particle Systems, 2nd ed. *World Scientific*, Singapore, 2004. MR2091955

[4] Carrillo, J. A., McCann, R. J. and Villani, C.: Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates, *Rev. Mat. Iberoam.*, **19** (2003), 971–1018. MR2053570

[5] Chayes, L. and Panferov, V.: The McKean-Vlasov equation in finite volume, *J. Stat. Phys.*, **138** (2010), 351–380. MR2594901

[6] Dawson, D. A.: Critical dynamics and fluctuations for a mean-field model of cooperative behavior, *J. Stat. Phys.*, **31** (1983), 29–85. MR0711469

[7] Deimling, K.: Nonlinear Functional Analysis. *Springer-Verlag*, Berlin Heidelberg, 1985. MR0787404

[8] Delgadino, M. G., Gvalani, R. S. and Pavliotis, G. A.: On the diffusive-mean field limit for weakly interacting diffusions exhibiting phase transitions, *Arch. Ration. Mech. Anal.*, **241** (2021), 91–148. MR4271956

[9] Delgadino, M. G., Yan, X. and Yao, Y.: Uniqueness and nonuniqueness of steady states of aggregation-diffsion equations, *Comm. Pure Appl. Math.*, **75** (2022), 3–59. MR4373166

[10] Duong, M. H. and Tugaut, J.: Stationary solutions of the Vlasov-Fokker-Planck equation: Existence, characterization and phase-transition, *Applied Mathematics Letters*, **52** (2016), 38–45. MR3416384

[11] Eberle, A., Guillin, A. and Zimmer, R.: Quantitative Harris-type theorems for diffusions and McKean-Vlasov processes, *Trans. Amer. Math. Soc.*, **371** (2019), 7135–7173. MR3939573

[12] Feng, S. and Zheng, X. G.: Solutions of a class of non-linear Master equations, *Stoch. Proc. Appl.*, **43** (1992), 65–84. MR1190907

[13] Goldys, B. and Maslowski, B.: Exponential ergodicity for stochastic reaction-diffusion equations. In: Stochastic Partial Differential Equations and Applications, XVII, Lect. Notes Pure Appl. Math., vol. 245, pp. 115–131. *Chapman Hall/CRC*, Boca Raton, FL, 2006. MR2227225

[14] Herrmann, S. and Tugaut, J.: Non-uniqueness of stationary measures for self-stabilizing processes, *Stoch. Proc. Appl.*, **120** (2010), 1215–1246. MR2639745

[15] Herrmann, S. and Tugaut, J.: Stationary measures for self-stabilizing processes: Asymptotic analysis in the small noise limit, *Electron. J. Probab.*, **15** (2010), 2087–2116. MR2745727

[16] Huang, X. and Wang, F.-Y.: Distribution dependent SDEs with singular coefficients, *Stoch. Proc. Appl.*, **129** (2019), 4747–4770. MR4013879

[17] Huang, X., Ren, P. and Wang, F.-Y.: Distribution dependent stochastic differential equations, *Front. Math. China*, **16** (2021), 257–301. MR4254653

[18] Khasminskii, R.: Stochastic Stability of Differential Equations, 2nd ed., Stochastic Modelling and Applied Probability, vol. 66. *Springer*, Heidelberg, 2012. MR2894052

[19] Liu, W. and Röckner, M.: Stochastic Partial Differential Equations: An Introduction. *Springer*, Switzerland, 2015. MR3410409

[20] McKean, H. P. Jr.: Propagation of chaos for a class of non-linear parabolic equations. In: Stochastic Differential Equations, Lecture Series in Differential Equations, Session 7, Catholic Univ., 41–57. *Air Force Office Sci. Res.*, Arlington, VA, 1967. MR0233437

[21] Röckner, M., Zhang, X.: Well-posedness of distribution dependent SDEs with singular drifts, *Bernoulli*, **27** (2021), 1131–1158. MR4255229

[22] Sznitman, A.-S.: Topics in propagation of chaos. In: École d'Été de Probabilités de Saint-Flour XIX-1989, Lecture Notes in Math., vol. 1464, pp. 165–251. *Springer*, Berlin, 1991. MR1108185

[23] Tamura, Y.: On asymptotic behaviors of the solution of a nonlinear diffusion equation, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, **31** (1984), 195–221. MR0743525

[24] Tugaut, J.: Convergence to the equilibria for self-stabilizing processes in double well landscape, *Ann. Probab.*, **41** (2010), 1427–1460. MR3098681

[25] Tugaut, J.: Self-stabilizing processes in multi-wells landscape in $\mathbb{R}^d$-convergence, *Stoch. Proc. Appl.*, **123** (2013), 1780–1801. MR3027901

[26] Tugaut, J.: Phase transitions of McKean-Vlasov processes in double-wells landscape, *Stochastics*, **86** (2014), 257–284. MR3180036

[27] Tugaut, J.: Self-stabilizing processes in multi-wells landscape in $\mathbb{R}^d$-invariant probabilities, *J. Theoret. Probab.*, **27** (2014), 57–79. MR3174216

[28] Wang, F.-Y.: Distribution dependent SDEs for Landau type equations, *Stoch. Proc. Appl.*, **128** (2018), 595–621. MR3739509

[29] Wang, F.-Y.: Exponential ergodicity for fully non-dissipative McKean-Vlasov SDEs, arXiv:2101.12562

[30] Xia, P., Xie, L., Zhang, X. and Zhao, G.: $L^q(L^p)$-theory of stochastic differential equations, *Stoch. Proc. Appl.*, **130** (2020), 5188–5211. MR4108486

[31] Xie, L. and Zhang, X.: Ergodicity of stochastic differential equations with jumps and singular coefficients, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, **56** (2020), 175–229. MR4058986

[32] Zhang, S.-Q. and Yuan, C.: A Zvonkin's transformation for stochastic differential equations with singular drift and related applications, *Journal of Differential Equations*, **297** (2021), 277–319. MR4280266

[33] Zhang, X.: Weak solutions of McKean-Vlasov SDEs with supercritical drifts, arXiv:2010.15330

[34] Zvonkin, A. K.: A transformation of the phase space of a diffusion process that will remove the drift, *Mat. Sb.*, **93** (1974), 129–149. MR0336813

Acknowledgments. The author thanks referees for their useful comments.