

July 2026

“Well-Posedness of McKean-Vlasov SDEs
with Density-Dependent Drift”

Anh-Dung Le and Stéphane Villeneuve

WELL-POSEDNESS OF MCKEAN-VLASOV SDES WITH DENSITY-DEPENDENT DRIFT

ANH-DUNG LE AND STÉPHANE VILLENEUVE

ABSTRACT. In this paper, we study well-posedness of McKean-Vlasov stochastic differential equations (SDE) whose drift depends pointwisely on marginal density and satisfies a local integrability condition in time-space variables. The drift and noise coefficients are assumed to be Lipschitz continuous in distribution variable with respect to Wasserstein metric W_p . Our approach is by approximation with mollifiers. We prove strong existence of a solution. Weak and strong uniqueness are obtained when $p = 1$, the drift coefficient is bounded, and the diffusion coefficient is distribution free.

Keywords: McKean-Vlasov SDEs, density-dependent SDEs, local integrability

2020 MSC: 60B10, 60H10

CONTENTS

1. Introduction	1
2. Main results	3
3. Preliminaries	6
4. Proof of Theorem 2.2	13
5. Proof of Theorem 2.3	21
6. Acknowledgment	23
References	24
7. Appendix	27

1. INTRODUCTION

The study of distribution-dependent SDEs started with McKean's seminal work [McK66] about Vlasov equation of plasma which had been proposed in [Kac56]. Classical results about the solvability of McKean-Vlasov equations include [Fun84; Szn84; Sch87]. Since then, the literature on the well-posedness of McKean-Vlasov SDEs has been extended significantly. For recent results, we refer to [MV20; Cha20; HW21; RZ21; HW22; CF22], the survey [HRW21], and references therein. For applications on mean-field games, see the two-volume monograph [CD18].

In this paper, we consider the density-dependent McKean-Vlasov SDEs studied in [BR20]. More precisely, let $(p, d) \in [1, \infty) \times \mathbb{N}$ and $\mathcal{P}_p(\mathbb{R}^d)$ be the space of Borel probability measures on $\mathbb{R}^d$ with finite p -th moment. We endow $\mathcal{P}_p(\mathbb{R}^d)$ with the Wasserstein metric W_p . Let $T \in (0, \infty)$. We consider measurable functions

$$b : [0, T] \times \mathbb{R}^d \times \mathbb{R}_+ \times \mathcal{P}_p(\mathbb{R}^d) \rightarrow \mathbb{R}^d,$$

Date: July 31, 2026.

This work was supported by ANR MaSDOL (No. ANR-19-CE23-0017); Air Force Office of Scientific Research, Air Force Material Command, USAF (No. FA9550-19-7026); and ANR Chess (No. ANR-17-EURE-0010).

$$\sigma : [0, T] \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d) \rightarrow \mathbb{R}^d \otimes \mathbb{R}^d.$$

Let (B_t) be a d -dimensional Brownian motion (d -BM) and $(\mathcal{F}_t)$ an admissible filtration (AF) on a probability space (PS) $(\Omega, \mathcal{A}, \mathbb{P})$. For a random variable Y taking values in $\mathbb{R}^d$, we denote by $\mathcal{L}(Y)$ its probability distribution. Our object of study is the SDE

$$\begin{cases} dX_t = b(t, X_t, \ell_t(X_t), \mu_t) dt + \sigma(t, X_t, \mu_t) dB_t, \\ \nu = \mathcal{L}(X_0), \quad \mu_t = \mathcal{L}(X_t), \quad \mu_t(dx) = \ell_t(x) dx. \end{cases} \quad (1.1)$$

In particular, $\ell_t : \mathbb{R}^d \rightarrow \mathbb{R}_+$ is the density of X_t . In the paper, we always mean by “density” the probability density function.

First, we give a brief survey of the related literature. In [HRZ23; IR23; Iss+24], σ is constant and b belongs to a Besov space. On the other hand, [BRR10; BRR11; BR12; BCR13] studied generalized porous media equations where $b = 0$ and σ is density-dependent but time-homogeneous. In [JM98; JR13; BJ19; BR18; BR20; BR21a; BR21b; BR23c; BR23b; Reh23; Gru23], b and σ are density-dependent but time-homogeneous. More generally, [Gru24; BR23a] allow b and σ to be depend on time. However, b and σ are distribution-independent in those previously mentioned papers. For a recent survey on existence result of density-dependent SDEs, we refer to [IOR19].

Our paper is closely related to [HRZ24; Wan23] where b is both distribution-dependent and density-dependent. In particular, [HRZ24] used the total variation metric on $\mathcal{P}(\mathbb{R}^d)$ while [Wan23] used supremum norm on the space of bounded densities. The current paper contributes in three aspects:

- (1) We use a mollifying argument whereas [HRZ24; Wan23] employed a Picard-iteration argument. We argue that our approach is more flexible because it can be extended to those situations where reasonable Hölder estimates are available. An interesting case for future study is when b is allowed to grow linearly in space as in [MPZ21].
- (2) For existence result, the conditions on σ in [Wan23; HRZ24] are more restrictive than ours. First, [HRZ24] assumed that $\sigma(t, x, \mu) = \sigma(t, x)$. Second, [Wan23] assumed that σ is Lipschitz in space, that $\nabla \sigma$ is Hölder continuous, and that $\|\sigma(t, \cdot, \mu) - \sigma(t, \cdot, \nu)\|_{C_b^\alpha} \lesssim \|\ell_\mu - \ell_\nu\|_\infty$. Here ℓ_μ, ℓ_ν are the densities of μ, ν respectively.
- (3) To be more aligned with existing literature of McKean-Vlasov SDEs, we use Wasserstein metric for assumptions of continuity. This makes estimating supremum norm between marginal densities (as in [HRZ24; Wan23]) (of two weak solutions) not applicable in our case. However, using Wasserstein metric to estimate the difference between marginal distributions (of two weak solutions) is also not applicable due to the presence of pointwise density $\ell_t(X_t)$ in b . To overcome these difficulties, we will estimate weighted total variation distance between marginal densities.

We recall notions of a solution:

Definition 1.1. (1) A *strong solution* to (1.1) is a continuous $\mathbb{R}^d$ -valued process (X_t) on $(\Omega, \mathcal{A}, \mathbb{P})$ such that for $t \in [0, T]$: X_t is $\mathcal{F}_t$ -adapted, X_t has a distribution $\mu_t \in \mathcal{P}_p(\mathbb{R}^d)$, X_t admits a density ℓ_t , and

$$\begin{aligned} X_t &= X_0 + \int_0^t b(s, X_s, \ell_s(X_s), \mu_s) ds + \int_0^t \sigma(s, X_s, \mu_s) dB_s \quad \mathbb{P}\text{-a.s.}, \\ \int_0^t \mathbb{E}[|b(s, X_s, \ell_s(X_s), \mu_s)| + |\sigma(s, X_s, \mu_s)|^2] ds &< \infty. \end{aligned}$$

- (2) A *weak solution* to (1.1) is a continuous $\mathbb{R}^d$ -valued process (X_t) on some PS $(\Omega, \mathcal{A}, \mathbb{P})$ on which there exist some d -BM (B_t) and some AF $(\mathcal{F}_t)$ such that the conditions in (1) are satisfied.

- (3) SDE (1.1) has *strong uniqueness* if, whenever the PS, the AF and the d -BM are fixed, two strong solutions (X_t) and (X'_t) such that $X_0 = X'_0$ coincide $\mathbb{P}$ -a.s. on the path space $C([0, T]; \mathbb{R}^d)$. SDE (1.1) has *weak uniqueness* if two weak solutions with the same initial distribution induce the same distribution on $C([0, T]; \mathbb{R}^d)$.
- (4) SDE (1.1) is *strongly well-posed* if it has strong solution and strong uniqueness. SDE (1.1) is *weakly well-posed* if it has weak solution and weak uniqueness. SDE (1.1) is *well-posed* if it is both strongly and weakly well-posed.

In Section 2, we state our main results about existence and uniqueness of a solution to (1.1). In Section 3, we remind some facts about optimal transport. Also, we recall estimates of marginal density and establish those of marginal distribution for classical SDEs. We prove our theorems in Section 4 and Section 5 respectively.

Throughout this paper, we use the following conventions:

- (1) We denote by $\mathbb{R}^d \otimes \mathbb{R}^d$ the space of d -dimensional square matrices with real entries. For $x, y \in \mathbb{R}^d \otimes \mathbb{R}^d$, let xy be their matrix product, $\langle x, y \rangle$ their Frobenius inner product, and $|x|$ the induced Frobenius norm of x .
- (2) We denote by $\mathbb{N}$ the set of strictly positive integers. Let $\mathbb{R}_+ := \{x \in \mathbb{R} : x \geq 0\}$. We denote by $\mathcal{B}(\mathbb{R}^d)$ the Borel σ -algebra on $\mathbb{R}^d$. For brevity, we write ∞ for $+\infty$. We denote $x \vee y := \max\{x, y\}$ and $x \wedge y := \min\{x, y\}$ for $x, y \in \mathbb{R}$.
- (3) We denote by ∇, ∇^2 the gradient and the Hessian with respect to (w.r.t) the spatial variable. We denote by ∂_t the derivative w.r.t time. Let $L^p(\mathbb{R}^d)$ be the Lebesgue space of real-valued p -integrable functions on $\mathbb{R}^d$.
- (4) Let $\mathcal{P}(\mathbb{R}^d)$ be the space of Borel probability measures on $\mathbb{R}^d$. The weak topology (and thus weak convergence $\rightharpoonup$) of $\mathcal{P}(\mathbb{R}^d)$ is the topology induced by $C_b(\mathbb{R}^d)$. The weak* topology (and thus weak* convergence $\rightharpoonup^*$) of $\mathcal{P}(\mathbb{R}^d)$ is the topology induced by $C_c(\mathbb{R}^d)$.

2. MAIN RESULTS

We recall a class of functions locally integrable in space-time. Let $p, q \in [1, \infty]$. The localized version $\tilde{L}^p(\mathbb{R}^d)$ of $L^p(\mathbb{R}^d)$ is defined by the norm

$$\|f\|_{\tilde{L}^p} := \sup_{x \in \mathbb{R}^d} \|1_{B(x,1)} f\|_{L^p},$$

Above, $B(x, r)$ is the open ball centered at x with radius r . For $0 \leq t_0 < t_1 \leq T$, we define the Bochner space

$$L_q^p(t_0, t_1) := L^q([t_0, t_1]; L^p(\mathbb{R}^d)).$$

The localized version $\tilde{L}_q^p(t_0, t_1)$ of $L_q^p(t_0, t_1)$ is defined by the norm

$$\|g\|_{\tilde{L}_q^p(t_0, t_1)} := \sup_{x \in \mathbb{R}^d} \|1_{B(x,1)} g\|_{L_q^p(t_0, t_1)}.$$

Then $\|g\|_{\tilde{L}_\infty^p(t_0, t_1)} = \|g\|_{L_\infty^p(t_0, t_1)}$. It holds for $p, q \in [1, \infty)$ that

$$\begin{aligned} \|g\|_{L_q^p(t_0, t_1)} &= \left(\int_{t_0}^{t_1} \left(\int_{\mathbb{R}^d} |g(s, y)|^p dy \right)^{\frac{q}{p}} ds \right)^{\frac{1}{q}}, \\ \|g\|_{\tilde{L}_q^p(t_0, t_1)} &= \sup_{x \in \mathbb{R}^d} \left(\int_{t_0}^{t_1} \left(\int_{B(x,1)} |g(s, y)|^p dy \right)^{\frac{q}{p}} ds \right)^{\frac{1}{q}}. \end{aligned}$$

For brevity, we denote

$$L_q^p(t) := L_q^p(0, t), \quad \tilde{L}_q^p(t) := \tilde{L}_q^p(0, t), \quad L_q^p := L_q^p(0, T), \quad \tilde{L}_q^p := \tilde{L}_q^p(0, T).$$

The class $\mathcal{K}$ of exponent parameter is defined by

$$\mathcal{K} := \left\{ (p, q) \in (2, \infty]^2 : \frac{d}{p} + \frac{2}{q} < 1 \right\}. \quad (2.1)$$

We denote by $M_p(\varrho)$ the p -th moment of $\varrho \in \mathcal{P}(\mathbb{R}^d)$, i.e., $M_p(\varrho) := \int_{\mathbb{R}^d} |x|^p d\varrho(x)$.

2.1. Main results. Below, we introduce the main assumption about the initial distribution and the coefficients of (1.1). Let $a := \sigma\sigma^\top$. We denote $b_t(x, r, \varrho) := b(t, x, r, \varrho)$, $\sigma_t(x, \varrho) := \sigma(t, x, \varrho)$ and $a_t(x, \varrho) := a(t, x, \varrho)$.

Assumption 2.1. Let $l \geq 1$ be an integer. There exist constants $p \in [1, \infty)$, $\beta \in (0, 1)$, $C > 0$ and, for $i \in \{0, 1, \dots, l\}$, function $1 \leq f_i \in \tilde{L}_{q_i}^{p_i}$ with $(p_i, q_i) \in \mathcal{K}$ such that for all $t \in [0, T]$; $x, y \in \mathbb{R}^d$; $r, r' \in \mathbb{R}_+$ and $\varrho, \varrho' \in \mathcal{P}_p(\mathbb{R}^d)$:

- (1) $a_t(x, \varrho)$ is invertible.
- (2) $|b_t(x, r, \varrho)| \leq f_0(t, x)$.
- (3) $x \mapsto \sigma_t(x, \varrho)$ is weakly differentiable and $|\nabla \sigma_t(x, \varrho)| \leq \sum_{i=1}^l f_i(t, x)$.
- (4) $\nu \in \mathcal{P}_p(\mathbb{R}^d)$ has a density $\ell_\nu \in L^\infty(\mathbb{R}^d)$.
- (5)

$$\begin{aligned} |\sigma_t(x, \varrho)| + |a_t^{-1}(x, \varrho)| &\leq C, \\ |b_t(x, r, \varrho) - b_t(x, r', \varrho')| &\leq C\{|r - r'| + W_p(\varrho, \varrho')\}, \\ |\sigma_t(x, \varrho) - \sigma_t(y, \varrho')| &\leq C\{|x - y|^\beta + W_p(\varrho, \varrho')\}. \end{aligned}$$

We gather parameters about (b, σ) in [Assumption 2.1](#):

$$\Theta_1 := (p, d, T, \beta, C, l, (p_i, q_i, \|f_i\|_{\tilde{L}_{q_i}^{p_i}})_{i=0}^l).$$

There is no continuity condition w.r.t spatial variable of b . [Assumption 2.1\(2\)](#) means that marginal density and marginal distribution do not affect local integrability of the drift. If b is bounded, then it satisfies [Assumption 2.1\(2\)](#).

Our main results are the following:

Theorem 2.2 (Existence). *Let [Assumption 2.1](#) hold. Then the following two statements hold:*

- (1) *The SDE (1.1) has a strong solution (X_t) whose marginal distribution is denoted by (μ_t) and marginal density is denoted by (ℓ_t) .*
- (2) *There exist constants $c_1 > 0$ (depending on Θ_1), $c_2 > 0$ (depending on Θ_1, ν), and $\delta \in (0, \frac{1}{2})$ (depending on q_0) such that*

$$\begin{aligned} \sup_{t \in [0, T]} \|\ell_t\|_\infty &\leq c_1 \|\ell_\nu\|_\infty, \\ W_p(\mu_s, \mu_t) &\leq c_2 |t - s|^\delta \quad \text{for } s, t \in [0, T]. \end{aligned}$$

Theorem 2.3 (Uniqueness). *Let [Assumption 2.1](#) hold. Assume in addition that $p = 1$, $\sigma_t(x, \varrho) = \sigma_t(x)$ and $|b_t(x, r, \varrho)| \leq C$ for all $(t, x, r, \varrho) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}_+ \times \mathcal{P}_p(\mathbb{R}^d)$. Then the following two statements hold:*

(1) For $k \in \{1, 2\}$, let $(X_t^k, t \in [0, T])$ be a weak solution to (1.1), ν_k its initial distribution satisfying Assumption 2.1(4) and $(\ell_t^k, t \in [0, T])$ its marginal density. Then there exists an increasing function $\Lambda : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ (depending on Θ_1) such that

$$\begin{aligned} & \sup_{t \in [0, T]} \int_{\mathbb{R}^d} (1 + |x|) |\ell_t^1(x) - \ell_t^2(x)| \, dx \\ & \leq \Lambda(\|\ell_{\nu_1}\|_\infty + M_1(\nu_1)) \int_{\mathbb{R}^d} (1 + |x|) |\ell_{\nu_1}(x) - \ell_{\nu_2}(x)| \, dx. \end{aligned}$$

(2) The SDE (1.1) has both weak and strong uniqueness.

The condition $p = 1$ in Theorem 2.3 is crucial for an application of Gronwall's lemma in its proof.

2.2. Outline of the proofs.

Existence. We will summarize the main ideas behind our mollifying argument. Let (ρ^n) be a sequence of mollifiers. We consider the SDE

$$\begin{cases} dX_t^n = b(t, X_t^n, \{\rho^n * \mu_t^n\}(X_t^n), \mu_t^n) dt + \sigma(t, X_t^n, \mu_t^n) dB_t, \\ \nu = \mathcal{L}(X_0^n) \quad \text{and} \quad \mu_t^n = \mathcal{L}(X_t^n). \end{cases} \quad (2.2)$$

Above, $*$ is the convolution operator, i.e.,

$$\{\rho^n * \varrho\}(x) := \int_{\mathbb{R}^d} \rho^n(x - y) \, d\varrho(y) \quad \text{for every } \varrho \in \mathcal{P}(\mathbb{R}^d).$$

Then (2.2) is well-posed and each X_t^n admits a density ℓ_t^n . The sequence (ℓ^n) is locally Hölder continuous on $(0, T] \times \mathbb{R}^d$. By Arzelà-Ascoli theorem, we can extract a subsequence that converges to some function $\ell : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}_+$ uniformly on every strip $[R^{-1}, T] \times B(0, R)$ where $R > 0$. We then verify that $\ell_t := \ell(t, \cdot)$ is indeed a density whose induced distribution $\mu_t \in \mathcal{P}_p(\mathbb{R}^d)$. Also, μ_t^n converges to μ_t in W_p (as $n \rightarrow \infty$) uniformly for $t \in [R^{-1}, T]$ where $R > 0$. By Itô's lemma, $(\mu_t^n, t \in [0, T])$ satisfies the Fokker-Planck equation

$$\partial_t \ell_t^n(x) = -\partial_{x_i} \{b(t, x, \{\rho^n * \ell_t^n\}(x), \mu_t^n) \ell_t^n(x)\} + \frac{1}{2} \partial_{x_i} \partial_{x_j} \{a^{i,j}(t, x, \mu_t^n) \ell_t^n(x)\}.$$

Above, we adapt Einstein summation convention. By the convergence of ℓ_t^n to ℓ_t , that of μ_t^n to μ_t , and the continuity of $b(t, x, r, \varrho)$ in (r, ϱ) , we deduce that (μ_t) satisfies

$$\partial_t \ell_t(x) = -\partial_{x_i} \{b(t, x, \ell_t(x), \mu_t) \ell_t(x)\} + \frac{1}{2} \partial_{x_i} \partial_{x_j} \{a^{i,j}(t, x, \mu_t) \ell_t(x)\}. \quad (2.3)$$

Notice that (2.3) is the Fokker-Planck equation associated with (1.1). By superposition principle (e.g. [BR20, Section 2]), (1.1) has a weak solution.

Uniqueness. We use the integral representation of marginal densities given by Lemma 3.10. Specifically, if $(X_t^1, t \in [0, T])$ and $(X_t^2, t \in [0, T])$ are two weak solutions with the same initial distribution, their marginal densities ℓ_t^1 and ℓ_t^2 satisfy an integral inequality involving the gradient of the transition kernel $p_{s,t}^{0,\sigma}$. Using the gradient estimate from Lemma 3.8(2) and the Wasserstein control from Lemma 3.2, we derive a Gronwall-type inequality for the weighted L^1 distance $f(t) = \int_{\mathbb{R}^d} (1 + |x|) |\ell_t^1 - \ell_t^2| \, dx$. This allows us to conclude $\ell_t^1 = \ell_t^2$ for all t . Once the marginal densities coincide, the coefficients of both solutions are identical. The problem reduces to a classical SDE with fully determined coefficients. Weak and strong uniqueness follow by applying [HW22, Theorem 1.1(1)].

3. PRELIMINARIES

3.1. Some facts from optimal transport. Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$. The set of *transport plans* (or *couplings*) between μ and ν is defined by

$$\Gamma(\mu, \nu) := \{\varrho \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d) : \mu = \pi_{\sharp}^1 \varrho \text{ and } \nu = \pi_{\sharp}^2 \varrho\}.$$

Above, π^i is the projection of $\mathbb{R}^d \times \mathbb{R}^d$ onto its i -th coordinate, and $\pi_{\sharp}^i \varrho \in \mathcal{P}(\mathbb{R}^d)$ is the push-forward measure of ϱ through π^i , i.e., $\{\pi_{\sharp}^i \varrho\}(A) = \varrho(\{\pi^i\}^{-1}(A))$ for all $A \in \mathcal{B}(\mathbb{R}^d)$. For $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$, we define

$$W_p(\mu, \nu) := \inf_{\varrho \in \Gamma(\mu, \nu)} \left\{ \int_{\mathbb{R}^d} |x - y|^p d\varrho(x, y) \right\}^{1/p}.$$

By [Vil09, Theorem 6.18], $(\mathcal{P}_p(\mathbb{R}^d), W_p)$ is a Polish space. By [Vil09, Theorem 6.9], it holds for $\mu_n, \mu \in \mathcal{P}_p(\mathbb{R}^d)$ that $W_p(\mu_n, \mu) \rightarrow 0$ if and only if $\mu_n \rightarrow \mu$ and $M_p(\mu_n) \rightarrow M_p(\mu)$. For brevity, we denote $W_p^p(\mu, \nu) := (W_p(\mu, \nu))^p$. We denote by $|\mu - \nu|$ the variation of the signed measure $\mu - \nu$ as in [Rud87, Section 6.1]. By [Rud87, Theorems 6.2 and 6.4], $|\mu - \nu|$ is a non-negative finite measure. We recall properties needed to prove our theorems:

Lemma 3.1. (1) [Vil03, Theorem 1.14] It holds for $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^d)$ that

$$W_1(\mu, \nu) = \sup_{\substack{f \in L^1(|\mu - \nu|) \\ [f]_1 \leq 1}} \int_{\mathbb{R}^d} f(x) d\{\mu - \nu\}(x),$$

where

$$[f]_1 := \sup_{\substack{x, y \in \mathbb{R}^d \\ x \neq y}} \frac{|f(x) - f(y)|}{|x - y|}.$$

(2) [Vil03, Remark 7.1.2] It holds for $1 \leq p \leq q < \infty$ and $\mu, \nu \in \mathcal{P}_q(\mathbb{R}^d)$ that $W_p(\mu, \nu) \leq W_q(\mu, \nu)$.

Above, the first claim is called Kantorovich duality while the second one is called Kantorovich-Rubinstein theorem. For more information about optimal transport, we refer to [FG21; Mag23; Vil09; Vil03; AG13; San15; ABS21; Tho23]. The next result states that W_p^p is controlled by weighted total variation distance.

Lemma 3.2. Let $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$ be absolutely continuous with corresponding densities ℓ_μ, ℓ_ν . Then

$$W_p^p(\mu, \nu) \leq (1 \vee 2^{p-1}) \int_{\mathbb{R}^d} |x|^p \times |\ell_\mu - \ell_\nu|(x) dx.$$

The proof of Lemma 3.2 is straightforward. For the sake of completeness, we include its proof in Section 7.

3.2. Moment estimates of marginal distribution. In the rest of Section 3, we consider measurable functions

$$\begin{aligned} b &: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d, \\ \sigma &: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^d. \end{aligned}$$

Let $a := \sigma \sigma^\top$. We denote $b_t := b(t, \cdot)$, $\sigma_t := \sigma(t, \cdot)$ and $a_t := a(t, \cdot)$. Let $\nu \in \mathcal{P}(\mathbb{R}^d)$. We consider the SDE

$$\begin{cases} dX_t = b(t, X_t) dt + \sigma(t, X_t) dB_t, \\ \nu = \mathcal{L}(X_0). \end{cases} \quad (3.1)$$

We consider the following set of assumption:

Assumption 3.3. The following conditions hold:

- (1) $a_t(x)$ is invertible for every $(t, x) \in [0, T] \times \mathbb{R}^d$.
- (2) There exist measurable maps $b^{(0)} : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $b^{(1)} : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that $b_t(x) = b_t^{(0)}(x) + b_t^{(1)}(x)$ for every $(t, x) \in [0, T] \times \mathbb{R}^d$.
- (3) There exists $1 \leq f_0 \in \tilde{L}_{q_0}^{p_0}$ with $(p_0, q_0) \in \mathcal{K}$ such that $|b_t^{(0)}(x)| \leq f_0(t, x)$ for every $(t, x) \in [0, T] \times \mathbb{R}^d$.
- (4) There exist constants $\beta \in (0, 1), C > 0$ such that for all $t \in [0, T]$ and $x, y \in \mathbb{R}^d$:

$$\begin{aligned} |b_t^{(1)}(x) - b_t^{(1)}(y)| &\leq C|x - y|, \\ |b_t^{(1)}(0)| + |\sigma_t(x)| + |a_t^{-1}(x)| &\leq C, \\ |\sigma_t(x) - \sigma_t(y)| &\leq C|x - y|^\beta. \end{aligned}$$

Under [Assumption 3.3](#), (3.1) has a unique weak solution (see e.g. [\[Wan23, Theorem 2.1\(1\)\]](#)). [Assumption 3.3](#) is appealing because it is a general but sufficient condition to obtain Krylov's and Khasminskii's estimates [\[Kry08\]](#), which are a key ingredient for establishing the other estimates in the remaining of [Section 3](#). We gather parameters in [Assumption 3.3](#):

$$\Theta_2 := (d, T, \beta, C, p_0, q_0, \|f_0\|_{\tilde{L}_{q_0}^{p_0}}).$$

We define the class $\bar{\mathcal{K}}$ of exponent parameter by

$$\bar{\mathcal{K}} := \left\{ (p, q) \in (1, \infty)^2 : \frac{d}{p} + \frac{2}{q} < 2 \right\}.$$

Remark 3.4. If $f \in \tilde{L}_q^p$ for some $(p, q) \in \mathcal{K}$ then there exists $(\bar{p}, \bar{q}) \in \bar{\mathcal{K}}$ such that $|f|^2 \in \tilde{L}_{\bar{q}}^{\bar{p}}$.

First, we establish the following estimates:

Lemma 3.5. *Let (b, σ) satisfy [Assumption 3.3](#) and (X_t) be the unique weak solution to (3.1). We fix $(p, q) \in \bar{\mathcal{K}}$. Then the following two statements hold:*

- (1) (Khasminskii's estimate) *There exist constants $c > 0, k > 1$ (depending on Θ_2, p, q) such that for $0 \leq t_0 < t_1 \leq T$ and $g \in \tilde{L}_q^p(t_0, t_1)$:*

$$\mathbb{E} \left[\exp \left(\int_{t_0}^{t_1} |g(s, X_s)| ds \right) \middle| \mathcal{F}_{t_0} \right] \leq \exp(c(1 + \|g\|_{\tilde{L}_q^p(t_0, t_1)}^k)). \quad (3.2)$$

- (2) (Krylov's estimate) *For $j \in [1, \infty)$, there exists a constant $c > 0$ (depending on Θ_2, p, q, j) such that for $0 \leq t_0 < t_1 \leq T$ and $g \in \tilde{L}_q^p(t_0, t_1)$:*

$$\mathbb{E} \left[\left(\int_{t_0}^{t_1} |g(s, X_s)| ds \right)^j \middle| \mathcal{F}_{t_0} \right] \leq c \|g\|_{\tilde{L}_q^p(t_0, t_1)}^j. \quad (3.3)$$

Proof of Lemma 3.5. (1) There exists $\bar{q} \in (1, q)$ such that $(p, \bar{q}) \in \bar{\mathcal{K}}$. By [\[ZY21, Theorem 3.1\]](#), there exists a constant $c_1 > 0$ (depending on $\Theta_2, p, \bar{q}$) such that for $0 \leq t_0 < t_1 \leq T$, stopping time τ and $g \in \tilde{L}_{\bar{q}}^p(t_0, t_1)$:

$$\mathbb{E} \left[\int_{t_0 \wedge \tau}^{t_1 \wedge \tau} |g(s, X_s)| ds \middle| \mathcal{F}_{t_0} \right] \leq c_1 \|g\|_{\tilde{L}_{\bar{q}}^p(t_0, t_1)}. \quad (3.4)$$

Let $\delta := \frac{1}{q} - \frac{1}{q} \in (0, 1)$. By Hölder's inequality, it holds for $0 \leq t_0 < t_1 \leq T$ and $g \in \tilde{L}_q^p(t_0, t_1)$ that

$$\|g\|_{\tilde{L}_q^p(t_0, t_1)} \leq (t_1 - t_0)^\delta \|g\|_{\tilde{L}_q^p(t_0, t_1)}. \quad (3.5)$$

We denote by I_j^n the open interval $(\frac{(j-1)(t_1-t_0)}{n}, \frac{j(t_1-t_0)}{n})$ for $j = 1, \dots, n$. We fix $g \in \tilde{L}_q^p(t_0, t_1) \subset \tilde{L}_q^p(t_0, t_1)$. Let $n \geq 2$ be the smallest integer such that

$$\|g\|_{\tilde{L}_q^p(I_j^n)} \leq \frac{1}{2c_1} \quad \text{for } j = 1, \dots, n. \quad (3.6)$$

By (3.4) and [XZ20, Lemma 3.5],

$$\mathbb{E} \left[\exp \left(\int_{t_0}^{t_1} |g(s, X_s)| \, ds \right) \middle| \mathcal{F}_{t_0} \right] \leq 2^n.$$

By (3.6), there exists $\bar{j} \in \{1, \dots, n-1\}$ such that

$$\|g\|_{\tilde{L}_q^p(I_{\bar{j}}^{n-1})} > \frac{1}{2c_1}. \quad (3.7)$$

By (3.5) and (3.7),

$$\left(\frac{t_1 - t_0}{n-1} \right)^\delta \|g\|_{\tilde{L}_q^p(I_{\bar{j}}^{n-1})} > \frac{1}{2c_1}.$$

Then

$$n < 1 + T(2c_1)^{-\frac{1}{\delta}} \|g\|_{\tilde{L}_q^p(t_0, t_1)}^{\frac{1}{\delta}}.$$

The estimate (3.2) then follows with $k := \frac{1}{\delta}$.

- (2) We follow an elegant idea in [HW22, Lemma 2.3]. Let $C_j := e^{j-1}$. We define $h : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ by $h(r) := |\ln(C_j + r)|^j$. Then h is concave. We have

$$\begin{aligned} & \mathbb{E} \left[\left(\int_{t_0}^{t_1} |g(s, X_s)| \, ds \right)^j \middle| \mathcal{F}_{t_0} \right] \\ & \leq \mathbb{E} \left[\left\{ \ln \left(C_j + \exp \left(\int_{t_0}^{t_1} |g(s, X_s)| \, ds \right) \right) \right\}^j \middle| \mathcal{F}_{t_0} \right] \\ & \leq \left\{ \ln \left(C_j + \mathbb{E} \left[\exp \left(\int_{t_0}^{t_1} |g(s, X_s)| \, ds \right) \middle| \mathcal{F}_{t_0} \right] \right) \right\}^j \quad \text{by Jensen's inequality} \\ & \leq \{ \ln[C_j + \exp(c(1 + \|g\|_{\tilde{L}_q^p(t_0, t_1)}^k))] \}^j. \end{aligned}$$

Above, the constants $c, k > 0$ are given by (3.2). As a result, there exists a constant $\bar{C}_j > 0$ (depending on c, j) such that

$$\mathbb{E} \left[\left(\int_{t_0}^{t_1} |g(s, X_s)| \, ds \right)^j \middle| \mathcal{F}_{t_0} \right] \leq \bar{C}_j (1 + \|g\|_{\tilde{L}_q^p(t_0, t_1)}^k)^j.$$

Replacing g with $\frac{g}{\|g\|_{\tilde{L}_q^p(t_0, t_1)}}$ in above inequality, we obtain

$$\mathbb{E} \left[\left(\int_{t_0}^{t_1} |g(s, X_s)| \, ds \right)^j \middle| \mathcal{F}_{t_0} \right] \leq \bar{C}_j 2^j \|g\|_{\tilde{L}_q^p(t_0, t_1)}^j.$$

The estimate (3.3) then follows. This completes the proof. ■

Second, we establish the following moment estimates:

Theorem 3.6. *Let (b, σ) satisfy [Assumption 3.3](#) and (X_t) be the unique weak solution to [\(3.1\)](#). Then the following two statements hold:*

- (1) X_t admits a density for $t \in (0, T]$.
- (2) Let $p \in [1, \infty)$. There exist constants $c > 0$ (depending on Θ_2, p) and $\delta \in (0, \frac{1}{2})$ (depending on q_0) such that for $0 \leq u \leq t \leq T$:

$$\mathbb{E} \left[\sup_{s \in [u, t]} |X_s|^p \right] \leq c(1 + \mathbb{E}[|X_u|^p]), \quad (3.8)$$

$$\mathbb{E} \left[\sup_{s \in [u, t]} |X_s - X_u|^p \right] \leq c|t - u|^{\delta p} (1 + \mathbb{E}[|X_u|^p]). \quad (3.9)$$

Proof. By [Assumption 3.3\(2\)](#), there exist measurable maps $b^{(0)} : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $b^{(1)} : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that $b_t(x) = b_t^{(0)}(x) + b_t^{(1)}(x)$. By [Assumption 3.3\(3\)](#), there exists $f_0 \in \tilde{L}_{q_0}^{p_0}$ with $(p_0, q_0) \in \mathcal{K}$ such that $|b_t^{(0)}(x)| \leq f_0(t, x)$. We consider the SDE

$$d\bar{X}_t = b^{(1)}(t, \bar{X}_t) dt + \sigma(t, \bar{X}_t) dB_t. \quad (3.10)$$

Above, the distribution of $\bar{X}_0$ is ν . Clearly, $(b^{(1)}, \sigma)$ satisfies [Assumption 3.3](#), so [\(3.10\)](#) is well-posed. We define

$$\begin{aligned} \xi_t &:= \{\sigma_t^\top a_t^{-1} b_t^{(0)}\}(\bar{X}_t), \\ \bar{B}_t &:= B_t - \int_0^t \xi_s ds, \\ R_t &:= \exp \left(\int_0^t \xi_s^\top dB_s - \frac{1}{2} \int_0^t |\xi_s|^2 ds \right), \\ I_t &:= \mathbb{E} \left[\exp \left(\frac{1}{2} \int_0^t |\xi_s|^2 ds \right) \right]. \end{aligned}$$

By uniform boundedness of $\sigma_t^\top a_t^{-1}$, [Remark 3.4](#) and [Lemma 3.5\(1\)](#), we have $I_T < \infty$. So R_T is an exponential martingale with $\mathbb{E}[R_T] = 1$. By Girsanov theorem, $(\bar{B}_t, t \in [0, T])$ is a d -dimensional Brownian motion under the probability measure $\bar{\mathbb{P}} := R_T \mathbb{P}$. We denote by $\bar{\mathbb{E}}$ the expectation w.r.t $\bar{\mathbb{P}}$. Clearly, [\(3.10\)](#) can be written under $\bar{\mathbb{P}}$ as

$$d\bar{X}_t = b(t, \bar{X}_t) dt + \sigma(t, \bar{X}_t) d\bar{B}_t. \quad (3.11)$$

- (1) By [Assumption 3.3\(4\)](#) and [[MPZ21](#), Theorem 1.2], the distribution of $\bar{X}_t$ under $\bar{\mathbb{P}}$ admits a density. Notice that $\bar{\mathbb{P}}$ and $\mathbb{P}$ are equivalent, so the distribution of $\bar{X}_t$ under $\bar{\mathbb{P}}$ also admits a density. Because $\bar{X}_0$ is $\mathcal{F}_0$ -measurable, it holds for $\varphi \in C_c^\infty(\mathbb{R}^d)$ that

$$\bar{\mathbb{E}}[\varphi(\bar{X}_0)] = \mathbb{E}[\varphi(\bar{X}_0)R_0] = \mathbb{E}[\varphi(\bar{X}_0)].$$

Then ν is also the distribution of $\bar{X}_0$ under $\bar{\mathbb{P}}$. By weak uniqueness of [\(3.1\)](#) and [\(3.11\)](#), the distribution of X_t under $\mathbb{P}$ is the same as that of $\bar{X}_t$ under $\bar{\mathbb{P}}$. Thus the distribution of X_t under $\mathbb{P}$ admits a density.

- (2) The estimate (3.8) follows from [HW22, Theorem 1.1(1)]. It remains to prove (3.9). Recall that f_0 is a parameter in Assumption 3.3. We have

$$\begin{aligned} \mathbb{E}\left[\sup_{s \in [u, t]} |X_s - X_u|^p\right] &\lesssim \mathbb{E}\left[\left(\int_u^t f_0(r, X_r) \, dr\right)^p\right] \\ &\quad + \mathbb{E}\left[\left(\int_u^t |b^{(1)}(r, X_r)| \, dr\right)^p\right] \\ &\quad + \mathbb{E}\left[\sup_{s \in [u, t]} \left|\int_u^s \sigma(r, X_r) \, dB_r\right|^p\right] \\ &=: J_1 + J_2 + J_3. \end{aligned}$$

There exists $\bar{q}_0 \in (2, q_0)$ such that $(p_0, \bar{q}_0) \in \bar{K}$. Let $\delta := \frac{1}{\bar{q}_0} - \frac{1}{q_0} \in (0, \frac{1}{2})$. By Hölder's inequality,

$$\|f_0\|_{\tilde{L}_{\bar{q}_0}^{p_0}(u, t)} \leq (t - u)^\delta \|f_0\|_{\tilde{L}_{q_0}^{p_0}(u, t)}.$$

By Lemma 3.5(2),

$$J_1 \lesssim \|f_0\|_{\tilde{L}_{\bar{q}_0}^{p_0}(u, t)}^p \leq (t - u)^{\delta p} \|f_0\|_{\tilde{L}_{q_0}^{p_0}(u, t)}^p \lesssim (t - u)^{\delta p}.$$

As for J_2 and J_3 , we have

$$\begin{aligned} J_2 &\lesssim |t - u|^{p-1} \mathbb{E}\left[\int_u^t (1 + |X_r|^p) \, dr\right] \\ &\lesssim |t - u|^p (1 + \mathbb{E}[|X_u|^p]) \quad \text{by (3.8),} \\ J_3 &\lesssim |t - u|^{\frac{p}{2}}. \end{aligned}$$

Then

$$\mathbb{E}\left[\sup_{s \in [u, t]} |X_s - X_u|^p\right] \lesssim |t - u|^{\delta p} (1 + \mathbb{E}[|X_u|^p]).$$

The estimate (3.9) then follows. This completes the proof. ■

3.3. Heat kernel estimates. For $(s, x) \in [0, T] \times \mathbb{R}^d$, we consider the SDE

$$\begin{cases} dX_{s,t}^x &= b(t, X_{s,t}^x) \, dt + \sigma(t, X_{s,t}^x) \, dB_t \quad \text{for } t \in (s, T], \\ X_{s,s}^x &= x. \end{cases} \quad (3.12)$$

If $X_{s,t}^x$ admits a density, we denote it by $p_{s,t}^{b,\sigma}(x, \cdot)$. We consider the following set of assumption:

Assumption 3.7. The following conditions hold:

- (1) $a_t(x)$ is invertible for every $(t, x) \in [0, T] \times \mathbb{R}^d$.
- (2) There exists $1 \leq f_0 \in \tilde{L}_{q_0}^{p_0}$ with $(p_0, q_0) \in \mathcal{K}$ such that $|b_t(x)| \leq f_0(t, x)$ for every $(t, x) \in [0, T] \times \mathbb{R}^d$.
- (3) There exist constants $\beta \in (0, 1), C > 0$ such that for all $t \in [0, T]$ and $x, y \in \mathbb{R}^d$:

$$\begin{aligned} |\sigma_t(x)| + |a_t^{-1}(x)| &\leq C, \\ |\sigma_t(x) - \sigma_t(y)| &\leq C|x - y|^\beta. \end{aligned}$$

Notice that [Assumption 3.7](#) is [Assumption 3.3](#) with $b^{(1)} = 0$. As such, under [Assumption 3.7](#), (3.12) has a unique weak solution; and for $t \in (s, T]$, $X_{s,t}^x$ admits a density [by [Theorem 3.6\(1\)](#)]. In this case, $\{p_{s,t}^{b,\sigma} : 0 \leq s < t \leq T\}$ are transition densities associated with (3.12). We gather parameters from [Assumption 3.7](#):

$$\Theta_3 := (d, T, \beta, C, p_0, q_0, \|f_0\|_{\tilde{L}_{q_0}^{p_0}}).$$

For $\lambda > 0$ and $\gamma \in \mathbb{R}$, the heat kernel $p^{\gamma,\lambda}$ is defined for $t > 0$ and $x \in \mathbb{R}^d$ by

$$p_t^{\gamma,\lambda}(x) := \frac{1}{t^{(\gamma+d)/2}} e^{-\frac{\lambda|x|^2}{t}}.$$

First, we recall the following estimates:

Lemma 3.8. [[Zha24](#), Lemma 3.9] *Let (b, σ) satisfy [Assumption 3.7](#). Let $\gamma_0 := 1 - \frac{d}{p_0} - \frac{2}{q_0}$. Then the map $x \mapsto p_{s,t}^{b,\sigma}(x, y)$ is differentiable for every $(t, y) \in (s, T] \times \mathbb{R}^d$. Moreover, there exist constants $c_1, c_2, c_3 > 0$ and $\lambda \in (0, 1)$ depending on Θ_3 such that*

(1) (Gaussian estimate) for $0 \leq s < t \leq T$ and $x, y \in \mathbb{R}^d$:

$$c_1 p_{t-s}^{0,\lambda^{-1}}(y-x) \leq p_{s,t}^{b,\sigma}(x, y) \leq c_2 p_{t-s}^{0,\lambda}(y-x).$$

(2) (Gradient estimate in x) for $0 \leq s < t \leq T$ and $x, y \in \mathbb{R}^d$:

$$|\nabla_x p_{s,t}^{b,\sigma}(x, y)| \leq c_3 p_{t-s}^{1,\lambda}(y-x).$$

(3) (Hölder estimate in t and y) for $\gamma \in (0, \beta \wedge \gamma_0)$, there exists a constant $c_4 > 0$ depending on (Θ_3, γ) such that for $0 \leq s < t_1 < t_2 \leq T$ and $x, y, y_1, y_2 \in \mathbb{R}^d$:

$$\begin{aligned} |p_{s,t_2}^{b,\sigma}(x, y) - p_{s,t_1}^{b,\sigma}(x, y)| &\leq c_4 |t_1 - t_2|^{\frac{\gamma}{2}} \sum_{i=1}^2 p_{t_i-s}^{\gamma,\lambda}(x-y), \\ |p_{s,t}^{b,\sigma}(x, y_1) - p_{s,t}^{b,\sigma}(x, y_2)| &\leq c_4 |y_1 - y_2|^\gamma \sum_{i=1}^2 p_{t-s}^{\gamma,\lambda}(x-y_i). \end{aligned}$$

We will need the following results:

Corollary 3.9. *Let (b, σ) satisfy [Assumption 3.7](#) and (X_t) be the unique weak solution to (3.1). Let ν admit a density $\ell_\nu \in L^\infty(\mathbb{R}^d)$. We denote by ℓ_t the density of X_t . Then the following two statements hold:*

(1) *There exists a constant $c_1 > 0$ depending on Θ_3 such that:*

$$\sup_{t \in [0, T]} \|\ell_t\|_\infty \leq c_1 \|\ell_\nu\|_\infty.$$

(2) *Let $\gamma_0 := 1 - \frac{d}{p_0} - \frac{2}{q_0}$. For $\gamma \in (0, \beta \wedge \gamma_0)$, there exists a constant $c_2 > 0$ depending on (Θ_3, γ, ν) such that for $0 < s, t \leq T$ and $x, y \in \mathbb{R}^d$:*

$$\begin{aligned} |\ell_t(x) - \ell_s(x)| &\leq c_2 |t - s|^{\frac{\gamma}{2}} (t^{-\frac{\gamma}{2}} + s^{-\frac{\gamma}{2}}), \\ |\ell_t(x) - \ell_t(y)| &\leq c_2 |x - y|^\gamma t^{-\frac{\gamma}{2}}. \end{aligned}$$

Lemma 3.10. *Let (b, σ) satisfy [Assumption 3.7](#) and (X_t) be the unique weak solution to (3.1). Let ν admit a density ℓ_ν . We denote by ℓ_t the density of X_t . Then it holds for $t \in [0, T]$ and $x \in \mathbb{R}^d$:*

$$\ell_t(x) = \int_{\mathbb{R}^d} p_{0,t}^{0,\sigma}(y, x) \ell_\nu(y) dy + \int_0^t \int_{\mathbb{R}^d} \ell_s(y) \langle b_s(y), \nabla_y p_{s,t}^{0,\sigma}(y, x) \rangle dy ds.$$

Lemma 3.11. *Let $p \in \{0\} \cup [1, \infty)$ and $\nu \in \mathcal{P}_p(\mathbb{R}^d)$. Let (b, σ) satisfy [Assumption 3.7](#) and (X_t) be the unique weak solution to [\(3.1\)](#). We denote by μ_t the distribution of X_t . There exists a function $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ depending on (Θ_3, ν, p) such that $\lim_{R \rightarrow \infty} \phi(R) = 0$ and that*

$$\sup_{t \in [0, T]} \int_{B^c(0, R)} |x|^p d\mu_t(x) \leq \phi(R) \quad \text{for every } R \geq 0.$$

Above, $B^c(0, R) := \mathbb{R}^d \setminus B(0, R)$ where $B(0, R) := \{x \in \mathbb{R}^d : |x| \leq R\}$.

The proof of [Corollary 3.9](#) is a straightforward application of [Lemma 3.8](#) and the fact that $\ell_t(x) = \int_{\mathbb{R}^d} \ell_\nu(y) p_{0,t}^{b,\sigma}(y, x) dy$. The proof of [Lemma 3.10](#) is contained in that of [[Wan23](#), Lemma 4.1]. The proof of [Lemma 3.11](#) is straightforward application of Markov's inequality and [Lemma 3.8\(1\)](#). For the sake of completeness, we still include their proofs in [Section 7](#).

3.4. A well-posedness result for classical McKean-Vlasov SDEs. Our final ingredient is a well-posedness result from [[HW22](#)]. For $p \in [1, \infty)$, we consider measurable functions

$$\begin{aligned} b &: [0, T] \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d) \rightarrow \mathbb{R}^d, \\ \sigma &: [0, T] \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d) \rightarrow \mathbb{R}^d \otimes \mathbb{R}^d. \end{aligned}$$

and the corresponding classical McKean-Vlasov SDEs:

$$\begin{cases} dX_t = b(t, X_t, \mu_t) dt + \sigma(t, X_t, \mu_t) dB_t, \\ \nu = \mathcal{L}(X_0) \quad \text{and} \quad \mu_t = \mathcal{L}(X_t). \end{cases} \quad (3.13)$$

We denote by $a = \sigma\sigma^\top$ the diffusion matrix.

Theorem 3.12. [[HW22](#), Theorem 1.1(1)] *Assume that the following four conditions hold:*

(1) *There exist measurable maps $b^{(0)} : [0, T] \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d) \rightarrow \mathbb{R}^d$ and $b^{(1)} : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that $b_t(x, \varrho) = b_t^{(0)}(x, \varrho) + b_t^{(1)}(x)$ for every $(t, x, \varrho) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d)$.*

(2) *$\nu \in \mathcal{P}_p(\mathbb{R}^d)$ and*

$$\lim_{\varepsilon \downarrow 0} \sup_{\mu \in C([0, T]; \mathcal{P}_p(\mathbb{R}^d))} \sup_{t \in [0, T], |x-y| \leq \varepsilon} |a_t(x, \mu_t) - a_t(y, \mu_t)| = 0.$$

(3) *$a_t(x, \varrho)$ is invertible and $x \mapsto \sigma_t(x, \varrho)$ is weakly differentiable for every $(t, x, \varrho) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d)$.*

(4) *There exist constants $(l, C) \in \mathbb{N} \times \mathbb{R}_+$ and, for $i \in \{0, 1, \dots, l\}$, function $1 \leq f_i \in \tilde{L}_{q_i}^{p_i}$ with $(p_i, q_i) \in \mathcal{K}$ such that for all $(t, x) \in [0, T] \times \mathbb{R}^d$; $\varrho, \varrho' \in \mathcal{P}_p(\mathbb{R}^d)$ and $\mu \in C([0, T]; \mathcal{P}_p(\mathbb{R}^d))$:*

$$\begin{aligned} |b_t^{(1)}(0)| + |\sigma_t(x, \mu_t)| + |a_t^{-1}(x, \mu_t)| &\leq C, \\ |b_t^{(0)}(x, \mu_t)| &\leq f_0(t, x) + CW_p(\mu_t, \delta_0), \\ |\nabla \sigma_t(x, \mu_t)| &\leq \sum_{i=1}^l f_i(t, x), \\ |b_t^{(1)}(x) - b_t^{(1)}(y)| &\leq C|x - y|, \\ |\sigma_t(x, \varrho) - \sigma_t(x, \varrho')| + |b_t(x, \varrho) - b_t(x, \varrho')| &\leq W_p(\varrho, \varrho') \sum_{i=0}^l f_i(t, x). \end{aligned}$$

Then the SDE [\(3.13\)](#) is well-posed.

Recall from our conventions in [Section 1](#) that being well-posed means being both weakly and strongly well-posed. Also, the set $\mathcal{K}$ of exponent parameters is defined by [\(2.1\)](#). The following special case of [Theorem 3.12](#) suffices for our purpose:

Corollary 3.13. *Assume that*

- (1) $a_t(x, \varrho)$ is invertible and $x \mapsto \sigma_t(x, \varrho)$ is weakly differentiable for every $(t, x, \varrho) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d)$.
- (2) There exist constants $(l, C, \beta) \in \mathbb{N} \times \mathbb{R}_+ \times (0, 1)$ and, for $i \in \{0, 1, \dots, l\}$, function $1 \leq f_i \in \tilde{L}_{q_i}^{p_i}$ with $(p_i, q_i) \in \mathcal{K}$ such that for all $t \in [0, T]; x, y \in \mathbb{R}^d$ and $\varrho, \varrho' \in \mathcal{P}_p(\mathbb{R}^d)$:

$$\begin{aligned} |\sigma_t(x, \varrho)| + |a_t^{-1}(x, \varrho)| &\leq C, \\ |b_t(x, \varrho)| &\leq f_0(t, x), \\ |\nabla \sigma_t(x, \varrho)| &\leq \sum_{i=1}^l f_i(t, x), \\ |b_t(x, \varrho) - b_t(x, \varrho')| &\leq CW_p(\varrho, \varrho'), \\ |\sigma_t(x, \varrho) - \sigma_t(y, \varrho')| &\leq C\{|x - y|^\beta + W_p(\varrho, \varrho')\}. \end{aligned}$$

Then the SDE [\(3.13\)](#) is well-posed.

[Corollary 3.13](#) is simpler than [Theorem 3.12](#) in that

- b is not allowed to have Lipschitz component, i.e., $b^{(1)} = 0$.
- σ is Hölder continuous in space uniformly in time and distribution variables.
- b and σ is Lipschitz continuous in distribution variable uniformly in time and space variables.

Remark 3.14. In dimension $d = 1$, consider the SDE $dX_t = \sigma(X_t)dB_t$ where $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ satisfies the conditions of [Corollary 3.13](#) with $b = 0$. In particular, σ is weakly differentiable with $|\sigma'(x)| \leq f_1(x)$ for some $f_1 \in \tilde{L}^{p_1}$ with $(p_1, q_1) \in \mathcal{K}$. Hence $p_1 > 2$.

Since $\sigma \in W_{\text{loc}}^{1, p_1}(\mathbb{R})$, the fundamental theorem of calculus (see e.g. [\[Bre11, Theorem 8.2\]](#)) gives $\sigma(y) - \sigma(x) = \int_x^y \sigma'(t) dt$ for all $x, y \in \mathbb{R}$. For $|y - x| \leq 1$, Hölder's inequality implies

$$|\sigma(x) - \sigma(y)| \leq \int_x^y |f_1(t)| dt \leq \|f_1\|_{\tilde{L}^{p_1}} |x - y|^{1 - \frac{1}{p_1}}.$$

If $|y - x| > 1$ then $|x - y|^{1 - \frac{1}{p_1}} \geq 1$ and thus

$$|\sigma(x) - \sigma(y)| \leq 2\|\sigma\|_\infty \leq 2\|\sigma\|_\infty |x - y|^{1 - \frac{1}{p_1}}.$$

Thus $\sigma \in C^\beta(\mathbb{R})$ with $\beta = 1 - \frac{1}{p_1} > \frac{1}{2}$, since $p_1 > 2$. In particular, the counterexamples of [\[Bar82\]](#), which show failure of strong existence for $\beta < \frac{1}{2}$, do not apply in this setting.

4. PROOF OF [THEOREM 2.2](#)

We recall that $\Theta_1 = (p, d, T, \beta, C, l, (p_i, q_i, f_i)_{i=0}^l)$ contains parameters about (b, σ) from [Assumption 2.1](#). We write $M_1 \lesssim M_2$ if there exists a constant $c > 0$ (depending on Θ_1, ν) such that $M_1 \leq cM_2$. We construct a sequence (ρ^n) of mollifiers as follows. We fix a smooth density

$\rho : \mathbb{R}^d \rightarrow \mathbb{R}$ whose support is contained in $B(0, 1)$. For each $n \in \mathbb{N}$, we define $\rho^n : \mathbb{R}^d \rightarrow \mathbb{R}$ by $\rho^n(x) := n^d \rho(nx)$ and consider the McKean-Vlasov SDE

$$\begin{cases} dX_t^n = b(t, X_t^n, (\rho^n * \mu_t^n)(X_t^n), \mu_t^n) dt + \sigma(t, X_t^n, \mu_t^n) dB_t, \\ \nu = \mathcal{L}(X_0^n) \quad \text{and} \quad \mu_t^n = \mathcal{L}(X_t^n). \end{cases} \quad (4.1)$$

Then (4.1) is a mollified version of (1.1).

4.1. Stability estimates for mollified SDEs. We define the map

$$b^n : [0, T] \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d) \rightarrow \mathbb{R}^d$$

by

$$b^n(t, x, \varrho) := b(t, x, (\rho^n * \varrho)(x), \varrho).$$

Then (4.1) becomes

$$\begin{cases} dX_t^n = b^n(t, X_t^n, \mu_t^n) dt + \sigma(t, X_t^n, \mu_t^n) dB_t, \\ \nu = \mathcal{L}(X_0^n) \quad \text{and} \quad \mu_t^n = \mathcal{L}(X_t^n). \end{cases} \quad (4.2)$$

Lemma 4.1. *For each $n \in \mathbb{N}$, the pair (b^n, σ) satisfy the assumptions required by Corollary 3.13.*

Proof. We have

$$|b^n(t, x, \varrho) - b^n(t, x, \tilde{\varrho})| \lesssim \left| \int_{\mathbb{R}^d} \rho^n(x - y) d(\varrho - \tilde{\varrho})(y) \right| + W_p(\varrho, \tilde{\varrho}) \quad (4.3)$$

$$\leq \|\nabla \rho^n\|_\infty W_1(\varrho, \tilde{\varrho}) + W_p(\varrho, \tilde{\varrho}) \quad (4.4)$$

$$\leq (1 + \|\nabla \rho^n\|_\infty) W_p(\varrho, \tilde{\varrho}). \quad (4.5)$$

Above, (4.3) is due to Assumption 2.1(5), (4.4) due to Lemma 3.1(1), and (4.5) due to Lemma 3.1(2). It follows that b^n is Lipschitz in distribution variable uniformly in (t, x) . By Assumption 2.1(2), we have for any $(t, x, r, \varrho) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}_+ \times \mathcal{P}_p(\mathbb{R}^d)$ that $|b(t, x, r, \varrho)| \leq C f_0(t, x)$. Then the drift b^n satisfies

$$|b^n(t, x, \varrho)| = |b(t, x, (\rho^n * \varrho)(x), \varrho)| \leq C f_0(t, x).$$

By Assumption 2.1(3), σ is weakly differentiable in space such that $|\nabla \sigma_t(x, \varrho)| \leq \sum_{i=1}^l f_i(t, x)$. By Assumption 2.1(1), σ is invertible. By Assumption 2.1(5), σ is uniformly elliptic, Lipschitz in distribution variable (uniformly in time and space), and Hölder continuous in space (uniformly in time and distribution variables). This completes the proof. $\blacksquare$

By Lemma 4.1, all conditions of Corollary 3.13 are satisfied, implying that the SDE (4.2) has a unique strong solution for each $n \in \mathbb{N}$.

Remark 4.2. The Lipschitz continuity (4.5) of $\varrho \mapsto b^n(t, x, \varrho)$ is just for applying Corollary 3.13. The size of its Lipschitz constant does not play any role below.

We define the maps

$$\bar{b}^n : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d,$$

$$\bar{\sigma}^n : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^d,$$

by

$$\bar{b}^n(t, x) := b^n(t, x, \mu_t^n),$$

$$\bar{\sigma}^n(t, x) := \sigma(t, x, \mu_t^n).$$

Then (4.2) becomes

$$dX_t^n = \bar{b}^n(t, X_t^n) dt + \bar{\sigma}^n(t, X_t^n) dB_t. \quad (4.6)$$

Let $\bar{a}^n := \bar{\sigma}^n(\bar{\sigma}^n)^\top$. We denote

$$\bar{b}_t^n := \bar{b}^n(t, \cdot), \quad \bar{\sigma}_t^n := \bar{\sigma}^n(t, \cdot), \quad \bar{a}_t^n := \bar{a}^n(t, \cdot).$$

Lemma 4.3. *All pairs $(\bar{b}^n, \bar{\sigma}^n)_{n \in \mathbb{N}}$ satisfy Assumption 3.7 for the same set of parameters.*

Proof. By Assumption 2.1, we have for any $(t, r, \varrho) \in [0, T] \times \mathbb{R}_+ \times \mathcal{P}_p(\mathbb{R}^d)$ and $x, y \in \mathbb{R}^d$ that

$$\begin{aligned} |b(t, x, r, \varrho)| &\leq C f_0(t, x), \\ \|\sigma_t\|_\infty + \|a_t^{-1}\|_\infty &\leq C, \\ |\sigma_t(x, \varrho) - \sigma_t(y, \varrho)| &\leq C|x - y|^\beta. \end{aligned}$$

Then

$$\begin{aligned} |\bar{b}_t^n(x)| &= |b^n(t, x, \mu_t^n)| = |b(t, x, (\rho^n * \mu_t^n)(x), \mu_t^n)| \leq C f_0(t, x), \\ |\bar{\sigma}_t^n(x) - \bar{\sigma}_t^n(y)| &= |\sigma_t(x, \mu_t^n) - \sigma_t(y, \mu_t^n)| \leq C|x - y|^\beta. \end{aligned}$$

The uniform ellipticity of $\bar{\sigma}^n$ is inherited from that of σ . This completes the proof. $\blacksquare$

Since $(\bar{b}^n, \bar{\sigma}^n)$ satisfies Assumption 3.7 (hence Assumption 3.3) by Lemma 4.3, each X_t^n admits a density denoted by ℓ_t^n by Theorem 3.6(1).

By Theorem 3.6(2), there exists a constant $\delta \in (0, \frac{1}{2})$ depending on q_0 such that

$$\sup_{n \in \mathbb{N}} \sup_{t \in [0, T]} M_p(\mu_t^n) \lesssim 1, \quad (4.7)$$

$$\sup_{n \in \mathbb{N}} W_p(\mu_s^n, \mu_t^n) \lesssim |t - s|^\delta \quad \text{for all } s, t \in [0, T]. \quad (4.8)$$

Let $\gamma_0 := 1 - \frac{d}{p_0} - \frac{2}{q_0}$ and $\gamma := \frac{\beta \wedge \gamma_0}{2}$. By Corollary 3.9(1),

$$\sup_{n \in \mathbb{N}} \sup_{t \in [0, T]} \|\ell_t^n\|_\infty \lesssim \|\ell_\nu\|_\infty. \quad (4.9)$$

By Corollary 3.9(2), it holds for $t \in (0, T]$ that

$$\sup_{n \in \mathbb{N}} \sup_{s, r \in [t, T]} \sup_{\substack{x \in \mathbb{R}^d \\ s \neq r}} \frac{|\ell_s^n(x) - \ell_r^n(x)|}{|s - r|^{\frac{\gamma}{2}}} \lesssim t^{-\frac{\gamma}{2}}, \quad (4.10)$$

$$\sup_{n \in \mathbb{N}} \sup_{s \in [t, T]} \sup_{\substack{x, y \in \mathbb{R}^d \\ x \neq y}} \frac{|\ell_s^n(x) - \ell_s^n(y)|}{|x - y|^\gamma} \lesssim t^{-\frac{\gamma}{2}}. \quad (4.11)$$

By Lemma 3.11, there exists a function $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ depending on (Θ_1, ν) such that $\lim_{R \rightarrow \infty} \phi(R) = 0$ and that

$$\sup_{n \in \mathbb{N}} \sup_{t \in [0, T]} \int_{B^c(0, R)} (1 + |x|^p) d\mu_t^n(x) \leq \phi(R) \quad \text{for every } R \geq 0. \quad (4.12)$$

By (4.8), the map $[0, T] \rightarrow \mathcal{P}_p(\mathbb{R}^d)$, $t \mapsto \mu_t^n$ is continuous.

4.2. Convergence of marginal densities of mollified SDEs. By (4.10), (4.11), Arzelà-Ascoli theorem and diagonal extraction, there exist a sub-sequence (also denoted by (ℓ^n) for simplicity) and a continuous function $\ell : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}_+$ such that

$$\lim_n \sup_{t \in [R^{-1}, T]} \sup_{x \in B(0, R)} |\ell_t^n(x) - \ell_t(x)| = 0 \quad \text{for every } R > T^{-1}. \quad (4.13)$$

Above, $\ell_t := \ell(t, \cdot)$. Clearly, $\ell_0 = \ell_\nu$ and

$$\sup_{t \in [0, T]} \|\ell_t\|_\infty \lesssim \|\ell_\nu\|_\infty. \quad (4.14)$$

We remark that the constant in (4.14) depends on Θ_1 . Next we verify that ℓ_t is indeed a density for every $t \in (0, T]$. We have

$$\begin{aligned} \int_{B(0, R)} \ell_t^n(x) dx &= 1 - \int_{B^c(0, R)} \ell_t^n(x) dx \\ &\gtrsim 1 - \phi(R) \quad \text{by (4.12)}. \end{aligned}$$

By (4.13), (4.14) and dominated convergence theorem,

$$\int_{B(0, R)} \ell_t(x) dx = \lim_{n \rightarrow \infty} \int_{B(0, R)} \ell_t^n(x) dx.$$

It follows that

$$1 - \phi(R) \lesssim \int_{B(0, R)} \ell_t(x) dx \leq 1.$$

Then

$$\int_{\mathbb{R}^d} \ell_t(x) dx = \lim_{R \rightarrow \infty} \int_{B(0, R)} \ell_t(x) dx = 1.$$

Let $\mu_t \in \mathcal{P}(\mathbb{R}^d)$ be the probability measure induced by ℓ_t , i.e.,

$$\mu_t(B) := \int_B \ell_t(x) dx \quad \text{for every } B \in \mathcal{B}(\mathbb{R}^d).$$

Lemma 4.4. *We have for each $t \in [0, T]$ that $\mu_t^n \rightharpoonup \mu_t$ as $n \rightarrow \infty$.*

Proof. It suffices to consider $t \in (0, T]$. By (4.13),

$$\mu_t^n \xrightarrow{*} \mu_t \quad \text{as } n \rightarrow \infty. \quad (4.15)$$

Let $f \in C_b(\mathbb{R}^d)$ and $g \in C_c(\mathbb{R}^d)$ such that $0 \leq g \leq 1$. Then $gf \in C_c(\mathbb{R}^d)$ and $f = (1 - g)f + gf$. We have

$$\begin{aligned} \left| \int_{\mathbb{R}^d} f(x) d\{\mu_t^n - \mu_t\}(x) \right| &\leq \|f\|_\infty \int_{\mathbb{R}^d} \{1 - g(x)\} d\{\mu_t^n + \mu_t\}(x) \\ &\quad + \left| \int_{\mathbb{R}^d} g(x) f(x) d\{\mu_t^n - \mu_t\}(x) \right|. \end{aligned}$$

By (4.15),

$$\limsup_n \int_{\mathbb{R}^d} g(x) f(x) d\{\mu_t^n - \mu_t\}(x) = 0.$$

Then

$$\limsup_n \left| \int_{\mathbb{R}^d} f(x) d\{\mu_t^n - \mu_t\}(x) \right| \leq \|f\|_\infty \limsup_n \int_{\mathbb{R}^d} \{1 - g(x)\} d\{\mu_t^n + \mu_t\}(x).$$

Notice that

$$\limsup_n \int_{\mathbb{R}^d} \{1 - g(x)\} d\mu_t^n(x) = 1 - \liminf_n \int_{\mathbb{R}^d} g(x) d\mu_t^n(x)$$

$$\begin{aligned}
&= 1 - \int_{\mathbb{R}^d} g(x) \, d\mu_t(x) \quad \text{by (4.15)} \\
&= \int_{\mathbb{R}^d} \{1 - g(x)\} \, d\mu_t(x).
\end{aligned}$$

Thus

$$\limsup_n \left| \int_{\mathbb{R}^d} f(x) \, d\{\mu_t^n - \mu_t\}(x) \right| \leq 2\|f\|_\infty \int_{\mathbb{R}^d} \{1 - g(x)\} \, d\mu_t(x). \quad (4.16)$$

Because μ_t is a probability measure,

$$\sup \left\{ \int_{\mathbb{R}^d} g(x) \, d\mu_t(x) : g \in C_c(\mathbb{R}^d) \text{ and } 0 \leq g \leq 1 \right\} = 1. \quad (4.17)$$

The claim then follows from (4.16) and (4.17). ■

By monotone convergence theorem,

$$\begin{aligned}
\int_{B^c(0,R)} |x|^p \, d\mu_t(x) &= \lim_{K \rightarrow \infty} \int_{B^c(0,R) \cap B(0,K)} |x|^p \, d\mu_t(x) \\
&= \lim_{K \rightarrow \infty} \lim_{n \rightarrow \infty} \int_{B^c(0,R) \cap B(0,K)} |x|^p \, d\mu_t^n(x) \quad \text{by (4.13)} \\
&\lesssim \phi(R) \quad \text{by (4.12)}.
\end{aligned}$$

Then

$$\sup_{t \in [0,T]} \int_{B^c(0,R)} |x|^p \, d\mu_t(x) \lesssim \phi(R). \quad (4.18)$$

Clearly, $x \mapsto |x|^p$ is continuous and bounded from below. By Lemma 4.4 and Portmanteau's theorem,

$$\begin{aligned}
\sup_{t \in [0,T]} M_p(\mu_t) &\leq \sup_{t \in [0,T]} \liminf_n M_p(\mu_t^n) \\
&\lesssim 1 \quad \text{by (4.7)}.
\end{aligned}$$

Then $\mu_t \in \mathcal{P}_p(\mathbb{R}^d)$ for every $t \in [0, T]$. We have

$$W_p(\mu_s, \mu_t) \leq \liminf_n W_p(\mu_s^n, \mu_t^n) \quad (4.19)$$

$$\lesssim |t - s|^\delta. \quad (4.20)$$

Above, (4.19) is due to the lower semicontinuity of W_p w.r.t weak convergence; (4.20) is due to (4.8). Next we establish an essential result about convergence:

Lemma 4.5. *We have for every $R > T^{-1}$ that*

$$\lim_n \sup_{t \in [R^{-1}, T]} W_p(\mu_t^n, \mu_t) = 0.$$

Proof. We have

$$\begin{aligned}
W_p^p(\mu_t^n, \mu_t) &\lesssim \int_{\mathbb{R}^d} |x|^p \times |\ell_t^n(x) - \ell_t(x)| \, dx \quad \text{by Lemma 3.2} \\
&\leq \int_{B(0,k)} |x|^p \times |\ell_t^n(x) - \ell_t(x)| \, dx + \int_{B^c(0,k)} |x|^p \{ \ell_t^n(x) + \ell_t(x) \} \, dx \\
&=: I(t, n, k) + J(t, n, k) \quad \text{for every } k > 0.
\end{aligned}$$

By (4.13),

$$\lim_n \sup_{t \in [R^{-1}, T]} I(t, n, k) = 0.$$

By (4.12) and (4.18),

$$\sup_{n \in \mathbb{N}} \sup_{t \in [0, T]} J(t, n, k) \lesssim \phi(k).$$

As such,

$$\begin{aligned} \limsup_n \sup_{t \in [R^{-1}, T]} W_p^p(\mu_t^n, \mu_t) &\lesssim \limsup_n \sup_{t \in [R^{-1}, T]} I(t, n, k) \\ &\quad + \limsup_n \sup_{t \in [0, T]} J(t, n, k) \\ &\lesssim \phi(k). \end{aligned}$$

The claim then follows by taking the limit $k \rightarrow \infty$. ■

4.3. Existence of a weak solution. Notice that $\rho^n * \mu_t^n = \rho^n * \ell_t^n$. The Fokker-Planck equation (in distributional sense) associated with (4.1) is

$$\partial_t \ell_t^n(x) = -\partial_{x_i} \{b(t, x, \{\rho^n * \ell_t^n\}(x), \mu_t^n) \ell_t^n(x)\} + \frac{1}{2} \partial_{x_i} \partial_{x_j} \{a^{i,j}(t, x, \mu_t^n) \ell_t^n(x)\}.$$

This means for each $(\varphi, \psi) \in C_c^\infty(0, T) \times C_c^\infty(\mathbb{R}^d)$ that

$$\begin{aligned} & - \int_{[0, T]} \int_{\mathbb{R}^d} \varphi'(t) \psi(x) d\mu_t^n(x) dt \\ &= \sum_{i=1}^d \int_{[0, T]} \int_{\mathbb{R}^d} b(t, x, \{\rho^n * \ell_t^n\}(x), \mu_t^n) \varphi(t) \partial_{x_i} \psi(x) d\mu_t^n(x) dt \\ & \quad + \frac{1}{2} \sum_{i,j=1}^d \int_{[0, T]} \int_{\mathbb{R}^d} a^{i,j}(t, x, \mu_t^n) \varphi(t) \partial_{x_i} \partial_{x_j} \psi(x) d\mu_t^n(x) dt. \end{aligned} \tag{4.21}$$

Above, $a^{i,j}$ is the entry in the i -th row and j -th column of a . We recall from Lemma 4.5 and (4.13) that

$$W_p(\mu_t^n, \mu_t) \xrightarrow{n \rightarrow \infty} 0, \tag{4.22}$$

$$\sup_{x \in B(0, R)} |\ell_t^n(x) - \ell_t(x)| \xrightarrow{n \rightarrow \infty} 0 \quad \text{for every } (t, R) \in [0, T] \times \mathbb{R}_+. \tag{4.23}$$

We fix $(\varphi, \psi) \in C_c^\infty(0, T) \times C_c^\infty(\mathbb{R}^d)$. By (4.22), the boundedness of a , and the continuity of a w.r.t distribution variable,

$$\int_{\mathbb{R}^d} a^{i,j}(t, x, \mu_t^n) \partial_{x_i} \partial_{x_j} \psi(x) d\mu_t^n(x) \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}^d} a^{i,j}(t, x, \mu_t) \partial_{x_i} \partial_{x_j} \psi(x) d\mu_t(x).$$

Let $S := B(0, 1) + \text{supp } \psi$. Then S is bounded. By triangle inequality,

$$\begin{aligned} \|1_S\{\rho^n * \ell_t^n\} - \ell_t\|_\infty &\leq \|1_S\{\rho^n * [\ell_t^n - \ell_t]\}\|_\infty + \|1_S\{\rho^n * \ell_t - \ell_t\}\|_\infty \\ &\leq \|\rho^n * \{1_S[\ell_t^n - \ell_t]\}\|_\infty + \|1_S\{\rho^n * \ell_t - \ell_t\}\|_\infty \\ &\leq \|1_S[\ell_t^n - \ell_t]\|_\infty + \|1_S\{\rho^n * \ell_t - \ell_t\}\|_\infty. \end{aligned}$$

By (4.23), $\|1_S[\ell_t^n - \ell_t]\|_\infty \rightarrow 0$ as $n \rightarrow \infty$. By [Bre11, Proposition 4.21], $\|1_S\{\rho^n * \ell_t - \ell_t\}\|_\infty \rightarrow 0$ as $n \rightarrow \infty$. It follows that $\|1_S\{\rho^n * \ell_t^n\} - \ell_t\|_\infty \rightarrow 0$ as $n \rightarrow \infty$. This, together with (4.22) and Assumption 2.1(5), implies

$$\sup_{x \in S} |b(t, x, \{\rho^n * \ell_t^n\}(x), \mu_t^n) - b(t, x, \ell_t(x), \mu_t)| \xrightarrow{n \rightarrow \infty} 0. \tag{4.24}$$

Recall that $|b| \leq f_0$ and $f_0 \in \tilde{L}_{q_0}^{p_0}$. It follows from (4.22), (4.24) and dominated convergence theorem that

$$\int_{\mathbb{R}^d} b(t, x, \{\rho^n * \ell_t^n\}(x), \mu_t^n) \partial_{x_i} \psi(x) d\mu_t^n(x) \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}^d} b(t, x, \ell_t(x), \mu_t) \partial_{x_i} \psi(x) d\mu_t(x).$$

Taking the limit $n \rightarrow \infty$ in (4.21), we get

$$\begin{aligned} & - \int_{[0, T]} \int_{\mathbb{R}^d} \varphi'(t) \psi(x) d\mu_t(x) dt \\ &= \sum_{i=1}^d \int_{[0, T]} \int_{\mathbb{R}^d} b(t, x, \ell_t(x), \mu_t) \varphi(t) \partial_{x_i} \psi(x) d\mu_t(x) dt \\ & \quad + \frac{1}{2} \sum_{i,j=1}^d \int_{[0, T]} \int_{\mathbb{R}^d} a^{i,j}(t, x, \mu_t) \varphi(t) \partial_{x_i} \partial_{x_j} \psi(x) d\mu_t(x) dt. \end{aligned}$$

So ℓ satisfies the Fokker-Planck equation

$$\partial_t \ell_t(x) = -\partial_{x_i} \{b(t, x, \ell_t(x), \mu_t) \ell_t(x)\} + \frac{1}{2} \partial_{x_i} \partial_{x_j} \{a^{i,j}(t, x, \mu_t) \ell_t(x)\}.$$

Moreover, ℓ satisfies the following integrability estimate:

Lemma 4.6. *There exists a constant $c > 0$ (depending on Θ_1) such that*

$$\int_{[0, T]} \int_{\mathbb{R}^d} \{|b(t, x, \ell_t(x), \mu_t)| + |a(t, x, \mu_t)|\} d\mu_t(x) dt \leq c(1 + \|f_0\|_{\tilde{L}_{q_0}^{p_0}}).$$

Proof. By (4.6),

$$dX_t^n = \bar{b}^n(t, X_t^n) dt + \bar{\sigma}^n(t, X_t^n) dB_t.$$

Recall that Assumption 3.7 is a special case of Assumption 3.3. By Lemma 4.3, all pairs $(\bar{b}^n, \bar{\sigma}^n)_{n \in \mathbb{N}}$ satisfy Assumption 3.3 for the same set of parameters. Then

$$\begin{aligned} \int_{[0, T]} \int_{\mathbb{R}^d} f_0(t, x) d\mu_t^n(x) dt &= \mathbb{E} \left[\int_{[0, T]} f_0(t, X_t^n) dt \right] \quad \text{by Tonelli's theorem} \\ &\lesssim 1 + \|f_0\|_{\tilde{L}_{q_0}^{p_0}} \quad \text{by Lemma 3.5(2)}. \end{aligned} \tag{4.25}$$

We have

$$\int_{[0, T]} \int_{\mathbb{R}^d} f_0(t, x) d\mu_t(x) dt = \int_{[0, T]} \lim_k \int_{\mathbb{R}^d} 1_{B(0, k)}(x) f_0(t, x) \ell_t(x) dx dt \tag{4.26}$$

$$\leq \liminf_k \int_{[0, T]} \int_{\mathbb{R}^d} 1_{B(0, k)}(x) f_0(t, x) \ell_t(x) dx dt \tag{4.27}$$

$$= \liminf_k \int_{[0, T]} \lim_n \int_{\mathbb{R}^d} 1_{B(0, k)}(x) f_0(t, x) \ell_t^n(x) dx dt \tag{4.28}$$

$$\leq \liminf_k \liminf_n \int_{[0, T]} \int_{\mathbb{R}^d} 1_{B(0, k)}(x) f_0(t, x) \ell_t^n(x) dx dt \tag{4.29}$$

$$\begin{aligned} &\leq \liminf_n \int_{[0, T]} \int_{\mathbb{R}^d} f_0(t, x) \ell_t^n(x) dx dt \\ &\lesssim 1 + \|f_0\|_{\tilde{L}_{q_0}^{p_0}} \quad \text{by (4.25)}. \end{aligned} \tag{4.30}$$

Above, (4.26) is due to monotone convergence theorem; (4.27) and (4.29) are due to Fatou's lemma. We will justify how (4.28) follows from $f_0 \in \tilde{L}_{q_0}^{p_0}$ and dominated convergence theorem:

- (1) From (4.9), we get $1_{B(0, k)}(x) f_0(t, x) \ell_t^n(x) \lesssim 1_{B(0, k)}(x) f_0(t, x)$.

- (2) From (4.23), we get $1_{B(0,k)}(x)f_0(t,x)\ell_t^n(x) \rightarrow 1_{B(0,k)}(x)f_0(t,x)\ell_t(x)$ (as $n \rightarrow \infty$) for every $x \in \mathbb{R}^d$.

We denote by I the left-hand side of the inequality in the statement of Lemma 4.6. Then

$$\begin{aligned} I &\lesssim 1 + \int_{[0,T]} \int_{\mathbb{R}^d} f_0(t,x) d\mu_t(x) dt \\ &\lesssim 1 + \|f_0\|_{\tilde{L}_{q_0}^{p_0}} \quad \text{by (4.30).} \end{aligned}$$

This completes the proof. ■

We have

- (1) The maps $(t,x) \mapsto b(t,x,\ell_t(x),\mu_t)$ and $(t,x) \mapsto a(t,x,\mu_t)$ are measurable.

- (2) By Lemma 4.6,

$$\int_{[0,T]} \int_{\mathbb{R}^d} \{|b(t,x,\ell_t(x),\mu_t)| + |a(t,x,\mu_t)|\} d\mu_t(x) dt < \infty.$$

- (3) By (4.20), the map $[0,T] \rightarrow \mathcal{P}_p(\mathbb{R}^d)$, $t \mapsto \mu_t$ is continuous.

By superposition principle [Fig08; Tre16; BRS21] as in [BR20, Section 2], the SDE (1.1) has a weak solution whose marginal distribution is exactly (μ_t) .

4.4. Existence of a strong solution. By the previous subsection, there exists a PS $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$ on which there exist a d -BM $(\bar{B}_t)$, an AF $(\bar{\mathcal{F}}_t)$ and a continuous $(\bar{\mathcal{F}}_t)$ -adapted process $(\bar{X}_t)$ such that

$$\begin{cases} d\bar{X}_t = b(t, \bar{X}_t, \ell_t(\bar{X}_t), \mu_t) dt + \sigma(t, \bar{X}_t, \mu_t) d\bar{B}_t, \\ \nu = \mathcal{L}(\bar{X}_0), \quad \mu_t = \mathcal{L}(\bar{X}_t), \quad \mu_t(dx) = \ell_t(x) dx. \end{cases} \quad (4.31)$$

We define the map

$$\bar{b} : [0,T] \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d) \rightarrow \mathbb{R}^d$$

by

$$\bar{b}(t, x, \varrho) := b(t, x, \ell_t(x), \varrho).$$

Then

$$\begin{cases} d\bar{X}_t = \bar{b}(t, \bar{X}_t, \mu_t) dt + \sigma(t, \bar{X}_t, \mu_t) d\bar{B}_t, \\ \nu = \mathcal{L}(\bar{X}_0) \quad \text{and} \quad \mu_t = \mathcal{L}(\bar{X}_t). \end{cases} \quad (4.32)$$

We consider the McKean-Vlasov SDE

$$\begin{cases} dY_t = \bar{b}(t, Y_t, \bar{\mu}_t) dt + \sigma(t, Y_t, \bar{\mu}_t) dB_t, \\ \nu = \mathcal{L}(Y_0) \quad \text{and} \quad \bar{\mu}_t = \mathcal{L}(Y_t), \end{cases} \quad (4.33)$$

where (B_t) is the fixed d -BM on the fixed PS $(\Omega, \mathcal{A}, \mathbb{P})$ introduced in Section 1.

Let us verify that $(\bar{b}, \sigma)$ satisfies the conditions of Corollary 3.13. By Assumption 2.1[(1) and (3)], $a_t(x, \varrho)$ is invertible and $x \mapsto \sigma_t(x, \varrho)$ is weakly differentiable with $|\nabla \sigma_t(x, \varrho)| \leq \sum_{i=1}^l f_i(t, x)$. By Assumption 2.1(2), we have $|\bar{b}(t, x, \varrho)| = |b(t, x, \ell_t(x), \varrho)| \leq C f_0(t, x)$. By Assumption 2.1(5), since $\ell_t(x)$ is fixed,

$$|\bar{b}(t, x, \varrho) - \bar{b}(t, x, \varrho')| = |b(t, x, \ell_t(x), \varrho) - b(t, x, \ell_t(x), \varrho')| \leq C W_p(\varrho, \varrho').$$

The remaining requirement $|\sigma_t(x, \varrho) - \sigma_t(y, \varrho')| \leq C\{|x - y|^\beta + W_p(\varrho, \varrho')\}$ is given by Assumption 2.1(5).

As such, the SDE (4.33) is both strongly and weakly well-posed by Corollary 3.13. By strong well-posedness of (4.33), it has a strong solution (Y_t) on the fixed PS $(\Omega, \mathcal{A}, \mathbb{P})$ driven by the fixed d -BM (B_t) .

By (4.32), $(\bar{X}_t)$ is a weak solution of (4.33) on the PS $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$. By weak well-posedness of (4.33), we have $\mu_t = \bar{\mu}_t$ for every $t \in [0, T]$. Thus $\mathcal{L}(Y_t) = \mu_t$ and

$$dY_t = b(t, Y_t, \ell_t(Y_t), \mu_t) dt + \sigma(t, Y_t, \mu_t) dB_t,$$

Recall from (4.31) that $\mu_t(dx) = \ell_t(x) dx$. As such, (Y_t) is a strong solution to the SDE (1.1) on the fixed stochastic basis. This completes the proof.

5. PROOF OF THEOREM 2.3

For $k \in \{1, 2\}$, we consider the SDE

$$\begin{cases} dX_t^k = b(t, X_t^k, \ell_t^k(X_t^k), \mu_t^k) dt + \sigma(t, X_t^k) dB_t^k, \\ \nu_k = \mathcal{L}(X_0^k), \quad \mu_t^k = \mathcal{L}(X_t^k), \quad \mu_t^k(dx) = \ell_t^k(x) dx. \end{cases} \quad (5.1)$$

Above, $(B_t^k, t \geq 0)$ is a d -BM. We define measurable maps

$$b^k : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$$

by

$$b^k(t, x) := b(t, x, \ell_t^k(x), \mu_t^k).$$

5.1. Uniqueness of marginal density. Clearly, (b^k, σ) satisfies Assumption 3.7. We denote $b_t^k(x) := b^k(t, x)$. By Lemma 3.10,

$$\ell_t^k(x) = \int_{\mathbb{R}^d} p_{0,t}^{0,\sigma}(y, x) \ell_{\nu_k}(y) dy + \int_0^t \int_{\mathbb{R}^d} \ell_s^k(y) \langle b_s^k(y), \nabla_y p_{s,t}^{0,\sigma}(y, x) \rangle dy ds.$$

Then

$$\begin{aligned} |\ell_t^2(x) - \ell_t^1(x)| &\leq \int_{\mathbb{R}^d} p_{0,t}^{0,\sigma}(y, x) |\ell_{\nu_1}(y) - \ell_{\nu_2}(y)| dy \\ &\quad + \int_0^t \int_{\mathbb{R}^d} |b_s^2(y)| \times |\ell_s^2(y) - \ell_s^1(y)| \times |\nabla_y p_{s,t}^{0,\sigma}(y, x)| dy ds \\ &\quad + \int_0^t \int_{\mathbb{R}^d} \ell_s^1(y) |b_s^2(y) - b_s^1(y)| \times |\nabla_y p_{s,t}^{0,\sigma}(y, x)| dy ds. \end{aligned}$$

We write $M_1 \preccurlyeq M_2$ if there exists a constant $c > 0$ (depending on Θ_1) such that $M_1 \preccurlyeq cM_2$. Thus

$$\begin{aligned} |\ell_t^2(x) - \ell_t^1(x)| &\preccurlyeq \int_{\mathbb{R}^d} p_{0,t}^{0,\sigma}(y, x) |\ell_{\nu_1}(y) - \ell_{\nu_2}(y)| dy \\ &\quad + \int_0^t \int_{\mathbb{R}^d} |\ell_s^2(y) - \ell_s^1(y)| \times |\nabla_y p_{s,t}^{0,\sigma}(y, x)| dy ds \\ &\quad + \int_0^t \int_{\mathbb{R}^d} \ell_s^1(y) |b_s^2(y) - b_s^1(y)| \times |\nabla_y p_{s,t}^{0,\sigma}(y, x)| dy ds. \end{aligned} \quad (5.2)$$

By Corollary 3.9(1),

$$\sup_{t \in [0, T]} \|\ell_t^1\|_\infty \preccurlyeq \|\ell_{\nu_1}\|_\infty. \quad (5.3)$$

By Assumption 2.1(5),

$$|b_s^2(y) - b_s^1(y)| \preccurlyeq |\ell_s^2(y) - \ell_s^1(y)| + W_p(\mu_s^2, \mu_s^1). \quad (5.4)$$

By (5.2), (5.3) and (5.4),

$$\begin{aligned}
|\ell_t^2(x) - \ell_t^1(x)| &\leq \int_{\mathbb{R}^d} p_{0,t}^{0,\sigma}(y, x) |\ell_{\nu_1}(y) - \ell_{\nu_2}(y)| dy \\
&\quad + (1 + \|\ell_{\nu_1}\|_\infty) \int_0^t \int_{\mathbb{R}^d} |\ell_s^2(y) - \ell_s^1(y)| \times |\nabla_y p_{s,t}^{0,\sigma}(y, x)| dy ds \\
&\quad + \int_0^t W_p(\mu_s^2, \mu_s^1) \int_{\mathbb{R}^d} \ell_s^1(y) |\nabla_y p_{s,t}^{0,\sigma}(y, x)| dy ds \\
&=: I_1(t, x) + (1 + \|\ell_{\nu_1}\|_\infty) I_2(t, x) + I_3(t, x).
\end{aligned}$$

The pair $(0, \sigma)$ satisfies [Assumption 3.7](#). By [Lemma 3.8](#), there exists a constant $\lambda > 0$ (depending on Θ_1) such that for $i \in \{0, 1\}$, $0 \leq s < t \leq T$ and $x, y \in \mathbb{R}^d$:

$$|\nabla_y^i p_{s,t}^{0,\sigma}(y, x)| \leq p_{t-s}^{i,\lambda}(y - x). \quad (5.5)$$

Then

$$\begin{aligned}
\int_{\mathbb{R}^d} (|x|^p + 1) |\nabla_y^i p_{s,t}^{0,\sigma}(y, x)| dx &\leq \int_{\mathbb{R}^d} (|x|^p + 1) p_{t-s}^{i,\lambda}(y - x) dx \quad \text{by (5.5)} \\
&\leq (t - s)^{-\frac{i}{2}} (|y|^p + 1). \quad (5.6)
\end{aligned}$$

We define a measurable map $f : [0, T] \rightarrow \mathbb{R}_+$ by

$$f(s) := \int_{\mathbb{R}^d} (|x|^p + 1) |\ell_s^2(x) - \ell_s^1(x)| dx.$$

By (3.8), f is bounded. First,

$$\begin{aligned}
&\int_{\mathbb{R}^d} (|x|^p + 1) I_1(t, x) dx \\
&= \int_0^t \int_{\mathbb{R}^d} |\ell_{\nu_1}(y) - \ell_{\nu_2}(y)| \int_{\mathbb{R}^d} (|x|^p + 1) p_{0,t}^{0,\sigma}(y, x) dx dy ds \\
&\leq \int_0^t \int_{\mathbb{R}^d} |\ell_{\nu_1}(y) - \ell_{\nu_2}(y)| (|y|^p + 1) dy ds \quad \text{by (5.6)} \\
&\leq \int_{\mathbb{R}^d} (|y|^p + 1) |\ell_{\nu_1}(y) - \ell_{\nu_2}(y)| dy = f(0).
\end{aligned}$$

Second,

$$\begin{aligned}
&\int_{\mathbb{R}^d} (|x|^p + 1) I_2(t, x) dx \\
&= \int_0^t \int_{\mathbb{R}^d} |\ell_s^2(y) - \ell_s^1(y)| \int_{\mathbb{R}^d} (|x|^p + 1) |\nabla_y p_{s,t}^{0,\sigma}(y, x)| dx dy ds \\
&\leq \int_0^t (t - s)^{-\frac{1}{2}} \int_{\mathbb{R}^d} |\ell_s^2(y) - \ell_s^1(y)| (|y|^p + 1) dy ds \quad \text{by (5.6)} \\
&= \int_0^t (t - s)^{-\frac{1}{2}} f(s) ds.
\end{aligned}$$

Third,

$$\begin{aligned}
&\int_{\mathbb{R}^d} (|x|^p + 1) I_3(t, x) dx \\
&= \int_0^t W_p(\mu_s^2, \mu_s^1) \int_{\mathbb{R}^d} \ell_s^1(y) \int_{\mathbb{R}^d} (|x|^p + 1) |\nabla_y p_{s,t}^{0,\sigma}(y, x)| dx dy ds \\
&\leq \int_0^t (t - s)^{-\frac{1}{2}} W_p(\mu_s^2, \mu_s^1) \int_{\mathbb{R}^d} \ell_s^1(y) (|y|^p + 1) dy ds \quad \text{by (5.6)}
\end{aligned}$$

$$\begin{aligned}
&\preceq (1 + M_p(\nu_1)) \int_0^t (t-s)^{-\frac{1}{2}} W_p(\mu_s^2, \mu_s^1) ds \quad \text{by (3.8)} \\
&\preceq (1 + M_p(\nu_1)) \int_0^t (t-s)^{-\frac{1}{2}} |f(s)|^{\frac{1}{p}} ds \quad \text{by Lemma 3.2.}
\end{aligned}$$

To sum up,

$$f(t) \preceq f(0) + (1 + \|\ell_{\nu_1}\|_\infty + M_p(\nu_1)) \int_0^t (T-s)^{-\frac{1}{2}} (f(s) + |f(s)|^{\frac{1}{p}}) ds.$$

Because $p = 1$, we get

$$f(t) \preceq f(0) + (1 + \|\ell_{\nu_1}\|_\infty + M_1(\nu_1)) \int_0^t (T-s)^{-\frac{1}{2}} f(s) ds.$$

By Gronwall's lemma,

$$\sup_{t \in [0, T]} f(t) \preceq f(0) \exp \{2\sqrt{T}(1 + \|\ell_{\nu_1}\|_\infty + M_1(\nu_1))\}. \quad (5.7)$$

This implies the existence of the function Λ as required in Theorem 2.3(1).

5.2. Weak and strong uniqueness of a solution. Now we assume $\nu_1 = \nu_2 =: \nu$. By (5.7), $\ell_t^1 = \ell_t^2 =: \ell_t$ and $\mu_t^1 = \mu_t^2 =: \mu_t$ for $t \in [0, T]$. Then $b^1 = b^2 =: \mathbf{b}$. By (5.1),

$$dX_t^k = \mathbf{b}(t, X_t^k) dt + \sigma(t, X_t^k) dB_t^k.$$

We consider the SDE

$$\begin{cases} dY_t = \mathbf{b}(t, Y_t) dt + \sigma(t, Y_t) dB_t, \\ \nu = \mathcal{L}(Y_0). \end{cases} \quad (5.8)$$

Let us verify that $(\mathbf{b}, \sigma)$ satisfies the conditions of Corollary 3.13. By Assumption 2.1(2), the drift $\mathbf{b}$ satisfies

$$|\mathbf{b}(t, x)| = |b(t, x, \ell_t(x), \mu_t)| \leq C f_0(t, x).$$

The remaining requirements on σ by Corollary 3.13 are satisfied by Assumption 2.1[(1), (3) and (5)]. We thus conclude from Corollary 3.13 that the SDE (5.8) is both strongly and weakly well-posed. We note that $(\mathbf{b}, \sigma)$ is distribution-free, so (5.8) is a classical SDE; Corollary 3.13 applies with trivial distribution dependence.

Weak uniqueness. Both (X_t^1) and (X_t^2) are weak solutions of (5.8) with the same initial distribution ν . By weak uniqueness of (5.8), the SDE (1.1) has weak uniqueness.

Strong uniqueness. Assume now that $(B_t^1) = (B_t^2) = (B_t)$ (the same prescribed d -BM) and that $X_0^1 = X_0^2 = X_0$ $\mathbb{P}$ -a.s. Then (X_t^1) and (X_t^2) are both strong solutions of (5.8) on the same stochastic basis $(\Omega, \mathcal{A}, \mathbb{P}, (\mathcal{F}_t), (B_t))$ with the same initial condition. By strong well-posedness of (5.8), we conclude $X_t^1 = X_t^2$ for all $t \in [0, T]$, $\mathbb{P}$ -a.s. Hence the SDE (1.1) has strong uniqueness.

6. ACKNOWLEDGMENT

Anh-Dung Le is grateful to Professor Sébastien Gadat for his generous PhD funding, and to Professor Feng-Yu Wang for kindly addressing all questions about [Wan23]. The authors thank Professor Michael Röckner, Dr. Zimo Hao and Professor Francesco Russo for their useful comments and suggestions.

REFERENCES

- [ABS21] Luigi Ambrosio, Elia Brué, and Daniele Semola. *Lectures on optimal transport*. Springer, 2021.
- [AG13] Luigi Ambrosio and Nicola Gigli. “A User’s Guide to Optimal Transport”. In: *Modelling and Optimisation of Flows on Networks: Cetraro, Italy 2009, Editors: Benedetto Piccoli, Michel Rasle*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 1–155.
- [Bar82] M. T. Barlow. “One Dimensional Stochastic Differential Equations with No Strong Solution”. In: *Journal of the London Mathematical Society* s2-26.2 (Oct. 1982), pp. 335–347.
- [BCR13] Nadia Belaribi, François Cuvelier, and Francesco Russo. “Probabilistic and deterministic algorithms for space multidimensional irregular porous media equation”. en. In: *Stochastic Partial Differential Equations: Analysis and Computations* 1.1 (Mar. 2013), pp. 3–62.
- [BJ19] Mireille Bossy and Jean-François Jabir. “On the Wellposedness of Some McKean Models with Moderated or Singular Diffusion Coefficient”. en. In: *Frontiers in Stochastic Analysis–BSDEs, SPDEs and their Applications*. Ed. by Samuel N. Cohen et al. Vol. 289. Series Title: Springer Proceedings in Mathematics & Statistics. Cham: Springer International Publishing, 2019, pp. 43–87.
- [BR12] Nadia Belaribi and Francesco Russo. “Uniqueness for Fokker-Planck equations with measurable coefficients and applications to the fast diffusion equation”. In: *Electronic Journal of Probability* 17.none (Jan. 2012).
- [BR18] Viorel Barbu and Michael Röckner. “Probabilistic Representation for Solutions to Nonlinear Fokker-Planck Equations”. In: *SIAM Journal on Mathematical Analysis* 50.4 (2018), pp. 4246–4260.
- [BR20] Viorel Barbu and Michael Röckner. “From nonlinear Fokker-Planck equations to solutions of distribution dependent SDE”. In: *The Annals of Probability* 48.4 (2020), pp. 1902–1920.
- [BR21a] Viorel Barbu and Michael Röckner. “Solutions for nonlinear Fokker-Planck equations with measures as initial data and McKean-Vlasov equations”. In: *Journal of Functional Analysis* 280.7 (2021), p. 108926.
- [BR21b] Viorel Barbu and Michael Röckner. “Uniqueness for nonlinear Fokker-Planck equations and weak uniqueness for McKean-Vlasov SDEs”. In: *Stochastics and Partial Differential Equations: Analysis and Computations* 9 (2021), pp. 702–713.
- [BR23a] Viorel Barbu and Michael Röckner. “Nonlinear Fokker-Planck equations with time-dependent coefficients”. In: *SIAM Journal on Mathematical Analysis* 55.1 (2023), pp. 1–18.
- [BR23b] Viorel Barbu and Michael Röckner. “Uniqueness for nonlinear Fokker-Planck equations and for McKean-Vlasov SDEs: The degenerate case”. en. In: *Journal of Functional Analysis* 285.4 (Aug. 2023), p. 109980.
- [BR23c] Viorel Barbu and Michael Roeckner. “The evolution to equilibrium of solutions to nonlinear Fokker-Planck equation”. en. In: *Indiana University Mathematics Journal* 72.1 (2023), pp. 89–131.
- [Bre11] Haim Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. en. New York, NY: Springer New York, 2011.
- [BRR10] Philippe Blanchard, Michael Röckner, and Francesco Russo. “Probabilistic representation for solutions of an irregular porous media type equation”. In: *The Annals of Probability* 38.5 (Sept. 2010).
- [BRR11] Viorel Barbu, Michael Röckner, and Francesco Russo. “Probabilistic representation for solutions of an irregular porous media type equation: the degenerate case”. en. In: *Probability Theory and Related Fields* 151.1-2 (Oct. 2011), pp. 1–43.

- [BRS21] Vladimir I. Bogachev, Michael Röckner, and Stanislav V. Shaposhnikov. “On the Ambrosio–Figalli–Trevisan Superposition Principle for Probability Solutions to Fokker–Planck–Kolmogorov Equations”. en. In: *Journal of Dynamics and Differential Equations* 33.2 (June 2021), pp. 715–739.
- [CD18] René Carmona and François Delarue. *Probabilistic theory of mean field games with applications I-II*. Springer, 2018.
- [CF22] Paul-Eric Chaudru De Raynal and Noufel Frikha. “Well-posedness for some non-linear SDEs and related PDE on the Wasserstein space”. en. In: *Journal de Mathématiques Pures et Appliquées* 159 (Mar. 2022), pp. 1–167.
- [Cha20] P.E. Chaudru De Raynal. “Strong well posedness of McKean–Vlasov stochastic differential equations with Hölder drift”. en. In: *Stochastic Processes and their Applications* 130.1 (Jan. 2020), pp. 79–107.
- [FG21] Alessio Figalli and Federico Glaudo. *An invitation to optimal transport, Wasserstein distances, and gradient flows*. eng. EMS Textbooks in Mathematics. Berlin: EMS Press, 2021.
- [Fig08] Alessio Figalli. “Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients”. en. In: *Journal of Functional Analysis* 254.1 (Jan. 2008), pp. 109–153.
- [Fun84] Tadahisa Funaki. “A certain class of diffusion processes associated with nonlinear parabolic equations”. In: *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 67.3 (1984), pp. 331–348.
- [Gru23] Sebastian Grube. “Strong solutions to McKean–Vlasov SDEs with coefficients of Nemytskii-type”. In: *Electronic Communications in Probability* 28.none (Jan. 2023).
- [Gru24] Sebastian Grube. “Strong solutions to McKean–Vlasov SDEs with coefficients of Nemytskii type: the time-dependent case”. en. In: *Journal of Evolution Equations* 24.2 (June 2024), p. 37.
- [HRW21] Xing Huang, Panpan Ren, and Feng-Yu Wang. “Distribution dependent stochastic differential equations”. en. In: *Frontiers of Mathematics in China* 16.2 (Apr. 2021), pp. 257–301.
- [HRZ23] Zimo Hao, Michael Röckner, and Xicheng Zhang. *Second order fractional mean-field SDEs with singular kernels and measure initial data*. arXiv:2302.04392 [math]. Mar. 2023.
- [HRZ24] Zimo Hao, Michael Röckner, and Xicheng Zhang. “Strong Convergence of Propagation of Chaos for McKean–Vlasov SDEs with Singular Interactions”. en. In: *SIAM Journal on Mathematical Analysis* 56.2 (Apr. 2024), pp. 2661–2713.
- [HW21] Xing Huang and Feng-Yu Wang. “McKean–Vlasov sdes with drifts discontinuous under wasserstein distance”. en. In: *Discrete & Continuous Dynamical Systems* 41.4 (2021), pp. 1667–1679.
- [HW22] Xing Huang and Feng-Yu Wang. “Singular McKean–Vlasov (reflecting) SDEs with distribution dependent noise”. In: *Journal of Mathematical Analysis and Applications* 514.1 (2022), p. 126301.
- [IOR19] Lucas Izydorczyk, Nadia Oudjane, and Francesco Russo. “McKean Feynman–Kac probabilistic representations of non-linear partial differential equations”. In: *International Conference on Random Transformations and Invariance in Stochastic Dynamics*. Springer. 2019, pp. 187–212.
- [IR23] Elena Issoglio and Francesco Russo. “McKean SDEs with singular coefficients”. In: *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques* 59.3 (Aug. 2023).
- [Iss+24] Elena Issoglio et al. *Degenerate McKean–Vlasov equations with drift in anisotropic negative Besov spaces*. arXiv:2401.09165 [math]. Jan. 2024.
- [JM98] Benjamin Jourdain and Sylvie Meleard. “Propagation of chaos and fluctuations for a moderate model with smooth initial data”. en. In: *Annales de l’Institut Henri Poincaré (B) Probability and Statistics* 34.6 (1998), pp. 727–766.

- [JR13] Benjamin Jourdain and Julien Reygner. “Propagation of chaos for rank-based interacting diffusions and long time behaviour of a scalar quasilinear parabolic equation”. en. In: *Stochastic Partial Differential Equations: Analysis and Computations* 1.3 (Sept. 2013), pp. 455–506.
- [Kac56] Mark Kac. “Foundations of kinetic theory”. In: *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume III*. Berkeley: University of California Press, 1956, pp. 173–200.
- [Kry08] Nikolaj Vladimirovič Krylov. *Controlled diffusion processes*. Vol. 14. Springer Science & Business Media, 2008.
- [Mag23] Francesco Maggi. *Optimal Mass Transport on Euclidean Spaces*. Cambridge University Press, 2023.
- [McK66] Henry P McKean Jr. “A class of Markov processes associated with nonlinear parabolic equations”. In: *Proceedings of the National Academy of Sciences* 56.6 (1966), pp. 1907–1911.
- [MPZ21] Stéphane Menozzi, Antonello Pesce, and Xicheng Zhang. “Density and gradient estimates for non degenerate Brownian SDEs with unbounded measurable drift”. In: *Journal of Differential Equations* 272 (2021), pp. 330–369.
- [MV20] Yuliya Mishura and Alexander Veretennikov. “Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations”. In: *Theory of Probability and Mathematical Statistics* 103 (2020), pp. 59–101.
- [Reh23] Marco Rehmeier. *Weighted L^1 -semigroup approach for nonlinear Fokker–Planck equations and generalized Ornstein–Uhlenbeck processes*. arXiv:2308.09420 [math]. Aug. 2023.
- [Rud87] Walter Rudin. *Real and complex analysis*. 3rd. New York: McGraw-Hill, 1987.
- [RZ21] Michael Röckner and Xicheng Zhang. “Well-posedness of distribution dependent SDEs with singular drifts”. In: *Bernoulli* 27.2 (2021), pp. 1131–1158.
- [San15] Filippo Santambrogio. “Optimal transport for applied mathematicians”. In: *Birkhäuser, NY* 55.58-63 (2015), p. 94.
- [Sch87] Michael Scheutzow. “Uniqueness and non-uniqueness of solutions of Vlasov-McKean equations”. In: *Journal of the Australian Mathematical Society* 43.2 (1987), pp. 246–256.
- [Szn84] Alain-Sol Sznitman. “Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated”. In: *Journal of functional analysis* 56.3 (1984), pp. 311–336.
- [Tho23] Matthew Thorpe. *Introduction to optimal transport*. 2023.
- [Tre16] Dario Trevisan. “Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients”. In: *Electronic Journal of Probability* 21.none (Jan. 2016).
- [Vil03] Cédric Villani. *Topics in Optimal Transportation*. Graduate studies in mathematics. American Mathematical Society, 2003.
- [Vil09] Cédric Villani. *Optimal transport: old and new*. Vol. 338. Springer, 2009.
- [Wan23] Feng-Yu Wang. “Singular density dependent stochastic differential equations”. In: *Journal of Differential Equations* 361 (2023), pp. 562–589.
- [XZ20] Longjie Xie and Xicheng Zhang. “Ergodicity of stochastic differential equations with jumps and singular coefficients”. In: *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques* 56.1 (2020), pp. 175–229.
- [Zha24] Guohuan Zhao. “Existence and Uniqueness for McKean-Vlasov Equations with Singular Interactions”. en. In: *Potential Analysis* (June 2024).
- [ZY21] Shao-Qin Zhang and Chenggui Yuan. “A Zvonkin’s transformation for stochastic differential equations with singular drift and applications”. In: *Journal of Differential Equations* 297 (2021), pp. 277–319.

7. APPENDIX

Proof of Lemma 3.2. For $B \in \mathcal{B}(\mathbb{R}^d)$, we denote by $\Pi(B)$ the collection of all finite measurable partitions of B . This means $(B_1, \dots, B_n) \in \Pi(B)$ if and only if $\{B_1, \dots, B_n\} \subset \mathcal{B}(\mathbb{R}^d)$ are pairwise disjoint and $B = \bigcup_{k=1}^n B_k$. We have

$$\begin{aligned} |\mu - \nu|(B) &= \sup \left\{ \sum_{k=1}^n |(\mu - \nu)(B_k)| : (B_1, \dots, B_n) \in \Pi(B) \right\} \\ &= \sup \left\{ \sum_{k=1}^n \left| \int_{B_k} (\ell_\mu - \ell_\nu)(x) dx \right| : (B_1, \dots, B_n) \in \Pi(B) \right\} \\ &\leq \sup \left\{ \sum_{k=1}^n \int_{B_k} |\ell_\mu - \ell_\nu|(x) dx : (B_1, \dots, B_n) \in \Pi(B) \right\} \\ &= \int_B |\ell_\mu - \ell_\nu|(x) dx. \end{aligned}$$

On the other hand, we have from [Vil03, Proposition 7.10] that

$$W_p^p(\mu, \nu) \leq (1 \vee 2^{p-1}) \int_{\mathbb{R}^d} |x|^p d|\mu - \nu|(x).$$

The claim then follows. ■

Proof of Lemma 3.10. First, we recall some notions related to (3.12). The semigroup $(P_{s,t}^{b,\sigma})_{0 \leq s < t \leq T}$ is defined for $x \in \mathbb{R}^d$ and bounded measurable function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ by

$$P_{s,t}^{b,\sigma} f(x) := \mathbb{E}[f(X_{s,t}^x)] = \int_{\mathbb{R}^d} p_{s,t}^{b,\sigma}(x, y) f(y) dy.$$

The differential operator $(L_t^{b,\sigma})_{t \in [0, T]}$ is defined for $f \in C^2(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$ by

$$L_t^{b,\sigma} f(x) := \langle b_t(x), \nabla f(x) \rangle + \frac{1}{2} \text{tr}(a_t(x) \nabla^2 f(x)).$$

The backward Kolmogorov equation holds, i.e., for $f \in C_b^2(\mathbb{R}^d)$, $x \in \mathbb{R}^d$ and $0 \leq s < t \leq T$:

$$\partial_s P_{s,t}^{b,\sigma} f(x) + L_s^{b,\sigma} P_{s,t}^{b,\sigma} f(x) = 0. \quad (7.1)$$

Next we go on to prove the result. Let $f \in C_c^\infty(\mathbb{R}^d)$. Applying Itô's lemma on $[0, t] \times \mathbb{R}^d \rightarrow \mathbb{R}$, $(s, x) \mapsto (P_{s,t}^{0,\sigma} f)(x)$, we have

$$\begin{aligned} d\{(P_{s,t}^{0,\sigma} f)(X_s)\} &= \{(\partial_s + L_s^{b,\sigma})(P_{s,t}^{b,\sigma} f)\}(X_s) ds + dM_s \\ &= \{(\partial_s + L_s^{0,\sigma}) + (L_s^{b,\sigma} - L_s^{0,\sigma})\}(P_{s,t}^{0,\sigma} f)(X_s) ds + dM_s \\ &= \{(L_s^{b,\sigma} - L_s^{0,\sigma})(P_{s,t}^{0,\sigma} f)\}(X_s) ds + dM_s. \end{aligned} \quad (7.2)$$

Above, $M_0 = 0$ and $dM_s = \{\nabla(P_{s,t}^{0,\sigma} f)(X_s)\}^\top \sigma(s, X_s) dB_s$ for $s \in [0, t]$; and (7.2) is due to backward Kolmogorov equation (7.1). Then

$$\begin{aligned} f(X_t) &= (P_{0,t}^{0,\sigma} f)(X_0) + \int_0^t \langle b_s(X_s), \nabla(P_{s,t}^{0,\sigma} f)(X_s) \rangle ds \\ &\quad + \int_0^t \{\nabla(P_{s,t}^{0,\sigma} f)(X_s)\}^\top \sigma(s, X_s) dB_s. \end{aligned} \quad (7.3)$$

Clearly, f and thus $P_{0,t}^{0,\sigma} f$ are bounded. Let's prove that $\|\nabla(P_{s,t}^{0,\sigma} f)\|_\infty < \infty$. It suffices to verify that $P_{s,t}^{0,\sigma} f$ is Lipschitz. We consider the SDE

$$dY_{s,t}^x = \sigma(t, Y_{s,t}^x) dB_s, \quad t \in [s, T], Y_{s,s}^x = x.$$

By [HW22, Theorem 1.1(2)], there exists a constant $c_1 > 0$ such that

$$\begin{aligned} |P_{s,t}^{0,\sigma} f(x) - P_{s,t}^{0,\sigma} f(y)| &= |\mathbb{E}[f(Y_{s,t}^x)] - \mathbb{E}[f(Y_{s,t}^y)]| \\ &\leq \|\nabla f\|_\infty \mathbb{E}[|Y_{s,t}^x - Y_{s,t}^y|] \\ &\leq c_1 \|\nabla f\|_\infty |x - y|. \end{aligned}$$

We have

$$\begin{aligned} \mathbb{E} \left[\int_0^t |b_s(X_s)| \times |\nabla(P_{s,t}^{0,\sigma} f)(X_s)| \, ds \right] &\leq \|\nabla(P_{s,t}^{0,\sigma} f)\|_\infty \mathbb{E} \left[\int_0^t g(t, X_s) \, ds \right] \\ &\lesssim \|\nabla(P_{s,t}^{0,\sigma} f)\|_\infty \|g\|_{\tilde{L}_q^{\bar{p}}(t)} \quad \text{by Lemma 3.5(2)}. \end{aligned}$$

So each term in (7.3) is $\mathbb{P}$ -integrable. Then

$$\begin{aligned} \int_{\mathbb{R}^d} \ell_t(x) f(x) \, dx &= \int_{\mathbb{R}^d} \ell_\nu(x) (P_{0,t}^{0,\sigma} f)(x) \, dx \\ &\quad + \int_0^t \int_{\mathbb{R}^d} \ell_s(x) \langle b_s(x), \nabla(P_{s,t}^{0,\sigma} f)(x) \rangle \, dx \, ds. \end{aligned} \tag{7.4}$$

By Lemma 3.8(1) and Leibniz integral rule,

$$\nabla(P_{s,t}^{0,\sigma} f)(x) = \nabla_x \int_{\mathbb{R}^d} p_{s,t}^{0,\sigma}(x, y) f(y) \, dy = \int_{\mathbb{R}^d} \nabla_x p_{s,t}^{0,\sigma}(x, y) f(y) \, dy.$$

So (7.4) is equivalent to

$$\begin{aligned} \int_{\mathbb{R}^d} \ell_t(x) f(x) \, dx &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} p_{0,t}^{0,\sigma}(y, x) \ell_\nu(y) \, dy \right) f(x) \, dx \\ &\quad + \int_{\mathbb{R}^d} \left(\int_0^t \int_{\mathbb{R}^d} \ell_s(y) \langle b_s(y), \nabla_y p_{s,t}^{0,\sigma}(y, x) \rangle \, dy \, ds \right) f(x) \, dx. \end{aligned}$$

The required representation then follows. ■

Proof of Lemma 3.11. WLOG, we consider $R > 0$. We write $M_1 \lesssim M_2$ if there exists a constant $c > 0$ depending on (Θ_3, ν, p) such that $M_1 \leq cM_2$. We denote by ℓ_t the density of X_t . Then

$$\ell_t(x) = \int_{\mathbb{R}^d} p_{0,t}^{b,\sigma}(y, x) \, d\nu(y).$$

By Lemma 3.8(1), there exists a constant $\lambda \in (0, 1)$ depending on Θ_3 such that $p_{0,t}^{b,\sigma}(y, x) \lesssim p_t^{0,\lambda}(y - x)$. Then $\ell_t(x) \lesssim p_t^{0,\lambda} * \nu(x)$. Let Z be a standard normal random variable on $\mathbb{R}^d$. Let Y be a random variable on $\mathbb{R}^d$, independent of Z , with distribution ν . Let $c_t := \sqrt{\frac{t}{2\lambda}}$ and $s := \frac{1}{2}$. Then

$$\begin{aligned} \int_{B^c(0,R)} |x|^p \, d\mu_t(x) &\lesssim \mathbb{E}[1_{\{|c_t Z + Y| > R\}} |c_t Z + Y|^p] \\ &\lesssim \mathbb{E}[(1_{\{|c_t Z| > sR\}} + 1_{\{|Y| > (1-s)R\}})(|c_t|^p |Z|^p + |Y|^p)] \\ &\lesssim \mathbb{E}[1_{\{|Z| > \frac{sR}{c_t}\}} |Z|^p] + \mathbb{E}[1_{\{|Y| > \frac{sR}{c_t}\}} |Y|^p] \\ &\quad + \mathbb{E}[1_{\{|Y| > (1-s)R\}} |Z|^p] + \mathbb{E}[1_{\{|Y| > (1-s)R\}} |Y|^p] \\ &=: I_1(R) + I_2(R) + I_3(R) + I_4(R) \\ &=: \phi(R). \end{aligned}$$

By Markov's inequality,

$$\mathbb{P}[|Z| > sR/c_T] \leq \frac{c_T \mathbb{E}[|Z|]}{sR},$$

$$\mathbb{P}[|Y| > (1-s)R] \leq \frac{\mathbb{E}[|Y|]}{(1-s)R}.$$

We have $\mathbb{E}[|Z|^p] + \mathbb{E}[|Y|^p] < \infty$. By dominated convergence theorem,

$$\lim_{R \rightarrow \infty} I_1(R) = \lim_{R \rightarrow \infty} I_2(R) = \lim_{R \rightarrow \infty} I_3(R) = \lim_{R \rightarrow \infty} I_4(R) = 0.$$

This completes the proof. ■

TOULOUSE SCHOOL OF ECONOMICS, 1 ESPLANADE DE L'UNIVERSITÉ, 31000 TOULOUSE, FRANCE

Email address, Le: `leanhdung1994@gmail.com`

TOULOUSE SCHOOL OF ECONOMICS, 1 ESPLANADE DE L'UNIVERSITÉ, 31000 TOULOUSE, FRANCE

Email address, Villeneuve: `stephane.villeneuve@tse-fr.eu`