

# EULER SCHEME FOR DENSITY DEPENDENT STOCHASTIC DIFFERENTIAL EQUATIONS

ZIMO HAO, MICHAEL RÖCKNER AND XICHENG ZHANG

ABSTRACT. In this paper we show the existence and uniqueness for a class of density dependent SDEs with bounded measurable drift, where the existence part is based on Euler's approximation for density dependent SDEs and the uniqueness is based on the associated nonlinear Fokker-Planck equation. As an application, we obtain the well-posedness of a nonlinear Fokker-Planck equation.

## 1. INTRODUCTION

In this paper we consider the following density dependent stochastic differential equation (abbreviated as DDSDE):

$$dX_t = b(t, X_t, \rho_t(X_t))dt + \sqrt{2}dW_t, \quad X_0 \stackrel{(d)}{=} \nu_0, \quad (1.1)$$

where  $W_t$  is a standard  $d$ -dimensional Brownian motion on some probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ ,  $b : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}^d$  is a bounded Borel measurable vector field and for  $t > 0$ ,  $\rho_t(x) = \mathbb{P} \circ X_t^{-1}(dx)/dx$  is the distributional density of  $X_t$  with respect to the Lebesgue measure  $dx$  on  $\mathbb{R}^d$ ,  $\nu_0$  is a probability measure over  $\mathbb{R}^d$ . By Itô's formula, one sees that  $\rho_t$  solves the following nonlinear Fokker-Planck equation (FPE) in the distributional sense:

$$\partial_t \rho_t - \Delta \rho_t + \operatorname{div}(b(t, \cdot, \rho_t) \rho_t) = 0, \quad \lim_{t \downarrow 0} \rho_t = \nu_0 \text{ weakly.} \quad (1.2)$$

More precisely, for any  $\varphi \in C_0^\infty(\mathbb{R}^d)$ ,

$$\langle \rho_t, \varphi \rangle = \langle \nu_0, \varphi \rangle + \int_0^t \langle \rho_s, \Delta \varphi \rangle ds + \int_0^t \langle \rho_s, b(s, \cdot, \rho_s) \cdot \nabla \varphi \rangle ds, \quad (1.3)$$

where  $\langle \rho_t, \varphi \rangle := \int_{\mathbb{R}^d} \varphi(x) \rho_t(x) dx = \mathbb{E} \varphi(X_t)$ .

Since the coefficients of SDE (1.1) depend on the distributional density of the solution  $X_t$  evaluated at  $X_t$ , the dependence of  $b$  on the measure  $\rho_t(x)dx$ , is called "Nemytskii-type" dependence (cf. [1], [2]). Thus (1.1) can be also called McKean-Vlasov SDE of Nemytskii-type. Let us first recall the definition of a weak solution to DDSDE (1.1):

**Definition 1.1.** *Let  $\nu_0$  be a probability measure on  $\mathbb{R}^d$ . We call a filtered probability space  $(\Omega, \mathcal{F}, \mathbf{P}; (\mathcal{F}_t)_{t \geq 0})$  together with a pair of processes  $(X, W)$  defined on it a weak solution of SDE (1.1) with initial distribution  $\nu_0$ , if*

- (i)  $\mathbf{P} \circ X_0^{-1} = \nu_0$ ,  $W$  is a  $d$ -dimensional  $\mathcal{F}_t$ -Brownian motion.

---

*Date:* September 18, 2020.

*Keywords:* density dependent SDE, heat kernel, Strong solution.

This work is supported by NNSFC grant of China (No. 11731009) and the DFG through the CRC 1283 "Taming uncertainty and profiting from randomness and low regularity in analysis, stochastics and their applications".

(ii) for each  $t > 0$ ,  $\mathbf{P} \circ X_t^{-1}(dx)/dx = \rho_t(x)$  and

$$X_t = X_0 + \int_0^t b(s, X_s, \rho_s(X_s))ds + \sqrt{2}W_t, \quad \mathbf{P} - a.s.$$

To show the existence of a weak solution, we consider the following Euler scheme to DDSDE (1.1): Let  $T > 0$ ,  $N \in \mathbb{N}$  and  $h := T/N$ . For  $t \in [0, h]$ , define

$$X_t^N := X_0 + \sqrt{2}W_t,$$

and for  $t \in [kh, (k+1)h]$  with  $k = 1, \dots, N$ , we inductively define  $X_t^N$  by

$$X_t^N := X_{kh}^N + \int_{kh}^t b(s, X_{kh}^N, \rho_{kh}^N(X_{kh}^N))ds + \sqrt{2}(W_t - W_{kh}), \quad (1.4)$$

where  $\rho_{kh}^N(x)$  is the distributional density of  $X_{kh}^N$ , whose existence is easily seen from the construction. We have the following existence and uniqueness result.

**Theorem 1.2.** *Assume that  $b$  is bounded measurable and for any  $t > 0$ ,*

$$\lim_{u \rightarrow u_0} \sup_{|x| < R} |b(t, x, u) - b(t, x, u_0)| = 0, \quad \forall R > 0. \quad (1.5)$$

**(Existence)** *For any  $T > 0$  and initial distribution  $\nu_0$ , there are a subsequence  $N_k$  and a weak solution  $(X, W)$  to DDSDE (1.1) in the sense of Definition 1.1 so that for any bounded measurable  $f$  and  $t \in (0, T]$ ,*

$$\lim_{k \rightarrow \infty} \mathbb{E}f(X_t^{N_k}) = \mathbf{E}f(X_t).$$

Moreover, for each  $t \in (0, T]$ ,  $X_t$  admits a density  $\rho_t$  satisfying the estimate

$$\rho_t(y) \leq Ct^{-d/2} \int_{\mathbb{R}^d} e^{-\frac{|x-y|^2}{\lambda t}} \nu_0(dx), \quad y \in \mathbb{R}^d,$$

where  $C, \lambda \geq 1$ , and the following  $L^1$ -convergence holds:

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}^d} |\rho_t^{N_k}(y) - \rho_t(y)| dy = 0. \quad (1.6)$$

**(Uniqueness)** *Suppose that  $\nu_0(dx) = \rho_0(x)dx$  with  $\rho_0 \in (L^1 \cap L^q)(\mathbb{R}^d)$  for some  $q \in (d, \infty]$ , and there is a  $C > 0$  such that for all  $t, x, u, u'$ ,*

$$|b(t, x, u) - b(t, x, u')| \leq C|u - u'|. \quad (1.7)$$

Then weak and strong uniqueness hold for SDE (1.1).

**Remark 1.3.** If the uniqueness holds, then limit (1.6) holds for the whole sequence.

As a consequence of Theorem 1.2, we have the following well-posedness of the nonlinear FPE (1.2).

**Corollary 1.4.** *Let  $\nu_0$  be a probability measure over  $\mathbb{R}^d$ .*

(i) *Assume  $b$  is bounded and Borel measurable such that (1.5) holds. Then there is a weak solution  $\rho_t$  to PDE (1.2) in the sense (1.3) with  $\int \rho_t = 1$  and*

$$0 \leq \rho_t(x) \leq Ct^{-d/2} \int_{\mathbb{R}^d} e^{-\frac{|x-y|^2}{\lambda t}} \nu_0(dy), \quad x \in \mathbb{R}^d, \quad t \in (0, T]. \quad (1.8)$$

(ii) *Assume that (1.7) holds and that  $\nu_0(dx) = \rho_0(x)dx$  with  $\rho_0 \in (L^1 \cap L^q)(\mathbb{R}^d)$  for some  $q \in (d, \infty]$ . Then the solution in assertion (i) is the unique weak solution to (1.2).*

*Proof.* (i) The existence of a weak solution follows from the existence part of Theorem 1.2 and Itô's formula.

(ii) The uniqueness follows from the uniqueness part of Theorem 1.2 by Section 2 in [2].  $\square$

McKean-Vlasov SDE of Nemytskii-type (density dependent SDE), i.e.

$$dX_t = b\left(t, X_t, \frac{d\mathcal{L}_{X_t}}{dx}(X_t)\right)dt + \sigma\left(t, X_t, \frac{d\mathcal{L}_{X_t}}{dx}(X_t)\right)dW_t, \quad X_0 \stackrel{(d)}{=} \nu_0,$$

were first introduced in [1, Section 2]. In [1], [2], for a large class of time independent coefficients  $b, \sigma$ , Barbu together with the second named author obtained the existence of weak solutions for such (possibly degenerate) density dependent SDEs (see Section 2). The strategy in [1] or [2] is to solve the associated nonlinear Fokker-Planck equation and then by the well-known superposition principle (cf. [11], generalizing [9] and [7]) to establish the existence of a weak solution to DDSDE (1.1). Later, in [3], the same authors prove the uniqueness of the weak solution to DDSDE (1.1), which is a consequence of the uniqueness of the corresponding nonlinear Fokker-Planck equation. Recently, in [4], they also consider the existence of solutions to a class of nonlinear Fokker-Planck equations with measure-valued initial data. The strategy of this paper is completely different from that in [1–4]. We start directly from DDSDE (1.1) and prove weak existence of solutions through Euler's approximation for DDSDE (1.1) and using simple heat kernel estimates. In other words, we do not use the superposition principle. Moreover, our assumptions on the drift are weaker. Especially, there is no regularity assumption of  $b$  in  $x$ . On the other hand we only consider the case  $\sigma = \text{identity}$ .

As explained above we obtain weak (i.e. in the sense of Schwartz distributions) solutions to FPE (1.2). Let us mention here that Chen and Perthame in [6] studied the Cauchy problem for a general nonlinear degenerate parabolic-hyperbolic equation of second order in the framework of kinetic and entropy solutions, in the case the coefficient do not depend explicitly on  $x$  and  $t$ .

However, all the above results do not cover Theorem 1.2 above. In particular, we use a purely probabilistic method to show the existence of weak solutions for the nonlinear FPE (1.2). This is the main contribution of the present paper. Here, for simplicity, we only consider the additive noise case. For the case of uniformly elliptic diffusion coefficients, it would also work by the corresponding heat kernel estimates (see [5] and [10]).

This paper is organized as follows: In Section 2, we give some necessary preliminaries about heat kernel estimates for the Euler scheme with bounded measurable drifts. In Section 3, we prove our main Theorem 1.2 by Euler's type approximation (cf. [13]). Note that the usual Picard iteration does not seem to work for DDSDE (1.1), since  $x \rightarrow b(t, x, \rho_t(x))$  is too singular.

Throughout this paper, we use the following conventions: The letter  $C$  denotes a constant, whose value may change in different places. We also use  $A \lesssim B$  to denote  $A \leq CB$  for some unimportant constant  $C > 0$ .

## 2. HEAT KERNEL OF EULER SCHEME

In this section we show heat kernel estimates for Euler's scheme of usual SDEs. First of all, we recall some basic properties about the Gaussian heat kernel. Let

$$g(t, x) := (4\pi t)^{-\frac{d}{2}} e^{-\frac{|x|^2}{4t}}, \quad t > 0, x \in \mathbb{R}^d, \quad (2.1)$$

which is the fundamental solution of  $\Delta$ , i.e.,

$$\partial_t g(t, x) = \Delta g(t, x).$$

Moreover, we have the following Chapman-Kolmogorov equations:

$$(g(t) * g(s))(x) := \int_{\mathbb{R}^d} g(t, x - z)g(s, z)dz = g(t + s, x), \quad t, s > 0, \quad (2.2)$$

and the following easy facts,

$$g(t, x + y) \leq 2^{\frac{d}{2}} g(2t, x) e^{-\frac{|y|^2}{4t}}, \quad |\nabla g|(t, x) \leq \frac{2^{d/2}}{\sqrt{t}} g(2t, x). \quad (2.3)$$

The following lemma is straightforward and elementary. For the readers' convenience, we provide a detail proof.

**Lemma 2.1.** *For any  $T > 0$ ,  $\beta \in (0, 1)$  and  $j = 0, 1$ , there is a constant  $C = C(T, \beta, j, d) > 0$  such that for any  $0 < t \leq T$  and  $x_1, x_2 \in \mathbb{R}^d$ ,*

$$|\nabla^j g(t, x_1) - \nabla^j g(t, x_2)| \leq C |x_1 - x_2|^{\beta} t^{-\frac{j}{2} - \beta} \sum_{i=1,2} g(4t, x_i), \quad (2.4)$$

and for any  $0 < t_1 < t_2 \leq T$  and  $x \in \mathbb{R}^d$ ,

$$|\nabla^j g(t_1, x) - \nabla^j g(t_2, x)| \leq C |t_2 - t_1|^{\frac{\beta}{2}} \sum_{i=1,2} t_i^{-\frac{j+\beta}{2}} g(2t_i, x). \quad (2.5)$$

*Proof.* (i) By definition (2.1), it is easy to see that for  $k = 1, 2, 3$ , there is a constant  $C > 0$  only depending on  $k, d$  such that

$$|\nabla^k g(t, x)| \leq C(4\pi t)^{-\frac{d}{2}} t^{\frac{k}{2}} e^{-\frac{|x|^2}{8t}} = C 2^{\frac{d}{2}} t^{-\frac{k}{2}} g(2t, x).$$

Thus, for  $j = 0, 1$  and  $\beta \in (0, 1)$ , if  $|x_1 - x_2| > \sqrt{t}$ , then

$$\begin{aligned} |\nabla^j g(t, x_1) - \nabla^j g(t, x_2)| &\lesssim t^{-\frac{j}{2}} (g(2t, x_1) + g(2t, x_2)) \\ &\lesssim |x_1 - x_2|^{\beta} t^{-\frac{j}{2} - \beta} (g(2t, x_1) + g(2t, x_2)); \end{aligned}$$

if  $|x_1 - x_2| \leq \sqrt{t}$ , then by the mean-value formula,

$$\begin{aligned} |\nabla^j g(t, x_1) - \nabla^j g(t, x_2)| &\leq |x_1 - x_2| \int_0^1 |\nabla^{j+1} g(t, x_1 + \theta(x_2 - x_1))| d\theta \\ &\lesssim |x_1 - x_2| t^{-\frac{j+1}{2}} \int_0^1 g(2t, x_1 + \theta(x_2 - x_1)) d\theta \\ &\lesssim |x_1 - x_2| t^{-\frac{j+1}{2}} g(4t, x_1) \lesssim |x_1 - x_2|^{\beta} t^{-j/2 - \beta} g(4t, x_1). \end{aligned}$$

Combining the above calculations, we get (2.4).

(ii) If  $t_2 - t_1 \leq t_1$ , then by the mean-value formula,

$$\begin{aligned} |\nabla^j g(t_1, x) - \nabla^j g(t_2, x)| &\leq |t_1 - t_2| \int_0^1 |\nabla^j \partial_t g|(t_1 + \theta(t_2 - t_1), x) d\theta \\ &= |t_1 - t_2| \int_0^1 |\nabla^j \Delta g|(t_1 + \theta(t_2 - t_1), x) d\theta \\ &\lesssim |t_1 - t_2| \int_0^1 \frac{g(2(t_1 + \theta(t_2 - t_1)), x)}{(t_1 + \theta(t_2 - t_1))^{1+j/2}} d\theta \\ &\lesssim |t_1 - t_2| t_1^{-1 - \frac{j}{2}} g(2t_2, x) \lesssim |t_1 - t_2|^{\frac{\beta}{2}} t_2^{-\frac{\beta}{2}} g(2t_2, x); \end{aligned}$$

if  $t_2 - t_1 > t_1$ , then  $t_2 \leq 2(t_2 - t_1)$  and

$$|\nabla^j g(t_1, x) - \nabla^j g(t_2, x)| \lesssim t_1^{-\frac{j}{2}} g(2t_1, x) + t_2^{-\frac{j}{2}} g(2t_2, x)$$

$$\lesssim |t_1 - t_2|^{\frac{\beta}{2}} \left( t_1^{-\frac{j+\beta}{2}} g(2t_1, x) + t_2^{-\frac{j+\beta}{2}} g(2t_2, x) \right).$$

The proof is complete.  $\square$

Let  $b : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  be a bound measurable function. Fix  $T > 0$  and  $x \in \mathbb{R}^d$ . Let  $X_t^N = X_t^N(x)$  be defined by the following Euler scheme:

$$X_t^N = x + \int_0^t b(s, X_{\phi_N(s)}^N) ds + \sqrt{2}W_t, \quad t \in [0, T], \quad (2.6)$$

where  $\phi_N(s) := kT/N$  for  $s \in [kT/N, (k+1)T/N)$ . We have the following Duhamel formula.

**Lemma 2.2.** *For each  $t \in (0, T]$  and  $x \in \mathbb{R}^d$ ,  $X_t^N(x)$  admits a density  $p_x^N(t, y)$  which satisfies the following Duhamel formula:*

$$p_x^N(t, y) = g(t, x - y) + \int_0^t \mathbb{E} \left[ b(s, X_{\phi_N(s)}^N) \cdot \nabla g(t - s, y - X_s^N) \right] ds. \quad (2.7)$$

*Proof.* Fix  $t \in (0, T]$  and  $f \in C_c^\infty(\mathbb{R}^d)$ . For  $s \in [0, t]$ , let  $u(s, x) := g(t - s, \cdot) * f(x)$ . Since  $(\partial_s + \Delta)u \equiv 0$  and  $u(t, x) = f(x)$ , by Itô's formula, we have

$$\mathbb{E}f(X_t^N) = \mathbb{E}u(t, X_t^N) = u(0, x) + \int_0^t \mathbb{E} \left[ b(s, X_{\phi_N(s)}^N) \cdot \nabla u(s, X_s^N) \right] ds.$$

From this, we derive the desired Duhamel formula.  $\square$

**Remark 2.3.** For a general initial value  $X_0^N = X_0 \in \mathcal{F}_0$  and each  $t \in (0, T]$ , since for each  $x \in \mathbb{R}^d$ ,  $X_t^N(x)$  is independent of  $X_0$ , the Euler scheme  $X_t^N$  defined by (2.6) with initial value  $X_0$  also has a density  $p_{X_0}^N(t, y)$  given by

$$p_{X_0}^N(t, y) = \int_{\mathbb{R}^d} p_x^N(t, y) \mathbb{P} \circ X_0^{-1}(dx). \quad (2.8)$$

The following Gaussian type estimate for  $p_x^N(t, y)$  was proved by Lemaire and Menozzi [8]. Since it is not difficult, for the readers' convenience, we provide a detailed proof here.

**Theorem 2.4.** *For any  $T > 0$ , there is a constant  $C = C(d, T, \|b\|_\infty)$  such that for all  $N \in \mathbb{N}$ ,  $t \in (0, T]$  and  $x, y \in \mathbb{R}^d$ ,*

$$p_x^N(t, y) \leq Cg(4t, x - y). \quad (2.9)$$

*Proof.* Let  $\varepsilon > 0$  be small enough so that

$$\ell_\varepsilon := 2^{d+1} \sqrt{\varepsilon \|b\|_\infty^2} e^{\varepsilon \|b\|_\infty^2} \leq 1/2.$$

Fix  $T > 0$ . Without loss of generality, we assume

$$N \geq (\|b\|_\infty^2 T / (4 \log 2)) \vee (T/\varepsilon). \quad (2.10)$$

For simplicity we shall write

$$h := T/N, \quad M := \lceil \varepsilon/h \rceil \in \mathbb{N}.$$

*Step 1:* In this step we use induction to show that for all  $k = 1, \dots, M \wedge N$ ,

$$p_x^N(kh, y) \leq 2^{d+1} g(4kh, x - y). \quad (2.11)$$

First of all, for  $k = 1$ , since  $X_h^N = x + W_h + \int_0^h b(s, x) ds$ , by (2.3) and (2.10) we have

$$p_x^N(h, y) = g(h, y - \int_0^h b(s, x) ds - x) \leq 2^{d/2} e^{\|b\|_\infty^2 h/4} g(2h, x - y)$$

$$\leq 2^d e^{\|b\|_\infty^2 T/(4N)} g(4h, x - y) \leq 2^{d+1} g(4h, x - y).$$

Suppose now that (2.11) holds for  $j = 1, 2, \dots, k-1$ . By Duhamel's formula (2.7), we have

$$\begin{aligned} p_x^N(kh, y) - g(kh, x - y) &= \int_0^{kh} \mathbb{E} \left[ b(s, X_{\phi_N(s)}^N) \cdot \nabla g(kh - s, y - X_s^N) \right] ds \\ &= \sum_{j=0}^{k-1} \int_{jh}^{(j+1)h} \mathbb{E} \left[ b(s, X_{jh}^N) \cdot \nabla g(kh - s, y - X_s^N) \right] ds. \end{aligned} \quad (2.12)$$

Note that for  $s \in (jh, (j+1)h)$ ,

$$X_s^N = X_{jh}^N + \sqrt{2}(W_s - W_{jh}) + \int_{jh}^s b(r, X_{jh}^N) dr.$$

Since  $\sqrt{2}(W_s - W_{jh})$  is independent of  $X_{jh}^N$  and has density  $g(s - jh, y)$ , by the C-K equations (2.2) we have

$$\begin{aligned} \mathcal{I}_j(s) &:= \mathbb{E} \left[ b(s, X_{jh}^N) \cdot \nabla g(kh - s, X_s^N - y) \right] \\ &= \mathbb{E} \left[ b(s, X_{jh}^N) \cdot \nabla g(kh - s) * g(s - jh) \left( X_{jh}^N + \int_{jh}^s b(r, X_{jh}^N) dr - y \right) \right] \\ &= \mathbb{E} \left[ b(s, X_{jh}^N) \cdot \nabla g(kh - jh, X_{jh}^N + \int_{jh}^s b(r, X_{jh}^N) dr - y) \right] \\ &\leq \|b\|_\infty \int_{\mathbb{R}^d} |\nabla g(kh - jh, z + \int_{jh}^s b(r, z) dr - y)| p_x^N(jh, z) dz. \end{aligned}$$

By (2.3) and induction hypothesis, we further have for  $s \in (jh, (j+1)h)$ ,

$$\begin{aligned} \mathcal{I}_j(s) &\leq \frac{\|b\|_\infty 2^{d/2}}{\sqrt{kh - jh}} \int_{\mathbb{R}^d} g(2(kh - jh), z + \int_{jh}^s b(r, z) dr - y) p_x^N(jh, z) dz \\ &\leq \frac{\|b\|_\infty 2^d e^{(k-j)h \|b\|_\infty^2/4}}{\sqrt{kh - jh}} \int_{\mathbb{R}^d} g(4(kh - jh), z - y) \cdot 2^{d+1} g(4jh, x - z) dz \\ &\leq \frac{\|b\|_\infty 2^{2d+1} e^{\varepsilon \|b\|_\infty^2/4}}{\sqrt{kh - s}} g(4kh, x - y) = \frac{2^d \ell_\varepsilon / \sqrt{\varepsilon}}{\sqrt{kh - s}} g(4kh, x - y), \end{aligned}$$

where we have used  $kh \leq Mh \leq \varepsilon$ . Substituting this into (2.12), we obtain

$$\begin{aligned} |p_x^N(kh, y) - g(kh, x - y)| &\leq \frac{2^d \ell_\varepsilon}{\sqrt{\varepsilon}} g(4kh, x - y) \sum_{j=0}^{k-1} \int_{jh}^{(j+1)h} \frac{1}{\sqrt{kh - s}} ds \\ &\leq \frac{2^d \ell_\varepsilon}{\sqrt{\varepsilon}} g(4kh, x - y) \int_0^{kh} \frac{1}{\sqrt{kh - s}} ds \\ &= \frac{2^d \ell_\varepsilon}{\sqrt{\varepsilon}} g(4kh, x - y) 2\sqrt{kh} \leq 2^{d+1} \ell_\varepsilon g(4kh, x - y), \end{aligned}$$

which implies, since  $g(t, x) \leq 2^d g(4t, x)$  and  $2\ell_\varepsilon \leq 1$ , that

$$p_x^N(kh, y) \leq 2^d (1 + 2\ell_\varepsilon) g(4kh, x - y) \leq 2^{d+1} g(4kh, x - y).$$

*Step 2:* Next we assume  $M < N$  and consider  $k = M + 1 \dots, 2M$ . Note that

$$\begin{aligned} X_{t+Mh}^N &= X_{Mh}^N + W_{t+Mh} - W_{Mh} + \int_{Mh}^{t+Mh} b(s, X_{\phi_N(s)}^N) ds \\ &= X_{Mh}^N + W_{t+Mh} - W_{Mh} + \int_0^t b(s + Mh, X_{\phi_N(s)+Mh}^N) ds, \end{aligned}$$

where we have used that  $\phi_N(s + Mh) = \phi_N(s) + Mh$ . In particular, if we let

$$\bar{X}_t^N := X_{t+Mh}^N, \quad \bar{W}_t := W_{t+Mh} - W_{Mh},$$

then for  $t \in [0, Mh]$ ,

$$\bar{X}_t^N = X_{Mh}^N + \bar{W}_t + \int_0^t b(s + Mh, \bar{X}_{\phi_N(s)}^N) ds.$$

Let  $\bar{p}_x^N(kh, y)$  be the density of  $\bar{X}_t^N$  with  $\bar{X}_0^N = x$ . By Step 1, we have

$$\bar{p}_x^N(kh, y) \leq 2^{d+1} g(4kh, x - y), \quad k = 1, \dots, M.$$

Thus, for  $k = 1, \dots, M$ , by (2.8) we have

$$\begin{aligned} p_x^N((k+M)h, y) &= \int_{\mathbb{R}^d} \bar{p}_z^N(kh, y) p_x^N(Mh, z) dz \\ &\leq 4^{d+1} \int_{\mathbb{R}^d} g(4kh, z - y) g(4Mh, x - z) dz \\ &= 4^{d+1} g(4(k+M)h, x - y). \end{aligned}$$

Repeating the above procedure  $\lceil \frac{T}{\varepsilon} \rceil + 1$ -times, we obtain that for some  $C > 0$  independent of  $N$ ,

$$p_x^N(kh, y) \leq C g(kh, x - y), \quad k = 1, \dots, N.$$

*Step 3:* Note that for  $t \in (kh, (k+1)h)$ ,

$$X_t^N = X_{kh}^N + W_t - W_{kh} + \int_{kh}^t b(s, X_{kh}^N) ds,$$

where  $W_t - W_{kh}$  is independent of  $X_{kh}^N$ . Hence,

$$\begin{aligned} p_x^N(t, y) &= \int_{\mathbb{R}^d} g(t - kh, z + \int_{kh}^t b(s, z) ds - y) p_x^N(kh, z) dz \\ &\leq C e^{(t-kh)\|b\|_\infty^2/4} \int_{\mathbb{R}^d} g(4(t-kh), y - z) g(4kh, x - z) dz \\ &\leq C e^{T\|b\|_\infty^2/4} g(4t, x - y). \end{aligned}$$

This completes the proof.  $\square$

The following corollary is a combination of Theorem 2.4 and Lemma 2.1.

**Corollary 2.5.** *Let  $\nu_0(dy) = \mathbb{P} \circ X_0^{-1}(dy)$  be the distribution of  $X_0$ .*

- (i) *For any  $T > 0$ , there is a constant  $C = C(d, T, \|b\|_\infty)$  such that for all  $N \in \mathbb{N}$ ,  $t \in (0, T)$  and  $y \in \mathbb{R}^d$*

$$p_{X_0}^N(t, y) \leq C \int_{\mathbb{R}^d} g(4t, x - y) \nu_0(dx). \quad (2.13)$$

- (ii) *For any  $T > 0$  and  $\beta \in (0, 1)$ , there is a constant  $C = C(d, T, \|b\|_\infty, \beta)$  such that for all  $N \in \mathbb{N}$ ,  $t \in (0, T)$  and  $y_1, y_2 \in \mathbb{R}^d$ ,*

$$|p_{X_0}^N(t, y_1) - p_{X_0}^N(t, y_2)| \leq C |y_1 - y_2|^\beta t^{-\frac{\beta}{2}} \sum_{j=1,2} \int_{\mathbb{R}^d} g(4t, x - y_j) \nu_0(dx).$$

- (iii) *For any  $T > 0$  and  $\beta \in (0, 1)$ , there is a constant  $C = C(d, T, \|b\|_\infty, \beta)$  such that for all  $N \in \mathbb{N}$ ,  $t_1, t_2 \in (0, T)$  and  $y \in \mathbb{R}^d$ ,*

$$|p_{X_0}^N(t_1, y) - p_{X_0}^N(t_2, y)| \leq C |t_1 - t_2|^{\beta/2} \sum_{j=1,2} t_j^{-\beta/2} \int_{\mathbb{R}^d} g(2t_j, x - y) \nu_0(dx).$$

*Proof.* (i) is a direct consequence of (2.8) and Theorem 2.4. We only show (iii) since (ii) is similar by (2.4). Suppose  $t_1 < t_2$ . By (2.7), we have

$$\begin{aligned} |p_{X_0}^N(t_1, y) - p_{X_0}^N(t_2, y)| &\leq \int_{\mathbb{R}^d} |g(t_1, x - y) - g(t_2, x - y)| \nu_0(dx) \\ &+ \|b\|_\infty \int_0^{t_1} \int_{\mathbb{R}^d} |\nabla g(t_1 - s, y - z) - \nabla g(t_2 - s, y - z)| p_{X_0}^N(s, z) dz ds \\ &+ \|b\|_\infty \int_{t_1}^{t_2} \int_{\mathbb{R}^d} |\nabla g(t_1 - s, y - z)| p_{X_0}^N(s, z) dz ds =: I_1 + I_2 + I_3. \end{aligned}$$

For  $I_1$ , by (2.5), we have

$$I_1 \lesssim |t_1 - t_2|^{\frac{\beta}{2}} \sum_{j=1,2} t_j^{-\frac{\beta}{2}} \int_{\mathbb{R}^d} g(2t_j, x - y) \nu_0(dx).$$

For  $I_2$ , by (i), (2.5) and the C-K equations (2.2), we have

$$\begin{aligned} I_2 &\lesssim |t_1 - t_2|^{\frac{\beta}{2}} \sum_{j=1,2} \int_0^{t_1} \left[ \int_{\mathbb{R}^d} (t_j - s)^{-\frac{1+\beta}{2}} g(4(t_j - s), z - y) \right. \\ &\quad \left. \times \int_{\mathbb{R}^d} g(4s, x - z) \nu_0(dx) dz \right] ds \\ &= |t_1 - t_2|^{\frac{\beta}{2}} \sum_{j=1,2} \int_0^{t_1} (t_j - s)^{-\frac{1+\beta}{2}} \int_{\mathbb{R}^d} g(4t_j, x - y) \nu_0(dx) ds \\ &\lesssim |t_1 - t_2|^{\frac{\beta}{2}} \sum_{j=1,2} \int_{\mathbb{R}^d} g(4t_j, x - y) \nu_0(dx). \end{aligned}$$

For  $I_3$ , by (i), (2.5) and the C-K equations, we have

$$\begin{aligned} I_3 &\lesssim \int_{t_1}^{t_2} \int_{\mathbb{R}^d} (t_1 - s)^{-\frac{1}{2}} g(4(t_1 - s), z - y) \int_{\mathbb{R}^d} g(4s, x - z) \nu_0(dx) dz ds \\ &= \int_{t_1}^{t_2} \int_{\mathbb{R}^d} (t_1 - s)^{-\frac{1}{2}} g(4t_1, x - y) \nu_0(dx) ds \lesssim |t_2 - t_1|^{\frac{1}{2}} \int_{\mathbb{R}^d} g(4t_1, x - y) \nu_0(dx). \end{aligned}$$

Combining the above calculations, we obtain the desired estimate.  $\square$

### 3. PROOF OF THEOREM 1.2

Let  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  be a complete filtered probability space,  $W_t$  a  $d$ -dimensional standard  $\mathcal{F}_t$ -Brownian motion, and  $X_0$  an  $\mathcal{F}_0$ -measurable random variable with distribution  $\nu_0$ . Let  $T > 0$ ,  $N \in \mathbb{N}$  and  $h := T/N$ . Let  $X_t^N$  be the Euler approximation of DDSDE (1.1) constructed in the introduction. From the construction (1.4), it is easy to see that  $X_t^N$  solves the following SDE:

$$X_t^N = X_0 + \int_0^t b^N(s, X_{\phi_N(s)}^N) ds + \sqrt{2} W_t, \quad (3.1)$$

where

$$b^N(s, x) = \mathbf{1}_{\{s \geq h\}} b\left(s, x, \rho_{\phi_N(s)}^N(x)\right) \quad (3.2)$$

and

$$\phi_N(s) := \sum_{j=0}^{\infty} j h \mathbf{1}_{[jh, (j+1)h)}(s). \quad (3.3)$$

The following lemma is easy by (3.1) and since  $\|b^N\|_\infty \leq \|b\|_\infty$ .

**Lemma 3.1.** *For any  $T > 0$ , there is a constant  $C > 0$  such that for all  $s, t \in [0, T]$ ,*

$$\sup_N \mathbb{E} |X_t^N - X_s^N|^4 \leq C |s - t|^2.$$

Let  $p_x^N(t, y)$  be the distributional density of the Euler scheme  $X_t^N(x)$  of SDE (3.1) starting from  $x$  at time 0. Since for each  $x \in \mathbb{R}^d$ ,  $X_t^N(x)$  is independent of  $X_0$ , the distributional density  $\rho_t^N(y)$  of  $X_t^N$  with initial distribution  $\nu_0$  is given by

$$\rho_t^N(y) = \int_{\mathbb{R}^d} p_x^N(t, y) \nu_0(dx). \quad (3.4)$$

The following lemma is crucial for the existence of a solution to DDSDE (1.1).

**Lemma 3.2.** *For fixed  $T > 0$ , there are a subsequence  $(N_k)_{k \in \mathbb{N}}$  and a continuous function  $\rho \in C((0, T] \times \mathbb{R}^d)$  such that for any  $M \in \mathbb{N}$ ,*

$$\lim_{k \rightarrow \infty} \sup_{|y| \leq M} \sup_{1/M \leq t \leq T} |\rho_t^{N_k}(y) - \rho_t(y)| = 0. \quad (3.5)$$

*Proof.* First of all, by the upper-bound estimate (2.9) for  $p_x^N(t, y)$ , we have

$$\sup_{|y| \leq M} \sup_{1/M \leq t \leq T} |\rho_t^N(y)| \leq C \int_{\mathbb{R}^d} \sup_{|y| \leq M} \sup_{1/M \leq t \leq T} |g(4t, x - y)| \nu_0(dx) \leq C_M,$$

where  $C_M$  is independent of  $N$ . Moreover, by Corollary 2.5, we have for any  $\beta < 1$ ,  $t_1, t_2 \in [1/M, T]$  and  $y_1, y_2 \in \mathbb{R}^d$ ,

$$\begin{aligned} |\rho_{t_1}^N(y_1) - \rho_{t_2}^N(y_2)| &\leq |\rho_{t_1}^N(y_1) - \rho_{t_2}^N(y_1)| + |\rho_{t_2}^N(y_1) - \rho_{t_2}^N(y_2)| \\ &\lesssim |t_2 - t_1|^{\frac{\beta}{2}} \sum_{i=1,2} \int_{\mathbb{R}^d} |g(2t_i, y_1 - x)| \nu_0(dx) \\ &\quad + |y_1 - y_2|^\beta \sum_{i=1,2} \int_{\mathbb{R}^d} |g(4t_2, y_i - x)| \nu_0(dx) \\ &\lesssim M^{-(d+1+\beta)/2} \left( |t_2 - t_1|^{\frac{\beta}{2}} + |y_1 - y_2|^\beta \right), \end{aligned} \quad (3.6)$$

where the implicit constants in the above  $\lesssim$  are independent of  $N$ . Thus, by Ascoli-Arzelà's theorem, we conclude the proof and have (3.5).  $\square$

Now we are in a position to give the

*Proof of Theorem 1.2. (Existence)* Fix  $T > 0$ . Let  $\mathbb{W}$  be the space of all continuous functions from  $[0, T]$  to  $\mathbb{R}^d$ . Let  $\mathbb{Q}_N$  be the law of  $(X^N, W)$  in  $\mathbb{W} \times \mathbb{W}$ . By Lemma 3.1 and Kolmogorov's criterion,  $\{\mathbb{Q}_N\}_{N \in \mathbb{N}}$  is tight. Therefore, by Prokhorov's theorem, there are a subsequence  $(N_k)_{k \in \mathbb{N}}$  and a probability measure  $\mathbb{Q}$  on  $\mathbb{W} \times \mathbb{W}$  so that

$$\mathbb{Q}_{N_k} \rightarrow \mathbb{Q} \text{ weakly.}$$

Without loss of generality, we assume that the subsequence is the same as that in Lemma 3.2. Below, we still denote the above subsequence by  $\mathbb{Q}_N$ ,  $N \in \mathbb{N}$  for simplicity. Now, by Skorokhod's representation theorem, there are probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$  and random variables  $(\tilde{X}^N, \tilde{W}^N)$  and  $(\tilde{X}, \tilde{W})$  thereon such that

$$(\tilde{X}^N, \tilde{W}^N) \rightarrow (\tilde{X}, \tilde{W}), \quad \tilde{\mathbb{P}}\text{-a.s.} \quad (3.7)$$

and

$$\tilde{\mathbb{P}} \circ (\tilde{X}^N, \tilde{W}^N)^{-1} = \mathbb{Q}_N = \mathbb{P} \circ (X^N, W)^{-1}, \quad \tilde{\mathbb{P}} \circ (\tilde{X}, \tilde{W})^{-1} = \mathbb{Q}. \quad (3.8)$$

In particular, the distributional density of  $\widetilde{X}_t^N$  is  $\rho_t^N$ . Moreover, by Lemma 3.2 and (3.7), for any  $t \in (0, T)$  and  $\varphi \in C_0^\infty(\mathbb{R}^d)$ ,

$$\mathbb{E}\varphi(\widetilde{X}_t) = \lim_{N \rightarrow \infty} \mathbb{E}\varphi(\widetilde{X}_t^N) = \lim_{N \rightarrow \infty} \int_{\mathbb{R}^d} \varphi(y) \rho_t^N(y) dy = \int_{\mathbb{R}^d} \varphi(y) \rho_t(y) dy.$$

In other words,  $\rho_t$  is the density of  $\widetilde{X}_t$ . Define  $\widetilde{\mathcal{F}}_t^N := \sigma(\widetilde{X}^N, \widetilde{W}^N; s \leq t)$ . We note that

$$\mathbb{P}[W_t - W_s \in \cdot | \mathcal{F}_s] = \mathbb{P}\{W_t - W_s \in \cdot\},$$

hence,

$$\widetilde{\mathbb{P}}[\widetilde{W}_t^N - \widetilde{W}_s^N \in \cdot | \widetilde{\mathcal{F}}_s^N] = \widetilde{\mathbb{P}}\{\widetilde{W}_t^N - \widetilde{W}_s^N \in \cdot\},$$

which means that  $\widetilde{W}_t^N$  is an  $\widetilde{\mathcal{F}}_t^N$ -BM. Thus, by (3.1) and (3.8) we have

$$\widetilde{X}_t^N = \widetilde{X}_0^N + \int_0^t b^N(s, \widetilde{X}_{\phi_N(s)}^N) ds + \sqrt{2} \widetilde{W}_t^N. \quad (3.9)$$

Let us now show that

$$\int_0^t \mathbf{1}_{s \geq h} b(s, \widetilde{X}_{\phi_N(s)}^N, \rho_{\phi_N(s)}^N(\widetilde{X}_{\phi_N(s)}^N)) ds \rightarrow \int_0^t b(s, \widetilde{X}_s, \rho_s(\widetilde{X}_s)) ds, \quad (3.10)$$

in probability as  $N \rightarrow \infty$ . Let  $\Omega_0 \subset \widetilde{\Omega}$  be a measurable set so that  $\widetilde{\mathbb{P}}(\Omega_0) = 1$  and for each  $\omega \in \Omega_0$ ,

$$\lim_{N \rightarrow \infty} (\widetilde{X}_{\phi_N(\cdot)}^N(\omega), \widetilde{W}_{\cdot}^N(\omega)) = (\widetilde{X}_{\cdot}(\omega), \widetilde{W}_{\cdot}(\omega)). \quad (3.11)$$

In particular, for each fixed  $\omega \in \Omega_0$  and  $s \in (0, T)$ , by (3.5), (3.6) and (3.11), we have

$$\lim_{N \rightarrow \infty} |\rho_{\phi_N(s)}^N(\widetilde{X}_{\phi_N(s)}^N(\omega)) - \rho_s(\widetilde{X}_{\phi_N(s)}^N(\omega))| = 0, \quad a.s. \quad (3.12)$$

On the other hand, for any  $s > 0$ , by (1.5) we have

$$\lim_{|\delta| \rightarrow 0} \sup_{|u| \leq R} \sup_{|x| < R} |b(s, x, u + \delta) - b(s, x, u)| = 0, \quad \forall R > 0.$$

Thus, by (3.11) and (3.12), we have for each  $s > 0$  and  $\omega \in \Omega_0$ ,

$$\lim_{N \rightarrow \infty} |b(s, \widetilde{X}_{\phi_N(s)}^N(\omega), \rho_{\phi_N(s)}^N(\widetilde{X}_{\phi_N(s)}^N(\omega))) - b(s, \widetilde{X}_{\phi_N(s)}^N(\omega), \rho_s(\widetilde{X}_{\phi_N(s)}^N(\omega)))| = 0,$$

which, by the dominated convergence theorem, implies that

$$\begin{aligned} & \lim_{N \rightarrow \infty} \widetilde{\mathbb{E}} \int_0^t |\mathbf{1}_{s > h} b(s, \widetilde{X}_{\phi_N(s)}^N, \rho_{\phi_N(s)}^N(\widetilde{X}_{\phi_N(s)}^N)) - b(s, \widetilde{X}_{\phi_N(s)}^N, \rho_s(\widetilde{X}_{\phi_N(s)}^N))| ds \\ &= \widetilde{\mathbb{E}} \int_0^t \lim_{N \rightarrow \infty} |\mathbf{1}_{s > h} b(s, \widetilde{X}_{\phi_N(s)}^N, \rho_{\phi_N(s)}^N(\widetilde{X}_{\phi_N(s)}^N)) - b(s, \widetilde{X}_{\phi_N(s)}^N, \rho_s(\widetilde{X}_{\phi_N(s)}^N))| ds = 0. \end{aligned}$$

For proving (3.10), it remains to show

$$\lim_{N \rightarrow \infty} \widetilde{\mathbb{E}} \int_h^t |b(s, \widetilde{X}_{\phi_N(s)}^N, \rho_s(\widetilde{X}_{\phi_N(s)}^N)) - b(s, \widetilde{X}_s, \rho_s(\widetilde{X}_s))| ds = 0.$$

Let  $K_\varepsilon$  be a family of mollifiers in  $\mathbb{R}^d$ . Define

$$B_\varepsilon(t, x) = b(t, \cdot, \rho_t(\cdot)) * K_\varepsilon(x).$$

Clearly, for fixed  $\varepsilon > 0$ , by (3.11) we have

$$\lim_{N \rightarrow \infty} \widetilde{\mathbb{E}} \int_h^t |B_\varepsilon(s, \widetilde{X}_{\phi_N(s)}^N) - B_\varepsilon(s, \widetilde{X}_s)| ds = 0.$$

Below for notational convenience, we write  $\tilde{X}_t^\infty := \tilde{X}_t$  and  $\phi_\infty(s) := s$ . For  $N \in \mathbb{N} \cup \{\infty\}$ , we have

$$\begin{aligned} & \tilde{\mathbb{E}} \int_h^t |B_\varepsilon(s, \tilde{X}_{\phi_N(s)}^N) - b(s, \tilde{X}_{\phi_N(s)}^N, \rho_s(\tilde{X}_{\phi_N(s)}^N))| ds \\ & \leq \tilde{\mathbb{E}} \int_h^t \mathbf{1}_{|\tilde{X}_{\phi_N(s)}^N| \leq R} |B_\varepsilon(s, \tilde{X}_{\phi_N(s)}^N) - b(s, \tilde{X}_{\phi_N(s)}^N, \rho_s(\tilde{X}_{\phi_N(s)}^N))| ds \\ & \quad + 2\|b\|_\infty \int_h^t \tilde{\mathbb{P}}(|\tilde{X}_{\phi_N(s)}^N| > R) ds =: I_R^N(\varepsilon) + J_R^N. \end{aligned}$$

For  $I_R^N(\varepsilon)$ , by (3.4), (2.9) and Hölder's inequality with  $p > 2d$  and  $q = \frac{p}{p-1}$ , we have

$$\begin{aligned} I_R^N(\varepsilon) &= \int_h^t \int_{B_R} |B_\varepsilon(s, y) - b(s, y, \rho_s(y))| \rho_{\phi_N(s)}^N(y) dy ds \\ &\lesssim \int_h^t \int_{B_R} |B_\varepsilon(s, y) - b(s, y, \rho_s(y))| \int_{\mathbb{R}^d} g(4\phi_N(s), x - y) \nu_0(dx) dy ds \\ &\lesssim \int_h^t \left( \int_{B_R} |B_\varepsilon(s, y) - b(s, y, \rho_s(y))|^p dy \right)^{\frac{1}{p}} \\ &\quad \times \left( \int_{B_R} \left| \int_{\mathbb{R}^d} g(4\phi_N(s), x - y) \nu_0(dx) \right|^q dy \right)^{\frac{1}{q}} ds \\ &\lesssim \int_h^t \left( \int_{B_R} |B_\varepsilon(s, y) - b(s, y, \rho_s(y))|^p dy \right)^{\frac{1}{p}} \phi_N(s)^{-\frac{d}{p}} ds \\ &\lesssim \left( \int_h^t \left( \int_{B_R} |B_\varepsilon(s, y) - b(s, y, \rho_s(y))|^p dy \right)^{\frac{2}{p}} ds \right)^{\frac{1}{2}} \left( \int_h^t \phi_N(s)^{-\frac{2d}{p}} ds \right)^{\frac{1}{2}} \\ &\lesssim \left( \int_0^t \left( \int_{B_R} |B_\varepsilon(s, y) - b(s, y, \rho_s(y))|^p dy \right)^{\frac{2}{p}} ds \right)^{\frac{1}{2}} \left( \int_0^t s^{-\frac{2d}{p}} ds \right)^{\frac{1}{2}}, \end{aligned}$$

where the implicit constant in the above  $\lesssim$  is independent of  $N, R$  and  $\varepsilon$ . Hence, for each  $R > 0$ , by the dominated convergence theorem, we obtain

$$\lim_{\varepsilon \rightarrow 0} \sup_{N \in \mathbb{N} \cup \{\infty\}} I_R^N(\varepsilon) = 0.$$

For  $J_R^N$ , by Chebyshev's inequality and (3.1) and since  $\|b^N\|_\infty \leq \|b\|_\infty$ , we have

$$\begin{aligned} J_R^N &= 2\|b\|_\infty \int_h^t \mathbb{P}(|X_{\phi_N(s)}^N| > R) ds \\ &\leq 2\|b\|_\infty \int_0^t \mathbb{P}(|X_0| + s\|b\|_\infty + \sqrt{2}|W_{\phi_N(s)}| > R) ds \\ &\leq 2\|b\|_\infty \left( \int_0^t \mathbb{P}(|X_0| + s\|b\|_\infty > R/2) ds + \int_0^t \frac{2\phi_N(s)}{(R/2)^2} ds \right), \end{aligned}$$

which converges to zero uniformly in  $N$ , as  $R \rightarrow \infty$ . Combining the above calculations, we obtain

$$\lim_{\varepsilon \rightarrow 0} \sup_{N \in \mathbb{N} \cup \{\infty\}} \tilde{\mathbb{E}} \int_h^t |B_\varepsilon(s, \tilde{X}_{\phi_N(s)}^N) - b(s, \tilde{X}_{\phi_N(s)}^N, \rho_s(\tilde{X}_{\phi_N(s)}^N))| ds = 0.$$

Thus, (3.10) is proven and the existence of a solution to DDSDE (1.1) is obtained.

**(Uniqueness)** Let  $X_t$  and  $\bar{X}_t$  be two solutions of DDSDE (1.1) defined on the same probability space and with the same initial value  $X_0$ , where  $X_0$  has the distributional density  $\rho_0 \in L^q(\mathbb{R}^d)$  with  $q \in (d, \infty]$ . Let  $\rho_t(y)$  and  $\bar{\rho}_t(y)$  be the distributional density of  $X_t$  and  $\bar{X}_t$ , respectively. Clearly, these are two solutions of the nonlinear Fokker-Planck equation (1.2) with the same initial value  $\rho_0$ . Consider the following linearized SDE:

$$dX_t = B(t, X_t)dt + \sqrt{2}dW_t, \quad X_0 = x,$$

where  $B(t, x) := b(t, x, \rho_t(x))$ . It is well known that  $X_t(x)$  admits a density  $p_x(t, y)$  with Gaussian type estimate: For some  $\lambda, C > 0$ , it holds that for all  $t \in (0, T]$  and  $x, y \in \mathbb{R}^d$ ,

$$p_x(t, y) \leq Cg(\lambda t, x - y).$$

Note that by (2.9) and Hölder's inequality,

$$\begin{aligned} \rho_t(y) &= \int_{\mathbb{R}^d} p_x(t, y) \rho_0(x) dx \lesssim \int_{\mathbb{R}^d} g(\lambda t, x - y) \rho_0(x) dx \\ &\leq \|g(\lambda t, \cdot)\|_{q/(q-1)} \|\rho_0\|_q \lesssim t^{-d/(2q)} \|\rho_0\|_q. \end{aligned} \quad (3.13)$$

Let

$$\Gamma_t := \rho_t - \bar{\rho}_t, \quad B_t := b(t, \cdot, \rho_t) \rho_t - b(t, \cdot, \bar{\rho}_t) \bar{\rho}_t$$

and

$$\Gamma_t^\varepsilon(x) := \Gamma_t * K_\varepsilon(x),$$

where  $\{K_\varepsilon(x), \varepsilon \in (0, 1)\}$  is a family of mollifiers. By definition, it is easy to see that

$$\Gamma_t^\varepsilon = \int_0^t \Delta \Gamma_s^\varepsilon ds + \int_0^t \operatorname{div}(B_s * K_\varepsilon) ds.$$

Let  $\beta_\delta(r) := \sqrt{r^2 + \delta} - \sqrt{\delta}$ . For simplicity, we write  $\int f$  for  $\int_{\mathbb{R}^d} f(x) dx$ . By the chain rule and integration by parts, we have

$$\begin{aligned} \partial_t \int \beta_\delta(\Gamma_t^\varepsilon) &= \int \beta'_\delta(\Gamma_t^\varepsilon) (\Delta \Gamma_t^\varepsilon + \operatorname{div}(B_t * K_\varepsilon)) \\ &= \int \beta''_\delta(\Gamma_t^\varepsilon) \left( -|\nabla \Gamma_t^\varepsilon|^2 - (B_t * K_\varepsilon) \cdot \nabla \Gamma_t^\varepsilon \right) \\ &\leq \int \beta''_\delta(\Gamma_t^\varepsilon) \left( -\frac{1}{2} |\nabla \Gamma_t^\varepsilon|^2 + 2|B_t * K_\varepsilon|^2 \right), \end{aligned}$$

where we have used that  $\beta''_\delta(r) = \delta/(r^2 + \delta)^{3/2} > 0$ . Hence, by Fatou's lemma,

$$\int |\Gamma_t| \leq \lim_{\delta \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \int \beta_\delta(\Gamma_t^\varepsilon) \leq 2 \lim_{\delta \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \int_0^t \int \beta''_\delta(\Gamma_s^\varepsilon) (|B_s| * K_\varepsilon)^2. \quad (3.14)$$

For fixed  $\delta > 0$ , since  $\beta''_\delta(r) \leq \delta^{-1/2}$  and by (3.13),

$$|B_t(x)| \leq \|b\|_\infty (\rho_t(x) + \bar{\rho}_t(x)) \leq C(t^{-d/(2q)} + 1),$$

the dominated convergence theorem implies

$$\lim_{\varepsilon \rightarrow 0} \int_0^t \int \beta''_\delta(\Gamma_s^\varepsilon) (|B_s| * K_\varepsilon)^2 - |B_s|^2 \leq \delta^{-\frac{1}{2}} \int_0^t \lim_{\varepsilon \rightarrow 0} \int (|B_s| * K_\varepsilon)^2 - |B_s|^2 = 0$$

and also,

$$\lim_{\varepsilon \rightarrow 0} \int_0^t \int |\beta''_\delta(\Gamma_s^\varepsilon) - \beta''_\delta(\Gamma_s)| \cdot |B_s|^2 = \int_0^t \int \lim_{\varepsilon \rightarrow 0} |\beta''_\delta(\Gamma_s^\varepsilon) - \beta''_\delta(\Gamma_s)| \cdot |B_s|^2 = 0.$$

Therefore,

$$\lim_{\varepsilon \rightarrow 0} \int_0^t \int \beta''_\delta(\Gamma_s^\varepsilon) (|B_s| * K_\varepsilon)^2 = \int_0^t \int \beta''_\delta(\Gamma_s) |B_s|^2,$$

and by (3.14), thanks to  $\beta''_\delta(r) \leq 1/r$ ,

$$\int |\Gamma_t| \leq 2 \lim_{\delta \rightarrow 0} \int_0^t \int \beta''_\delta(\Gamma_s) |B_s|^2 \leq 2 \int_0^t \int \frac{|B_s|^2}{|\Gamma_s|}.$$

Moreover, noting that by (1.7) and (3.13),

$$|B_t(x)| \leq C|\Gamma_t(x)|(\|\rho_t\|_\infty + 1) \leq C|\Gamma_t(x)|(t^{-\frac{d}{2q}} + 1),$$

we further have for  $\gamma \in (1, \frac{q}{d})$ ,

$$\int |\Gamma_t| \lesssim \int_0^t (s^{-\frac{d}{q}} + 1) \int |\Gamma_s| \leq \left( \int_0^t (s^{-\frac{d}{q}} + 1)^\gamma ds \right)^{\frac{1}{\gamma}} \left( \int_0^t \left[ \int |\Gamma_s| \right]^{\frac{\gamma}{\gamma-1}} ds \right)^{\frac{\gamma-1}{\gamma}}.$$

Thus by Gronwall's inequality, we get

$$\int |\Gamma_t| \equiv 0,$$

which implies  $\rho_t = \bar{\rho}_t$ . Now the pathwise uniqueness of SDE (1.1) follows by the well-known pathwise uniqueness for SDE (1.1) with bounded measurable drift  $b(t, x, \rho_t(x))$  (cf. [12]).  $\square$

#### REFERENCES

- [1] V. Barbu and M. Röckner, Probabilistic representation for solutions to nonlinear Fokker-Planck equations. *SIAM J. Math. Anal.*, **50** (2018), 4246-4260.
- [2] V. Barbu and M. Röckner, From nonlinear Fokker-Planck equations to solutions of distribution dependent SDE. *Annals of Probability*, **48** (2020), 1902-1920.
- [3] V. Barbu and M. Röckner, Uniqueness for nonlinear Fokker-Planck equations and weak uniqueness for McKean-Vlasov SDEs. arXiv:1909.04464.
- [4] V. Barbu and M. Röckner, Solutions for nonlinear Fokker-Planck equations with measures as initial data and McKean-Vlasov equations. arXiv:2005.02311.
- [5] Z.-Q. Chen, E. Hu, L. Xie, and X. Zhang, Heat kernels for non-symmetric diffusion operators with jumps. *J. Differential Equations*, **263** (2017), 6576-6634.
- [6] G.Q. Chen and B. Perthame, Well-posedness for nonisotropic degenerate parabolic-hyperbolic equations. *Ann. Inst. H. Poincaré (C) Non Linear Analysis*, **20** (2003), 645-668.
- [7] A. Figalli, Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *J. Funct. Anal.*, **254** (2008), 109-153.
- [8] V. Lemaire and S. Menozzi, On some Non Asymptotic Bounds for the Euler Scheme. *Electron. J. Probab.*, **15** (2010), No. 53, 1645-1681.
- [9] T.G. Kurtz, Martingale problems for conditional distributions of Markov processes. *Elec. J. Probab.*, **3** (1998), 1-29.
- [10] S. Menozzi, A. Pesce, and X. Zhang, Density and gradient estimates for non degenerate Brownian SDEs with unbounded measurable drift. arXiv:Preprint.
- [11] D. Trevisan, Well-posedness of multidimensional diffusion process with weakly differentiable coefficients. *Electron. J. Probab.*, **21** (2016).
- [12] A Veretennikov, On the strong solutions of stochastic differential equations. *Theory Probab. Appl.*, **24** (1979), 354-366.
- [13] X. Zhang, A discretized version of Krylov's estimate and its applications. *Electron. J. Probab.*, **24** (2019), 1-17.

ZIMO HAO: SCHOOL OF MATHEMATICS AND STATISTICS, WUHAN UNIVERSITY, WUHAN, HUBEI 430072, P.R.CHINA, EMAIL: ZIMOHAO@WHU.EDU.CN

MICHAEL RÖCKNER: FAKULTÄT FÜR MATHEMATIK, UNIVERSITÄT BIELEFELD, 33615, BIELEFELD, GERMANY, EMAIL: ROECKNER@MATH.UNI-BIELEFELD.DE

XICHENG ZHANG: SCHOOL OF MATHEMATICS AND STATISTICS, WUHAN UNIVERSITY, WUHAN, HUBEI 430072, P.R.CHINA, EMAIL: XICHENGZHANG@GMAIL.COM