

# Uniqueness of degenerate Fokker–Planck equations with weakly differentiable drift whose gradient is given by a singular integral\*

Dejun Luo<sup>†</sup>

## Abstract

In this paper we prove the uniqueness of solutions to degenerate Fokker–Planck equations with bounded coefficients, under the additional assumptions that the diffusion coefficient has  $W_{loc}^{1,2}$  regularity, while the gradient of the drift coefficient is merely given by a singular integral.

**Keywords:** Fokker–Planck equation ; martingale solution ; maximal function ; singular integral operator.

**AMS MSC 2010:** 35Q84 ; 60H10.

Submitted to ECP on March 26, 2014, final version accepted on July 3, 2014.

## 1 Introduction

This short note is motivated by the work of Röckner and Zhang [21], where they proved the uniqueness of solutions to degenerate Fokker–Planck equations with bounded coefficients, satisfying a pointwise inequality. Before going to the details, we first introduce some notations. Let  $\sigma : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^m$  and  $b : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  be measurable functions. Define the second order differential operator

$$L_t \varphi(x) = \frac{1}{2} \sum_{i,j=1}^d \sum_{k=1}^m \sigma_t^{ik}(x) \sigma_t^{jk}(x) \partial_{ij} \varphi(x) + \sum_{i=1}^d b_t^i(x) \partial_i \varphi(x), \quad \varphi \in C_c^\infty(\mathbb{R}^d), \quad (1.1)$$

where  $\partial_i \varphi(x) = \frac{\partial \varphi}{\partial x_i}(x)$  and  $\partial_{ij} \varphi(x) = \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(x)$ ,  $1 \leq i, j \leq d$ . We consider the Fokker–Planck equation

$$\partial_t \mu_t = L_t^* \mu_t, \quad \mu|_{t=0} = \mu_0, \quad (1.2)$$

where  $L_t^*$  is the adjoint operator of  $L_t$ . Here is the rigorous meaning of this equation: for any  $\varphi \in C_c^\infty(\mathbb{R}^d)$ ,

$$\frac{d}{dt} \int_{\mathbb{R}^d} \varphi(x) d\mu_t(x) = \int_{\mathbb{R}^d} L_t \varphi(x) d\mu_t(x),$$

where the initial condition means that  $\mu_t$  weakly\* converges to  $\mu_0$  as  $t$  tends to 0. If  $\mu_t$  is absolutely continuous with respect to the Lebesgue measure with the density function  $u_t$  for all  $t \in [0, T]$ , then the density function  $u_t$  solves the PDE below in the weak sense:

$$\partial_t u_t = L_t^* u_t, \quad u|_{t=0} = u_0. \quad (1.3)$$

---

\*Support: Key Lab of RCSDS, CAS (No. 2008DP173182), NSFC (No. 11101407) and AMSS (Y129161ZZ1).

<sup>†</sup>Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China. E-mail: luodj@amss.ac.cn

We can now recall the main result of Röckner and Zhang [21]. They assume that the coefficients  $\sigma$  and  $b$  are bounded and for any  $R > 0$  and a.e.  $x, y \in B(R) := \{z \in \mathbb{R}^d : |z| \leq R\}$ ,

$$2\langle x - y, b_t(x) - b_t(y) \rangle + \|\sigma_t(x) - \sigma_t(y)\|^2 \leq (f_{R,t}(x) + f_{R,t}(y))|x - y|^2, \quad (1.4)$$

where  $f_R \in L^q([0, T] \times B(R))$  for some  $q \geq 1$ . Under these conditions, they proved the uniqueness of solutions to (1.3), in an integrability class depending on  $q$ , with probability density  $\rho$  as the initial value  $u_0$ . Their method is based on the natural connection between Fokker–Planck equations and stochastic differential equations (SDE), see Subsection 2.1 for more details. We mention that (1.4) is satisfied when  $b_t \in W_{loc}^{1,q}$  and  $\sigma_t \in W_{loc}^{1,2\vee q}$  with  $q > 1$  for a.e.  $t \in [0, T]$ , but is in general not so when  $q = 1$ . Our purpose in this work is to generalize Röckner and Zhang’s result to cover the case that  $b_t \in W_{loc}^{1,1}$ . Indeed, by employing Bouchut and Crippa’s estimate (see Theorem 2.15 of the current paper), we can treat more general situation where the drift coefficient  $b$  has a gradient given by a singular integral.

Here are our assumptions on the coefficients  $\sigma$  and  $b$ .

**Assumption 1.1.** *Assume that*

- (H1) *the functions  $\sigma$  and  $b$  are essentially bounded;*
- (H2)  *$\sigma \in L^2([0, T], W_{loc}^{1,2}(\mathbb{R}^d))$ ;*
- (H3) *for a.e.  $t \in (0, T)$  and for every  $i, j = 1, \dots, d$ , we have*

$$\partial_j b_t^i = \sum_{k=1}^{m_0} S_{jk}^i(g_{jk}^i(t)) \quad \text{holds in } \mathcal{D}'(\mathbb{R}^d); \quad (1.5)$$

where  $S_{jk}^i$  are singular integral operators of fundamental type in  $\mathbb{R}^d$  (see Definition 2.13 for the precise meaning) and the functions  $g_{jk}^i \in L^1((0, T) \times \mathbb{R}^d)$  for all  $i, j = 1, \dots, d$  and  $k = 1, \dots, m_0$ . In vectorial form, the above identity can be written as

$$\partial_j b_t = \sum_{k=1}^{m_0} S_{jk}(g_{jk}(t)) \quad \text{holds in } \mathcal{D}'(\mathbb{R}^d), \quad \text{for a.e. } t \in (0, T), \quad (1.6)$$

in which  $S_{jk}$  is a vector consisting of  $d$  singular integral operators and for each  $j = 1, \dots, d$  and  $k = 1, \dots, m_0$ , we have  $g_{jk} \in L^1((0, T) \times \mathbb{R}^d, \mathbb{R}^d)$ .

Some comments on the assumptions are in order. We assume  $\sigma$  and  $b$  are bounded because we shall make use of a representation formula by Figalli (see [16, Theorem 2.6] or Theorem 2.5 below), where such boundedness condition are imposed on the coefficients. The assumption (H2) on  $\sigma$  is natural, and it has already been used in [18, 21, 20].

The motivation for considering the condition (H3) on the drift  $b$  comes from the recent developments in the DiPerna–Lions theory, especially the papers [9, 10] by Bouchut and Crippa, where the authors established the existence and uniqueness of flows associated to such vector field  $b$ . This theory has its origin in the celebrated work of DiPerna and Lions [13], who proved that if  $b$  is a  $W_{loc}^{1,1}$  vector field with bounded divergence, then there exists a unique flow of measurable maps generated by  $b$  which leaves the Lebesgue measure quasi-invariant. Ambrosio [1] extended the main result in [13] to the case where the vector field has only BV spatial regularity, see [2, 3] for more details. In the recent preprint [5], Ambrosio and Trevisan developed the DiPerna–Lions theory in a rather general setting, that is, on metric measure spaces. This theory is indirect in the sense that the authors first established the well-posedness of the corresponding first order linear PDE (transport equation or continuity equation), from which

they deduced the results on ODE. See [4, 14] for the developments in the infinite dimensional Wiener space. Crippa and De Lellis [12] obtained some a-priori estimates on the flow in the Lagrangian formulation, which enables them to give a direct construction of the flow (see [23, 15] for the extension to the stochastic setting). While this approach works very well when the vector field  $b$  has  $W_{loc}^{1,p}$  regularity with  $p > 1$ , it is not so for the case  $p = 1$ . This motivates Bouchut and Crippa to further develop the direct method to cover the case  $b \in W_{loc}^{1,1}$ . Indeed, they are able to deal with more general vector fields  $b$  whose gradient is given by a singular integral, cf. [10]. Remark that this family of functions include the Sobolev space  $W^{1,1}$ , but does not contain the  $BV$  class, nor is contained in it.

We can now state the main result of this paper.

**Theorem 1.2** (Uniqueness of Fokker–Planck equations). *Under the assumptions (H1)–(H3), for any given probability density function  $\rho$  on  $\mathbb{R}^d$ , there is at most one weak solution  $u_t$  to the Fokker–Planck equation (1.3) in the class  $L^\infty([0, T], L^1 \cap L^\infty(\mathbb{R}^d))$  with  $u_0 = \rho$ .*

We recall some known results concerning the uniqueness of Fokker–Planck equations. Let  $\mathcal{P}(\mathbb{R}^d)$  be the set of probability measures on  $\mathbb{R}^d$ . In the *non-degenerate* case, it was shown in [6] that if in addition the diffusion coefficient  $\sigma$  is Lipschitz continuous and the drift vector field  $b$  is locally integrable and coercive, then the uniqueness holds for (1.2) in  $\mathcal{P}(\mathbb{R}^d)$  when the initial measure has finite entropy. On the other hand, Le Bris and Lions [18] established the well-posedness of degenerate Fokker–Planck type equations with coefficients fulfilling quite general Sobolev regularity, by extending the DiPerna–Lions theory to this setting. In [20], we slightly generalize the main result in [18] to the case where the drift  $b$  has only  $BV$  spatial regularity, in the spirit of [1]. The study of Fokker–Planck equations in the infinite dimensional setting can be found in [7, 19]. Bogachev et al. considered in the recent paper [8] a class of second order differential operators in divergence form, whose diffusion coefficient  $\sigma$  is written as the product of a nonnegative function  $\varrho$  (possibly unbounded and non-smooth) and a positive definite matrix  $A$ . They proved the uniqueness of solutions to (1.3) in a suitable class, provided that  $A$  is bounded and Lipschitzian, and the vector field  $b$  in the drift coefficient  $\sqrt{\varrho}b$  is bounded too. We stress that, in Theorem 1.2, we require neither the non-degeneracy condition nor Lipschitz continuity on  $\sigma$ , and the drift  $b$  has only very weak differentiability which is not included in the  $BV$  class.

**Remark 1.3.** *Before finishing this section, we give the following two remarks:*

- (i) *This paper is only concerned with the uniqueness of solutions to the Fokker–Planck equation (1.3). To show the existence of solutions to (1.3), one usually needs some assumptions on the divergence of the coefficients, for instance  $[\operatorname{div}(b)]^- \in L^1([0, T], L^\infty(\mathbb{R}^d))$ . Under such conditions, one can prove some a-priori estimates on the solution  $u$  to (1.3), see e.g. [18, Section 5.2, p.1289] for more details.*
- (ii) *The proof of Theorem 1.2 follows the line of arguments in [21, Theorem 1.1]. A close look at the proof reveals that this method allows us to prove the pathwise uniqueness of solutions to SDE (2.1), once we have some a-priori estimates on the distributions of solutions, cf. [11, Theorem 1.1].*

This paper is organized as follows. In Section 2, we first recall some well known results on the connection between Fokker–Planck equations and SDEs, then we introduce the pointwise estimate of weakly differentiable functions with gradient given by a singular integral. Finally we prove in Section 3 our main result by following the arguments in [21, 10].

## 2 Preliminary results

In this section we recall some known results which are necessary for proving our main result.

### 2.1 Connection between Fokker–Planck equations and SDEs

This subsection mainly follows the beginning parts of [21, Sections 1 and 2]. We first introduce some notations. Denote by  $\mathbb{W}_T^m = C([0, T]; \mathbb{R}^m)$  the space of continuous functions from  $[0, T]$  to  $\mathbb{R}^m$ . Let  $\mathcal{F}_t^m$  be the canonical filtration generated by coordinate process  $W_t(w) = w_t$ ,  $w \in \mathbb{W}_T^m$ . We write  $\nu$  for the standard Wiener measure on  $(\mathbb{W}_T^m, \mathcal{F}_T^m)$  so that  $t \rightarrow W_t(w)$  is an  $m$ -dimensional standard Brownian motion.

Given bounded measurable functions  $\sigma : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d \otimes \mathbb{R}^m$  and  $b : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ , we consider the Itô SDE

$$dX_t = \sigma_t(X_t) dW_t + b_t(X_t) dt, \quad X|_{t=0} = X_0. \tag{2.1}$$

Let  $\mu_t$  be the distribution of  $X_t$ . Then it is well known that, by Itô’s formula,  $\mu_t$  is a distributional solution to the Fokker–Planck equation (1.2).

Recall that  $\mathcal{P}(\mathbb{R}^d)$  is the set of probability measures on  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ . Here are two well known notions of solutions to (2.1) in the theory of SDEs, which are stated in detail to fix notations.

**Definition 2.1** (Martingale solution). *Given  $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$ , a probability measure  $P_{\mu_0}$  on  $(\mathbb{W}_T^d, \mathcal{F}_T^d)$  is called a martingale solution to SDE (2.1) with initial distribution  $\mu_0$  if  $P_{\mu_0} \circ w_0^{-1} = \mu_0$ , and for any  $\varphi \in C_c^\infty(\mathbb{R}^d)$ ,  $\varphi(w_t) - \varphi(w_0) - \int_0^t L_s \varphi(w_s) ds$  is an  $(\mathcal{F}_t^d)$ -martingale under  $P_{\mu_0}$ .*

**Definition 2.2** (Weak solution). *Let  $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$ . The SDE (2.1) is said to have a weak solution with initial law  $\mu_0$  if there exist a filtered probability space  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{0 \leq t \leq T}, P)$ , on which are defined a  $(\mathcal{G}_t)$ -adapted continuous process  $X_t$  taking values in  $\mathbb{R}^d$  and an  $m$ -dimensional standard  $(\mathcal{G}_t)$ -Brownian motion  $W_t$ , such that  $X_0$  is distributed as  $\mu_0$  and a.s.,*

$$X_t = X_0 + \int_0^t \sigma_s(X_s) dW_s + \int_0^t b_s(X_s) ds, \quad \forall t \in [0, T].$$

We denote this solution by  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{0 \leq t \leq T}, P; X, W)$ .

The next result can be found in the proof of [17, Chap. IV, Theorem 1.1].

**Proposition 2.3.** *Given two weak solutions  $(\Omega^{(i)}, \mathcal{G}^{(i)}, (\mathcal{G}_t^{(i)})_{0 \leq t \leq T}, P^{(i)}; X^{(i)}, W^{(i)})$ ,  $i = 1, 2$  to SDE (2.1), having the same initial law  $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$ , there exist a filtered probability space  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{0 \leq t \leq T}, P)$ , a standard  $m$ -dimensional  $(\mathcal{G}_t)$ -Brownian motion  $W_t$  and two  $\mathbb{R}^d$ -valued continuous  $(\mathcal{G}_t)$ -adapted processes  $Y^{(i)}$ ,  $i = 1, 2$ , such that  $P(Y_0^{(1)} = Y_0^{(2)}) = 1$  and for  $i = 1, 2$ ,  $X^{(i)}$  and  $Y^{(i)}$  have the same distributions in  $\mathbb{W}_T^d$ , and  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{0 \leq t \leq T}, P; Y^{(i)}, W)$  is a weak solution of SDE (2.1).*

The assertion below is a special case of [17, Chap. IV, Proposition 2.1].

**Proposition 2.4** (Existence of martingale solution implies that of weak solution). *Let  $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$  and  $P_{\mu_0}$  be a martingale solution of SDE (2.1). Then there exists a weak solution  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{0 \leq t \leq T}, P; X, W)$  to SDE (2.1) such that  $P \circ X^{-1} = P_{\mu_0}$ .*

Finally we remind the following result which is an easy consequence of Figalli’s representation theorem (see [16, Theorem 2.6]) for solutions to the Fokker–Planck equation (1.2).

**Theorem 2.5.** *Assume that  $\sigma$  and  $b$  are two bounded measurable functions. Given  $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$ , let  $\mu_t \in \mathcal{P}(\mathbb{R}^d)$  be a measure-valued solution to equation (1.2) with initial value  $\mu_0$ . Then there exists a martingale solution  $P_{\mu_0}$  to SDE (2.1) with initial law  $\mu_0$  such that for all  $\varphi \in C_c^\infty(\mathbb{R}^d)$ , one has*

$$\int_{\mathbb{R}^d} \varphi(x) d\mu_t(x) = \int_{\mathbb{W}_T^d} \varphi(w_t) dP_{\mu_0}(w), \quad \forall t \in [0, T].$$

## 2.2 Elements from harmonic analysis and Bouchut and Crippa’s estimate

In this subsection we first recall some basic facts in harmonic analysis, and then we introduce the pointwise estimate of Bouchut and Crippa on weakly differentiable functions whose gradients are given by singular integrals. The main reference is [10, Sections 2–4].

### 2.2.1 Weak Lebesgue spaces

Denote by  $\mathcal{L}^d$  the Lebesgue measure on  $\mathbb{R}^d$ , and  $B(R)$  the ball in  $\mathbb{R}^d$  centered at the origin with radius  $R$ .

**Definition 2.6.** *Let  $O \subset \mathbb{R}^d$  be an open set and  $u$  a measurable function (possibly vector valued) defined on  $O$ . For any  $p \in [1, \infty)$ , define*

$$\| \| u \| \|_{M^p(O)}^p = \sup_{\lambda > 0} \{ \lambda^p \mathcal{L}^d(\{x \in O : |u(x)| > \lambda\}) \}, \quad (2.2)$$

and denote by  $M^p(O)$  the totality of measurable functions  $u$  defined on  $O$  such that  $\| \| u \| \|_{M^p(O)} < \infty$ .  $M^p(O)$  is called the weak Lebesgue space. For  $p = \infty$ , we set  $M^\infty(O) = L^\infty(O)$  by convention.

It is worth mentioning that  $M^p(O)$  is not a Banach space, since  $\| \| \cdot \| \|_{M^p(O)}$  is not subadditive and hence not a norm. From the simple inequality below

$$\lambda^p \mathcal{L}^d(\{x \in O : |u(x)| > \lambda\}) \leq \int_{\{|u|>\lambda\}} |u(x)|^p dx \leq \| \| u \| \|_{L^p(O)}^p,$$

we conclude that  $L^p(O) \subset M^p(O)$  and  $\| \| u \| \|_{M^p(O)} \leq \| \| u \| \|_{L^p(O)}$ .

The following result (see [10, Lemma 2.2] for its proof) concerning the interpolation between  $M^1$  and  $M^p$  ( $p > 1$ ) is one of the key ingredient in the proof of Section 3.

**Lemma 2.7.** *Let  $O \subset \mathbb{R}^d$  be a set with finite Lebesgue measure and  $u : O \rightarrow \mathbb{R}_+$  a nonnegative measurable function. Then for any  $p \in (1, \infty)$ , it holds*

$$\| \| u \| \|_{L^1(O)} \leq \frac{p}{p-1} \| \| u \| \|_{M^1(O)} \left[ 1 + \log \left( \frac{\| \| u \| \|_{M^p(O)}}{\| \| u \| \|_{M^1(O)}} \mathcal{L}^d(O)^{1-\frac{1}{p}} \right) \right], \quad (2.3)$$

and for  $p = \infty$ ,

$$\| \| u \| \|_{L^1(O)} \leq \| \| u \| \|_{M^1(O)} \left[ 1 + \log \left( \frac{\| \| u \| \|_{L^\infty(O)}}{\| \| u \| \|_{M^1(O)}} \mathcal{L}^d(O) \right) \right]. \quad (2.4)$$

### 2.2.2 Maximal functions

We first introduce the notion of local maximal functions. Let  $R > 0$  and  $u : \mathbb{R}^d \rightarrow \mathbb{R}$  be a measurable function. Set for  $x \in \mathbb{R}^d$

$$M_R u(x) = \sup_{0 < r \leq R} \int_{B(x,r)} |u(y)| dy = \sup_{0 < r \leq R} \frac{1}{\mathcal{L}^d(B(x,r))} \int_{B(x,r)} |u(y)| dy, \quad (2.5)$$

where  $B(x, r)$  is the ball centered at  $x$  of radius  $r > 0$ . The following properties of the local maximal function are well known, see for instance [12, Lemmas A.2 and A.3]. In the sequel,  $C$  with subscripts  $d, p$  and so on means it is a positive constant depending on these parameters.

**Proposition 2.8.** *Fix any  $R, \rho > 0$ . If  $u \in L^1_{loc}(\mathbb{R}^d)$ , then it holds*

$$\|M_R u\|_{M^1(B(\rho))} \leq C_d \|u\|_{L^1(B(R+\rho))}, \tag{2.6}$$

and if  $u \in L^p_{loc}(\mathbb{R}^d)$  with  $p \in (1, \infty)$ , then

$$\|M_R u\|_{L^p(B(\rho))} \leq C_{d,p} \|u\|_{L^p(B(R+\rho))}. \tag{2.7}$$

Moreover, if  $u$  belongs to the Sobolev space  $W^{1,1}_{loc}(\mathbb{R}^d)$ , then there exist a constant  $C_d > 0$  and a negligible set  $N \subset \mathbb{R}^d$  such that for all  $x, y \in N^c$  with  $|x - y| \leq R$ , one has

$$|u(x) - u(y)| \leq C_d |x - y| (M_R |\nabla u|(x) + M_R |\nabla u|(y)). \tag{2.8}$$

We shall also need the so-called grand maximal function which is an important tool in the theory of Hardy spaces. Denote by  $L^\infty_c(\mathbb{R}^d)$  the space of bounded functions with compact support.

**Definition 2.9** (Grand maximal function). *Given a family of functions  $\{\rho^\alpha\}_\alpha \subset L^\infty_c(\mathbb{R}^d)$  and  $u \in L^1_{loc}(\mathbb{R}^d)$ , we define the grand maximal function of  $u$  relative to  $\{\rho^\alpha\}_\alpha$  as*

$$M_{\{\rho^\alpha\}} u(x) = \sup_\alpha \sup_{\varepsilon > 0} |(\rho^\alpha_\varepsilon * u)(x)| = \sup_\alpha \sup_{\varepsilon > 0} \left| \int_{\mathbb{R}^d} \rho^\alpha_\varepsilon(x - y) u(y) dy \right|, \tag{2.9}$$

where  $\rho^\alpha_\varepsilon(x) = \varepsilon^{-d} \rho^\alpha(\varepsilon^{-1}x)$ ,  $x \in \mathbb{R}^d$ . When the family  $\{\rho^\alpha\}_\alpha \subset C^\infty_c(\mathbb{R}^d)$ , the space of smooth functions with compact support, the same definition applies for distributions  $u \in \mathcal{D}'(\mathbb{R}^d)$ , more precisely,

$$M_{\{\rho^\alpha\}} u(x) = \sup_\alpha \sup_{\varepsilon > 0} |\langle u, \rho^\alpha_\varepsilon(x - \cdot) \rangle|.$$

**Remark 2.10.** *Here are two comments on the above definition.*

- (i) *Compared to the definition (2.5) of the local maximal function, we move the absolute value outside the integral sign. This allows some kind of cancellation effect when the grand maximal function is composed with the singular integral operator, see [10, Section 3] for more details.*
- (ii) *Taking  $\rho^\alpha(x) = [\mathcal{L}^d(B(1))]^{-1} \mathbf{1}_{B(1)}(x)$  and replacing  $\sup_{\varepsilon > 0}$  by  $\sup_{0 < \varepsilon \leq R}$  in (2.9), we get the local maximal function  $M_R u(x)$  defined in (2.5), except that the absolute value is outside the integral sign.*

### 2.2.3 Singular integral operators

We now recall some facts on singular kernels and singular integral operators, see [22, Chap. II] for details. Let  $\mathcal{S}(\mathbb{R}^d)$  be the Schwartz space and  $\mathcal{S}'(\mathbb{R}^d)$  the space of tempered distributions.

**Definition 2.11** (Singular kernel). *We call  $K$  a singular kernel on  $\mathbb{R}^d$  if*

- (i)  *$K \in \mathcal{S}'(\mathbb{R}^d)$  and its Fourier transform  $\hat{K} \in L^\infty(\mathbb{R}^d)$ ;*
- (ii) *the restriction  $K|_{\mathbb{R}^d \setminus \{0\}}$  of  $K$  outside the origin belongs to  $L^1_{loc}(\mathbb{R}^d \setminus \{0\})$  and there exists a constant  $A \geq 0$  such that*

$$\int_{\{|x| > 2|y|\}} |K(x - y) - K(x)| dx \leq A, \quad \text{for all } y \in \mathbb{R}^d.$$

**Theorem 2.12** (Calderón–Zygmund). *Let  $K$  be a singular kernel. For  $u \in L^2(\mathbb{R}^d)$ , define  $Su = K * u$  in the sense of multiplication in the Fourier variable. Then for every  $p \in (1, \infty)$ , the following strong estimate holds:*

$$\|Su\|_{L^p(\mathbb{R}^d)} \leq C_{d,p}(A + \|\hat{K}\|_{L^\infty})\|u\|_{L^p(\mathbb{R}^d)}, \quad u \in L^p \cap L^2(\mathbb{R}^d); \quad (2.10)$$

when  $p = 1$ , the weak estimate below holds:

$$\|Su\|_{M^1(\mathbb{R}^d)} \leq C_d(A + \|\hat{K}\|_{L^\infty})\|u\|_{L^1(\mathbb{R}^d)}, \quad u \in L^1 \cap L^2(\mathbb{R}^d). \quad (2.11)$$

As a direct consequence of the above theorem, for any  $1 < p < \infty$ , we can extend the domain of  $S$  to the whole  $L^p(\mathbb{R}^d)$  with values in  $L^p(\mathbb{R}^d)$ , and the inequality (2.10) holds for all  $u \in L^p(\mathbb{R}^d)$ ; furthermore,  $S$  can be extended to the whole of  $L^1(\mathbb{R}^d)$  with values in  $M^1(\mathbb{R}^d)$ , and the estimate (2.11) holds for all  $u \in L^1(\mathbb{R}^d)$ . The operator  $S$  constructed in this way is called the *singular integral operator* associated to the singular kernel  $K$ .

Following the terminology of [10], we introduce a special class of singular kernels.

**Definition 2.13** (Singular kernel of fundamental type). *We say that  $K$  is a singular kernel of fundamental type if it possesses the following properties:*

- (i)  $K|_{\mathbb{R}^d \setminus \{0\}} \in C^1(\mathbb{R}^d \setminus \{0\})$ ;
- (ii) there is a positive constant  $C_0$  such that  $|K(x)| \leq C_0/|x|^d$  for all  $x \neq 0$ ;
- (iii) there exists a positive constant  $C_1$  such that  $|\nabla K(x)| \leq C_1/|x|^{d+1}$  for all  $x \neq 0$ ;
- (iv) there is a constant  $A_2 \geq 0$  such that

$$\left| \int_{\{R_1 < |x| < R_2\}} K(x) dx \right| \leq A_2 \quad \text{for all } 0 < R_1 < R_2 < \infty.$$

### 2.2.4 Bouchut and Crippa’s estimate

Now we are ready to introduce the important pointwise estimate of Bouchut and Crippa on weakly differentiable functions whose gradient is given by a singular integral. First of all, we present the following result (cf. [10, Theorem 3.3]) on the cancellation effect between the singular integral and the maximal function introduced in Definition 2.9.

**Theorem 2.14.** *Let  $K$  be a singular kernel of fundamental type as in Definition 2.13 and set  $Su = K * u$  for  $u \in L^2(\mathbb{R}^N)$ . Let  $\{\rho^\alpha\}_\alpha$  be a family of kernels satisfying*

$$\text{supp}(\rho^\alpha) \subset B(1) \quad \text{and} \quad \|\rho^\alpha\|_{L^1(\mathbb{R}^d)} \leq Q_1 \quad \text{for every } \alpha. \quad (2.12)$$

Assume that for every  $\varepsilon > 0$  and every  $\alpha$ , it holds  $(\varepsilon^d K(\varepsilon \cdot)) * \rho^\alpha \in C_b(\mathbb{R}^d)$  with the uniform norm estimate

$$\|(\varepsilon^d K(\varepsilon \cdot)) * \rho^\alpha\|_{C_b(\mathbb{R}^d)} \leq Q_2 \quad \text{for every } \varepsilon > 0 \text{ and every } \alpha. \quad (2.13)$$

Then we have

- (i) there is a dimensional constant  $C_d$  such that for all  $u \in L^1 \cap L^2(\mathbb{R}^d)$ ,

$$\|M_{\{\rho^\alpha\}}(Su)\|_{M^1(\mathbb{R}^d)} \leq C_d[Q_2 + Q_1(C_0 + C_1 + \|\hat{K}\|_{L^\infty})]\|u\|_{L^1(\mathbb{R}^d)}, \quad (2.14)$$

where  $C_0$  and  $C_1$  are constants in Definition 2.13;

- (ii) if  $Q_3 := \sup_\alpha \|\rho^\alpha\|_{L^\infty(\mathbb{R}^d)}$  is finite, then there exists a constant  $C_d$  dependent on  $d$  such that

$$\|M_{\{\rho^\alpha\}}(Su)\|_{L^2(\mathbb{R}^d)} \leq C_d Q_3 \|\hat{K}\|_{L^\infty} \|u\|_{L^2(\mathbb{R}^d)} \quad \text{for all } u \in L^2(\mathbb{R}^d). \quad (2.15)$$

Finally we can introduce Bouchut and Crippa’s pointwise estimate (see [10, Proposition 4.2]).

**Theorem 2.15.** *Let  $u \in L^1_{loc}(\mathbb{R}^d)$  and assume that for every  $j = 1, \dots, d$ , it holds*

$$\partial_j u = \sum_{k=1}^{m_0} S_{jk} g_{jk} \quad \text{in } \mathcal{D}'(\mathbb{R}^d), \tag{2.16}$$

where  $S_{jk}$  are singular integral operators of fundamental type as in Definition 2.13 and  $g_{jk} \in L^1(\mathbb{R}^d)$  for all  $j = 1, \dots, d$  and  $k = 1, \dots, m_0$ . Then there exists a nonnegative function  $U \in M^1(\mathbb{R}^d)$  and a negligible set  $N \subset \mathbb{R}^d$  such that

$$|u(x) - u(y)| \leq |x - y|(U(x) + U(y)) \quad \text{for every } x, y \in \mathbb{R}^d \setminus N. \tag{2.17}$$

Moreover, the function  $U$  is explicitly given by

$$U = \sum_{j=1}^d \sum_{k=1}^{m_0} M_{\{\Lambda^{\xi,j}, \xi \in S^{d-1}\}}(S_{jk} g_{jk}), \tag{2.18}$$

where the maximal function relative to a family of kernels is defined in Definition 2.9, and the functions  $\Lambda^{\xi,j} \in C_c^\infty(\mathbb{R}^d)$  are explicitly defined as

$$\Lambda^{\xi,j}(x) = h\left(\frac{\xi}{2} - x\right) x_j, \quad \xi \in S^{d-1}, j = 1, \dots, d \tag{2.19}$$

and the kernel  $h$  satisfies

$$h \in C_c^\infty(\mathbb{R}^d), \quad \int_{\mathbb{R}^d} h(y) dy = 1 \quad \text{and} \quad \text{supp}(h) \subset B(1/2). \tag{2.20}$$

At the beginning of the proof of [10, Proposition 4.2], it has been checked that Theorem 2.14 now applies to the singular kernels  $S_{jk}$  and the family of mollifiers  $\Lambda^{\xi,j}$ , since they verify the conditions (2.12) and (2.13). We would like to mention that, in Section 3, we actually use the smooth version of the above theorem, that is,  $\{g_{jk} : 1 \leq j \leq d, 1 \leq k \leq m_0\} \subset C^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ . In this case, (2.17) holds for all  $x, y \in \mathbb{R}^d$  (cf. Step 1 of the proof of [10, Proposition 4.2]).

### 3 Proof of the main result

This section is devoted to the proof of Theorem 1.2, which is quite long and will be divided into several steps.

*Proof of Theorem 1.2.* We follow the idea of the proof of [21, Theorem 1.1]. Let  $u_t^{(i)}, i = 1, 2$  be two weak solutions to (1.3) in the class  $L^\infty([0, T], L^1 \cap L^\infty(\mathbb{R}^d))$  with the same initial value  $\rho$ . Set  $d\mu_0(x) = \rho(x) dx$ . Then by Theorem 2.5, there exist two martingale solutions  $P_{\mu_0}^{(i)}, i = 1, 2$  to the SDE (2.1) with the same initial probability distribution  $\mu_0$ , such that for all  $\varphi \in C_c^\infty(\mathbb{R}^d)$ ,

$$\int_{\mathbb{R}^d} \varphi(x) u_t^{(i)}(x) dx = \int_{\mathbb{W}_T^d} \varphi(w_t) dP_{\mu_0}^{(i)}(w), \quad i = 1, 2. \tag{3.1}$$

Applying Proposition 2.4, we obtain two weak solutions  $(\Omega^{(i)}, \mathcal{G}^{(i)}, (\mathcal{G}_t^{(i)})_{0 \leq t \leq T}, P^{(i)}; X^{(i)}, W^{(i)}) (i = 1, 2)$  to SDE (2.1) satisfying  $P^{(i)} \circ (X^{(i)})^{-1} = P_{\mu_0}^{(i)}, i = 1, 2$ . Finally by Proposition 2.3, we can find a common filtered probability space  $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{0 \leq t \leq T}, P)$ , on which

are defined a standard  $m$ -dimensional  $(\mathcal{G}_t)$ -Brownian motion  $W$  and two continuous  $(\mathcal{G}_t)$ -adapted processes  $Y^{(i)}$  ( $i = 1, 2$ ), such that  $P(Y_0^{(1)} = Y_0^{(2)}) = 1$  and  $Y^{(i)}$  is distributed as  $P_{\mu_0}^{(i)}$  on  $W_T^d$ ; moreover for  $i = 1, 2$ , it holds a.s. that

$$Y_t^{(i)} = Y_0^{(i)} + \int_0^t b_s(Y_s^{(i)}) ds + \int_0^t \sigma_s(Y_s^{(i)}) dW_s \quad \text{for all } t \leq T.$$

Set  $Z_t = Y_t^{(1)} - Y_t^{(2)}$  and for  $R > 0$ , define the stopping time  $\tau_R = \inf \{t \in [0, T] : |Y_t^{(1)}| \vee |Y_t^{(2)}| \geq R\}$  with the convention that  $\inf \emptyset = T$ . Since the coefficients  $\sigma$  and  $b$  are bounded, it is clear that

$$\lim_{R \rightarrow \infty} \tau_R(\omega) = T \quad \text{almost surely.} \tag{3.2}$$

Fix  $\delta > 0$ . We have by Itô’s formula that

$$\begin{aligned} \log \left( \frac{|Z_{t \wedge \tau_R}|^2}{\delta^2} + 1 \right) &= \int_0^{t \wedge \tau_R} \frac{2 \langle Z_s, b_s(Y_s^{(1)}) - b_s(Y_s^{(2)}) \rangle + \|\sigma_s(Y_s^{(1)}) - \sigma_s(Y_s^{(2)})\|^2}{|Z_s|^2 + \delta^2} ds \\ &\quad + 2 \int_0^{t \wedge \tau_R} \frac{\langle Z_s, [\sigma_s(Y_s^{(1)}) - \sigma_s(Y_s^{(2)})] dW_s \rangle}{|Z_s|^2 + \delta^2} \\ &\quad - 2 \int_0^{t \wedge \tau_R} \frac{\|[\sigma_s(Y_s^{(1)}) - \sigma_s(Y_s^{(2)})] Z_s\|^2}{(|Z_s|^2 + \delta^2)^2} ds. \end{aligned}$$

Taking expectation on both sides with respect to  $P$  yields

$$\begin{aligned} \mathbb{E} \log \left( \frac{|Z_{t \wedge \tau_R}|^2}{\delta^2} + 1 \right) &\leq 2 \mathbb{E} \int_0^{t \wedge \tau_R} \frac{\langle Z_s, b_s(Y_s^{(1)}) - b_s(Y_s^{(2)}) \rangle}{|Z_s|^2 + \delta^2} ds \\ &\quad + \mathbb{E} \int_0^{t \wedge \tau_R} \frac{\|\sigma_s(Y_s^{(1)}) - \sigma_s(Y_s^{(2)})\|^2}{|Z_s|^2 + \delta^2} ds \\ &=: I_1 + I_2. \end{aligned} \tag{3.3}$$

In the sequel we shall estimate the two terms separately.

*Step 1.* We first deal with the simpler term  $I_2$ . Choose  $\chi \in C_c^\infty(\mathbb{R}^d, \mathbb{R}_+)$  such that  $\text{supp}(\chi) \subset B(1)$  and  $\int_{\mathbb{R}^d} \chi(x) dx = 1$ . For  $\varepsilon \in (0, 1)$  let  $\chi_\varepsilon(x) = \varepsilon^{-d} \chi(x/\varepsilon)$ ,  $x \in \mathbb{R}^d$ . Define  $\sigma_s^\varepsilon = \sigma_s * \chi_\varepsilon$ . Then by (H1), for a.e.  $s \in [0, T]$ ,  $\sigma_s^\varepsilon \in C_b^\infty(\mathbb{R}^d)$  for every  $\varepsilon \in (0, 1)$ . By the triangular inequality, we have

$$\begin{aligned} I_2 &\leq 3 \mathbb{E} \int_0^{t \wedge \tau_R} \frac{\|\sigma_s^\varepsilon(Y_s^{(1)}) - \sigma_s^\varepsilon(Y_s^{(2)})\|^2}{|Z_s|^2 + \delta^2} ds \\ &\quad + 3 \mathbb{E} \int_0^{t \wedge \tau_R} \frac{\|\sigma_s^\varepsilon(Y_s^{(1)}) - \sigma_s(Y_s^{(1)})\|^2 + \|\sigma_s^\varepsilon(Y_s^{(2)}) - \sigma_s(Y_s^{(2)})\|^2}{|Z_s|^2 + \delta^2} ds \\ &=: I_{2,1} + I_{2,2}. \end{aligned} \tag{3.4}$$

To estimate the first term, we shall use (2.8). Note that  $\sigma_s^\varepsilon$  is now smooth, hence the inequality (2.8) holds without the exceptional set  $N$ . Thus

$$\begin{aligned} I_{2,1} &\leq 3C_d^2 \mathbb{E} \int_0^{t \wedge \tau_R} [M_{2R} |\nabla \sigma_s^\varepsilon|(Y_s^{(1)}) + M_{2R} |\nabla \sigma_s^\varepsilon|(Y_s^{(2)})]^2 ds \\ &\leq 6C_d^2 \mathbb{E} \int_0^t \left( [M_{2R} |\nabla \sigma_s^\varepsilon|(Y_s^{(1)})]^2 \mathbf{1}_{\{|Y_s^{(1)}| \leq R\}} + [M_{2R} |\nabla \sigma_s^\varepsilon|(Y_s^{(2)})]^2 \mathbf{1}_{\{|Y_s^{(2)}| \leq R\}} \right) ds. \end{aligned}$$

Recall that  $Y_s^{(i)}$  has the same law with  $X_s^{(i)}$ , which is distributed as  $u_s^{(i)}(x) dx$ ,  $i = 1, 2$ . Consequently,

$$\begin{aligned} I_{2,1} &\leq C \int_0^t \int_{B(R)} (M_{2R} |\nabla \sigma_s^\varepsilon|(x))^2 (u_s^{(1)}(x) + u_s^{(2)}(x)) dx ds \\ &\leq C \sum_{i=1}^2 \|u^{(i)}\|_{L^\infty([0,T], L^\infty(\mathbb{R}^d))} \int_0^t \int_{B(R)} (M_{2R} |\nabla \sigma_s^\varepsilon|(x))^2 dx ds \\ &\leq \tilde{C} \int_0^t \int_{B(3R)} (|\nabla \sigma_s^\varepsilon|(x))^2 dx ds \\ &\leq \tilde{C} \|\nabla \sigma\|_{L^2([0,T], L^2(B(3R+1)))}^2 < +\infty, \end{aligned}$$

where in the third inequality we have used (2.7). Note that the bound is independent of  $\varepsilon \in (0, 1)$ . In the same way,

$$\begin{aligned} I_{2,2} &\leq \frac{3}{\delta^2} \sum_{i=1}^2 \mathbb{E} \int_0^t \|\sigma_s^\varepsilon(Y_s^{(i)}) - \sigma_s(Y_s^{(i)})\|^2 \mathbf{1}_{\{|Y_s^{(i)}| \leq R\}} ds \\ &\leq \frac{3}{\delta^2} \sum_{i=1}^2 \int_0^t \int_{B(R)} \|\sigma_s^\varepsilon(x) - \sigma_s(x)\|^2 u_s^{(i)}(x) dx ds \\ &\leq \frac{3}{\delta^2} \sum_{i=1}^2 \|u^{(i)}\|_{L^\infty([0,T], L^\infty(\mathbb{R}^d))} \int_0^t \int_{B(R)} \|\sigma_s^\varepsilon(x) - \sigma_s(x)\|^2 dx ds \end{aligned}$$

which vanishes as  $\varepsilon \rightarrow 0$  by the assumption (H2). Substituting the above two estimates into (3.4) gives us

$$I_2 \leq \tilde{C} \|\nabla \sigma\|_{L^2([0,T], L^2(B(3R+1)))}^2 =: \tilde{C}_{T,R} < +\infty. \tag{3.5}$$

*Step 2.* Now we turn to the difficult term  $I_1$  for which we shall need Bouchut and Crippa’s estimate in Theorem 2.15. Again we set  $b_s^\varepsilon = b_s * \chi_\varepsilon \in C_b^\infty(\mathbb{R}^d)$  for any  $\varepsilon \in (0, 1)$  and a.e.  $s \in [0, T]$ . Then similar to (3.4), we have

$$\begin{aligned} I_1 &\leq 2 \mathbb{E} \int_0^{t \wedge \tau_R} \frac{|b_s(Y_s^{(1)}) - b_s(Y_s^{(2)})|}{\sqrt{|Z_s|^2 + \delta^2}} ds \\ &\leq 2 \mathbb{E} \int_0^{t \wedge \tau_R} \frac{|b_s^\varepsilon(Y_s^{(1)}) - b_s^\varepsilon(Y_s^{(2)})|}{\sqrt{|Z_s|^2 + \delta^2}} ds \\ &\quad + 2 \mathbb{E} \int_0^{t \wedge \tau_R} \frac{|b_s^\varepsilon(Y_s^{(1)}) - b_s(Y_s^{(1)})| + |b_s^\varepsilon(Y_s^{(2)}) - b_s(Y_s^{(2)})|}{\sqrt{|Z_s|^2 + \delta^2}} ds \\ &=: I_{1,1} + I_{1,2}. \end{aligned} \tag{3.6}$$

The estimate of the term  $I_{1,2}$  is analogous to that of  $I_{2,2}$ :

$$\begin{aligned} I_{1,2} &\leq \frac{2}{\delta} \sum_{i=1}^2 \mathbb{E} \int_0^t |b_s^\varepsilon(Y_s^{(i)}) - b_s(Y_s^{(i)})| \mathbf{1}_{\{|Y_s^{(i)}| \leq R\}} ds \\ &\leq \frac{2}{\delta} \sum_{i=1}^2 \int_0^t \int_{B(R)} |b_s^\varepsilon(x) - b_s(x)| u_s^{(i)}(x) dx ds \\ &\leq \frac{2}{\delta} \sum_{i=1}^2 \|u^{(i)}\|_{L^\infty([0,T], L^\infty(\mathbb{R}^d))} \int_0^t \int_{B(R)} |b_s^\varepsilon(x) - b_s(x)| dx ds \\ &\rightarrow 0 \quad \text{as } \varepsilon \downarrow 0 \end{aligned} \tag{3.7}$$

since  $b \in L^\infty([0, T], L^\infty(\mathbb{R}^d))$ .

Finally we deal with the term  $I_{1,1}$ . Fix any  $\eta > 0$ . From (H3), we have

$$\partial_j b_s^\varepsilon = \sum_{k=1}^{m_0} S_{jk}(g_{jk}(s) * \chi_\varepsilon).$$

Moreover, for the finite family  $\{g_{jk}; 1 \leq j \leq d, 1 \leq k \leq m_0\} \subset L^1((0, T) \times \mathbb{R}^d, \mathbb{R}^d)$ , we can find  $C_\eta > 0$  and a set  $A_\eta \subset (0, T) \times \mathbb{R}^d$  with finite measure such that for every  $j = 1, \dots, d$  and  $k = 1, \dots, m_0$ , we have the decomposition below:

$$g_{jk}(s, x) = g_{jk}^{(1)}(s, x) + g_{jk}^{(2)}(s, x)$$

satisfying

$$\|g_{jk}^{(1)}\|_{L^1((0,T) \times \mathbb{R}^d, \mathbb{R}^d)} \leq \eta, \quad \text{supp}(g_{jk}^{(2)}) \subset A_\eta \quad \text{and} \quad \|g_{jk}^{(2)}\|_{L^2((0,T) \times \mathbb{R}^d, \mathbb{R}^d)} \leq C_\eta. \quad (3.8)$$

Now by Theorem 2.15 (see in particular the remark after it),

$$|b_s^\varepsilon(x) - b_s^\varepsilon(y)| \leq |x - y|(U_s^\varepsilon(x) + U_s^\varepsilon(y)), \quad \text{for all } x, y \in \mathbb{R}^d, \quad (3.9)$$

where

$$\begin{aligned} U_s^\varepsilon &= \sum_{k=1}^{m_0} \sum_{j=1}^d M_{\{\Lambda^{\varepsilon,j}, \xi \in S^{d-1}\}} [S_{jk}(g_{jk}(s) * \chi_\varepsilon)] \\ &\leq \sum_{k=1}^{m_0} \sum_{j=1}^d \left( M_{\{\Lambda^{\varepsilon,j}, \xi \in S^{d-1}\}} [S_{jk}(g_{jk}^{(1)}(s) * \chi_\varepsilon)] + M_{\{\Lambda^{\varepsilon,j}, \xi \in S^{d-1}\}} [S_{jk}(g_{jk}^{(2)}(s) * \chi_\varepsilon)] \right) \\ &=: U_s^{\varepsilon,1} + U_s^{\varepsilon,2}. \end{aligned}$$

Therefore

$$\begin{aligned} I_{1,1} &\leq 2 \mathbb{E} \int_0^{t \wedge \tau_R} \min \left\{ \frac{|b_s^\varepsilon(Y_s^{(1)})| + |b_s^\varepsilon(Y_s^{(2)})|}{\delta}, \frac{|b_s^\varepsilon(Y_s^{(1)}) - b_s^\varepsilon(Y_s^{(2)})|}{|Y_s^{(1)} - Y_s^{(2)}|} \right\} ds \\ &\leq 2 \mathbb{E} \int_0^{t \wedge \tau_R} \min \left\{ \frac{2\|b_s\|_{L^\infty(\mathbb{R}^d)}}{\delta}, U_s^\varepsilon(Y_s^{(1)}) + U_s^\varepsilon(Y_s^{(2)}) \right\} ds \\ &\leq 2 \mathbb{E} \int_0^{t \wedge \tau_R} \min \left\{ \frac{2\|b_s\|_{L^\infty(\mathbb{R}^d)}}{\delta}, U_s^{\varepsilon,1}(Y_s^{(1)}) + U_s^{\varepsilon,1}(Y_s^{(2)}) \right\} ds \\ &\quad + 2 \mathbb{E} \int_0^{t \wedge \tau_R} \min \left\{ \frac{2\|b_s\|_{L^\infty(\mathbb{R}^d)}}{\delta}, U_s^{\varepsilon,2}(Y_s^{(1)}) + U_s^{\varepsilon,2}(Y_s^{(2)}) \right\} ds \\ &=: I_{1,1,1} + I_{1,1,2}. \end{aligned} \quad (3.10)$$

Similar to the treatment of  $I_{2,1}$ , we have

$$\begin{aligned} I_{1,1,2} &\leq 2 \mathbb{E} \int_0^{t \wedge \tau_R} [U_s^{\varepsilon,2}(Y_s^{(1)}) + U_s^{\varepsilon,2}(Y_s^{(2)})] ds \\ &\leq 2 \sum_{i=1}^2 \int_0^t \int_{B(R)} U_s^{\varepsilon,2}(x) u_s^{(i)}(x) dx ds \\ &\leq 2 \sum_{i=1}^2 \|u^{(i)}\|_{L^\infty([0,T], L^\infty(\mathbb{R}^d))} \int_0^t \int_{B(R)} U_s^{\varepsilon,2}(x) dx ds. \end{aligned}$$

By Theorem 2.14(ii), we can find a positive constant  $L_1 > 0$  such that

$$\begin{aligned} \|U^{\varepsilon,2}\|_{L^2([0,T],L^2(\mathbb{R}^d))} &= \left[ \int_0^T \int_{\mathbb{R}^d} |U_s^{\varepsilon,2}(x)|^2 dx ds \right]^{\frac{1}{2}} \\ &\leq L_1 \sum_{k=1}^{m_0} \sum_{j=1}^d \left[ \int_0^T \int_{\mathbb{R}^d} |g_{jk}^{(2)}(s,x)|^2 dx ds \right]^{\frac{1}{2}} \\ &\leq L_1 dm_0 C_\eta, \end{aligned}$$

where the last inequality follows from (3.8). Thus by Cauchy’s inequality,

$$I_{1,1,2} \leq C \sqrt{T \mathcal{L}^d(B(R))} \|U^{\varepsilon,2}\|_{L^2([0,T],L^2(B(R)))} \leq C_{d,T,R} C_\eta. \tag{3.11}$$

It remains to estimate the quantity  $I_{1,1,1}$  defined in (3.10). We have

$$\begin{aligned} I_{1,1,1} &\leq 2 \sum_{i=1}^2 \mathbb{E} \int_0^{t \wedge \tau_R} \min \left\{ \frac{2 \|b_s\|_{L^\infty(\mathbb{R}^d)}}{\delta}; U_s^{\varepsilon,1}(Y_s^{(i)}) \right\} ds \\ &\leq 2 \sum_{i=1}^2 \int_0^t \int_{B(R)} \min \left\{ \frac{2 \|b_s\|_{L^\infty(\mathbb{R}^d)}}{\delta}; U_s^{\varepsilon,1}(x) \right\} u_s^{(i)}(x) dx ds \\ &\leq \hat{C} \int_0^t \int_{B(R)} \min \left\{ \frac{2 \|b_s\|_{L^\infty(\mathbb{R}^d)}}{\delta}; U_s^{\varepsilon,1}(x) \right\} dx ds. \end{aligned} \tag{3.12}$$

For simplicity of notations, we denote by  $\psi_s(x)$  the integrand on the right hand side. Using the simple inequality

$$\| \|U^{\varepsilon,1}\| \|_{M_{s,x}^1} \leq \| \| \|U^{\varepsilon,1}\| \|_{M_x^1} \|_{L_s^1},$$

we deduce from Theorem 2.14(i) that there exists a positive constant  $L_2 > 0$ , such that

$$\| \|U^{\varepsilon,1}\| \|_{M^1((0,T) \times \mathbb{R}^d)} \leq \int_0^T L_2 \sum_{j=1}^d \sum_{k=1}^{m_0} \| \|g_{jk}^{(1)}(s)\| \|_{L^1(\mathbb{R}^d)} ds \leq L_2 dm_0 T \eta,$$

where the last inequality is due to (3.8). Therefore, by the definition of  $\psi$ ,

$$\| \|\psi\| \|_{M^1((0,T) \times B(R))} \leq \| \|U^{\varepsilon,1}\| \|_{M^1((0,T) \times \mathbb{R}^d)} \leq L_2 dm_0 T \eta =: \hat{L}_2 \eta. \tag{3.13}$$

On the other hand,

$$\|\psi\|_{L^\infty((0,T) \times (B(R)))} \leq \frac{2 \|b\|_{L^\infty([0,T] \times \mathbb{R}^d)}}{\delta}.$$

Combining this estimate with (3.13) and applying (2.4), we get

$$\|\psi\|_{L^1((0,T) \times B(R))} \leq \hat{L}_2 \eta \left[ 1 + \log \left( \frac{2 \|b\|_{L^\infty}}{\delta} \cdot \frac{T \mathcal{L}^d(B(R))}{\hat{L}_2 \eta} \right) \right],$$

in which we have used the fact that the function  $s \mapsto s(1 + \log^+(C/s))$  is nondecreasing on  $[0, \infty)$ . Substituting this inequality into (3.12) finally leads to

$$I_{1,1,1} \leq \hat{C} \hat{L}_2 \eta \left[ 1 + \log \left( \frac{2 \|b\|_{L^\infty}}{\delta} \cdot \frac{T \mathcal{L}^d(B(R))}{\hat{L}_2 \eta} \right) \right].$$

Combining the above estimate with (3.10) and (3.11), we obtain

$$I_{1,1} \leq C_{d,T,R} C_\eta + \hat{C} \hat{L}_2 \eta \left[ 1 + \log \left( \frac{2 \|b\|_{L^\infty}}{\delta} \cdot \frac{T \mathcal{L}^d(B(R))}{\hat{L}_2 \eta} \right) \right],$$

which, together with (3.6) and (3.7), yields

$$I_1 \leq C_{d,T,R}C_\eta + \hat{C}\hat{L}_2\eta \left[ 1 + \log \left( \frac{2\|b\|_{L^\infty}}{\delta} \cdot \frac{T\mathcal{L}^d(B(R))}{\hat{L}_2\eta} \right) \right]. \quad (3.14)$$

*Step 3.* Having the estimates (3.5) and (3.14) in hand, we are ready to complete the proof as follows. Substituting the two estimates (3.5) and (3.14) into (3.3), we get for any  $t \in [0, T]$  that

$$\mathbb{E} \log \left( \frac{|Z_{t \wedge \tau_R}|^2}{\delta^2} + 1 \right) \leq \tilde{C}_{T,R} + C_{d,T,R}C_\eta + \hat{C}\hat{L}_2\eta \left[ 1 + \log \left( \frac{2\|b\|_{L^\infty}}{\delta} \cdot \frac{T\mathcal{L}^d(B(R))}{\hat{L}_2\eta} \right) \right].$$

Fix any  $\theta > 0$ . The above inequality implies

$$\begin{aligned} P(|Z_{t \wedge \tau_R}| > \theta) &\leq \frac{\tilde{C}_{T,R} + C_{d,T,R}C_\eta + \hat{C}\hat{L}_2\eta}{\log \left[ \left(\frac{\theta}{\delta}\right)^2 + 1 \right]} + \frac{\hat{C}\hat{L}_2\eta}{\log \left[ \left(\frac{\theta}{\delta}\right)^2 + 1 \right]} \log \left( \frac{2\|b\|_{L^\infty}}{\delta} \right) \\ &\quad + \frac{\hat{C}\hat{L}_2\eta}{\log \left[ \left(\frac{\theta}{\delta}\right)^2 + 1 \right]} \log \left( \frac{T\mathcal{L}^d(B(R))}{\hat{L}_2\eta} \right). \end{aligned}$$

Notice that in the second term, the quantity

$$\frac{1}{\log \left[ \left(\frac{\theta}{\delta}\right)^2 + 1 \right]} \log \left( \frac{2\|b\|_{L^\infty}}{\delta} \right)$$

is bounded as  $\delta$  tends to 0. Therefore first letting  $\delta \downarrow 0$  and then  $\eta \downarrow 0$  we arrive at  $P(|Z_{t \wedge \tau_R}| > \theta) = 0$ . Since  $\theta$  can be arbitrarily small, it follows that  $Z_{t \wedge \tau_R} = 0$  almost surely. Finally, we conclude from (3.2) that for any  $t \in [0, T]$ ,  $Z_t = Y_t^{(1)} - Y_t^{(2)} = 0$  a.s. The continuity of the two processes  $Y_t^{(1)}$  and  $Y_t^{(2)}$  yields that, almost surely,  $Y_t^{(1)} = Y_t^{(2)}$  for all  $t \in [0, T]$ . Therefore  $P_{\mu_0}^{(1)} = P_{\mu_0}^{(2)}$ , which, together with the representation formula (3.1), leads to the uniqueness of solutions to (1.3).  $\square$

## References

- [1] Ambrosio, L.: Transport equation and Cauchy problem for BV vector fields. *Invent. Math.* **158**, (2004), 227–260. MR-2096794
- [2] Ambrosio, L.: Transport equation and Cauchy problem for non-Smooth Vector Fields. *Calculus of variations and nonlinear partial differential equations*, 1–41, Lecture Notes in Math., 1927, Springer, Berlin, 2008. MR-2408257
- [3] Ambrosio, L. and Crippa, G.: Existence, uniqueness, stability and differentiability properties of the flow associated to weakly differentiable vector fields. *Transport equations and multi-D hyperbolic conservation laws*, 3–57, Lect. Notes Unione Mat. Ital., 5, Springer, Berlin, 2008. MR-2409676
- [4] Ambrosio, L. and Figalli, A.: On flows associated to Sobolev vector fields in Wiener space: an approach à la DiPerna–Lions. *J. Funct. Anal.* **256**, (2009), 179–214. MR-2475421
- [5] Ambrosio, L. and Trevisan, D.: Well posedness of Lagrangian flows and continuity equations in metric measure spaces, arXiv:1402.4788
- [6] Bogachev, V. I., Da Prato, G., Röckner, M. and Stannat, W.: Uniqueness of solutions to weak parabolic equations for measures. *Bull. Lond. Math. Soc.* **39**, (2007), no. 4, 631–640. MR-2346944
- [7] Bogachev, V. I., Da Prato, G. and Röckner, M.: Existence and uniqueness of solutions for Fokker–Planck equations on Hilbert spaces. *J. Evol. Equ.* **10**, (2010), no. 3, 487–509. MR-2674056

- [8] Bogachev, V. I., Röckner, M. and Shaposhnikov, S. V.: On uniqueness of solutions to the Cauchy problem for degenerate Fokker–Planck–Kolmogorov equations. *J. Evol. Equ.* **13**, (2013), no. 3, 577–593. MR-3089794
- [9] Bouchut, F. and Crippa, G.: Équations de transport à coefficient dont le gradient est donné par une intégrale singulière. Séminaire: Équations aux Dérivées Partielles, Exp. No.1 2007–2008, 1–13, 2009. MR-2532938
- [10] Bouchut, F. and Crippa, G.: Lagrangian flows for vector fields with gradient given by a singular integral. *J. Hyperbolic Differ. Equ.* **10**, (2013), 235–282. MR-3078074
- [11] Champagnat, N. and Jabin, P. E.: Strong solutions to stochastic differential equations with rough coefficients, arXiv:1303.2611
- [12] Crippa, G. and De Lellis, C.: Estimates and regularity results for the DiPerna–Lions flows. *J. Reine Angew. Math.* **616**, (2008), 15–46. MR-2369485
- [13] DiPerna, R. J. and Lions, P. L.: Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.* **98**, (1989), 511–547. MR-1022305
- [14] Fang, S. and Luo, D.: Transport equations and quasi-invariant flows on the Wiener space. *Bull. Sci. Math.* **134**, (2010), 295–328. MR-2607935
- [15] Fang, S., Luo, D. and Thalmaier, A.: Stochastic differential equations with coefficients in Sobolev spaces. *J. Funct. Anal.* **259**, (2010), 1129–1168. MR-2652184
- [16] Figalli, A.: Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *J. Funct. Anal.* **254**, (2008), 109–153. MR-2375067
- [17] Ikeda, N. and Watanabe, S.: Stochastic differential equations and diffusion processes, 2nd ed., North-Holland, 1989. MR-1011252
- [18] Le Bris, C. and Lions, P. L.: Existence and uniqueness of solutions to Fokker–Planck type equations with irregular coefficients. *Comm. Partial Differential Equations* **33**, (2008), 1272–1317. MR-2450159
- [19] Luo, D.: Well-posedness of Fokker–Planck type equations on the Wiener space. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **13**, (2010), no. 2, 273–304. MR-2669050
- [20] Luo, D.: Fokker–Planck type equations with Sobolev diffusion coefficients and BV drift coefficients. *Acta Math. Sin. (Engl. Ser.)* **29** (2013), no. 2, 303–314. MR-3016531
- [21] Röckner, M. and Zhang, X.: Weak uniqueness of Fokker–Planck equations with degenerate and bounded coefficients. *C. R. Math. Acad. Sci. Paris* **348** (2010), 435–438. MR-2607035
- [22] Stein, E. M.: Singular integrals and differentiability properties of functions. *Princeton University Press*, 1970. MR-0290095
- [23] Zhang, X.: Stochastic flows of SDEs with irregular coefficients and stochastic transport equations. *Bull. Sci. Math.* **134**, (2010), 340–378. MR-2651896

**Acknowledgments.** The author is indebted to the anonymous referee for his/her valuable suggestions which not only improve the presentation of the current paper but also inspire future researches.