From nonlinear Fokker-Planck equations to solutions of distribution dependent SDE

Viorel Barbu* Michael Röckner†‡

Abstract

We construct weak solutions to the McKean-Vlasov SDE

$$dX(t) = b\left(X(t), \frac{d\mathcal{L}_{X(t)}}{dx}(X(t))\right)dt + \sigma\left(X(t), \frac{d\mathcal{L}_{X(t)}}{dt}(X(t))\right)dW(t)$$

on $\mathbb{R}^d$ for possibly degenerate diffusion matrices σ with $X(0)$ having a given law, which has a density with respect to Lebesgue measure, dx. Here $\mathcal{L}_{X(t)}$ denotes the law of $X(t)$. Our approach is to first solve the corresponding nonlinear Fokker-Planck equations and then use the well known superposition principle to obtain weak solutions of the above SDE.

Keywords: Fokker-Planck equation, Kolmogorov operator, probability density, m-accretive operator, Wiener process.
Mathematics Subject Classification (2000): 60H30, 60H10, 60G46, 35C99, 58J165.

1 Introduction

Recently there has been an increasing interest in distribution dependent stochastic differential equations (DDSDE for short) of type

$$\begin{aligned} &dX(t) = b(t, X(t), \mathcal{L}_{X(t)})dt + \sigma(t, X(t), \mathcal{L}_{X(t)})dW(t) \\ &X(0) = \xi_0, \end{aligned} \tag{1.1}$$

*Octav Mayer Institute of Mathematics of the Romanian Academy, Iaşi, Romania
†Fakultät für Mathematik, Universität Bielefeld, D-33501 Bielefeld, Germany
‡Academy of Mathematics and Systems Science, CAS, Beijing

on $\mathbb{R}^d$, where $W(t),\ t \geq 0$, is an $(\mathcal{F}_t)$-Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ with normal filtration $(\mathcal{F}_t)_{t\geq 0}$. The coefficients b, σ are defined on $[0,\infty) \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$ are $\mathbb{R}^d$ and $d \times d$-matrix valued, respectively (satisfying conditions to be specified below). Here $\mathcal{P}(\mathbb{R}^d)$ denotes the set of all probability measures on $\mathbb{R}^d$. In (1.1), $\mathcal{L}_{X(t)}$ denotes the law of $X(t)$ under P and ξ_0 is an $\mathcal{F}_0$-measurable $\mathbb{R}^d$-valued map. Equations as in (1.1) are also referred to as McKean-Vlasov SDEs. Here we refer to the classical papers [16], [21], [22], [26], [28], and, e.g., the more recent papers [13], [17], [18], [19], [23], [24] and [30].

By Itô's formula, under quite general conditions on the coefficients, the time marginal laws $\mu_t := \mathcal{L}_{X(t)},\ t \geq 0$, with $\mu_0 :=$ law of ξ_0, of the solution $X(t),\ t \geq 0$, to (1.1) satisfy a *nonlinear* Fokker-Planck equation (FPE for short). More precisely, for all $\varphi \in C_0^2(\mathbb{R}^d)$ (= all twice differentiable real-valued functions of compact support) and, for all $t \geq 0$,

$$\int_{\mathbb{R}^d} \varphi(x)\mu_t(dx) = \int_{\mathbb{R}^d} \varphi(x)\mu_0(dx) + \int_0^t \int_{\mathbb{R}^d} L_{\mu_s}\varphi(s,x)\mu_s(dx)ds, \tag{1.2}$$

where, for $x \in \mathbb{R}^d,\ t \geq 0,\ a_{ij} := (\sigma\sigma^T)_{i,j},\ 1 \leq i,j \leq d$,

$$L_{\mu_t}\varphi(t,x) := \frac{1}{2}\sum_{i,j=1}^d a_{ij}(t,x,\mu_t)\,\frac{\partial^2}{\partial x_i \partial x_j}\,\varphi(x) + \sum_{i=1}^d b_i(t,x,\mu_t)\,\frac{\partial}{\partial x_i}\,\varphi(x), \tag{1.3}$$

is the corresponding Kolmogorov operator. For equations of type (1.2), we refer the reader, e.g., to [10]. We note that (1.2) is also shortly written as

$$\partial_t \mu_t = L^*_{\mu_t}\mu_t \text{ with } \mu_0 \text{ given.} \tag{1.4}$$

Hence, if one can solve (1.1), one obtains a solution to (1.2) this way.

In the special case where the solutions $\mu_t,\ t \geq 0$, to (1.2) have densities with respect to the Lebesgue measure dx, i.e., $\mu_t(dx) = u(t,x)dx,\ t \geq 0$, (1.2) can be rewritten (in the sense of Schwartz distributions) as (cf.[15])

$$\begin{aligned} \frac{\partial}{\partial t}\,u(t,x) &= \frac{1}{2}\sum_{i,j=1}^d \frac{\partial^2}{\partial x_i \partial x_j}\,[a_{ij}(t,x,u(t,\cdot)dx)u(t,x)] \\ &\quad - \sum_{i=1}^d \frac{\partial}{\partial x_i}[b_i(t,x,u(t,\cdot)dx)u(t,x)] \\ u(0,x) &= u_0(x)\ \left(= \frac{d\mu_0}{dx}\,(x)\right), \end{aligned} \tag{1.5}$$

$x \in \mathbb{R}^d,\ t \geq 0$, or, shortly,

$$\partial_t u = \frac{1}{2} \partial_i \partial_j (a_{ij}(u)u) - \partial_i (b_i(u)u),\ u(0, \cdot) = u_0. \tag{1.6}$$

In this paper, we want to go in the opposite direction, that is, we first want to solve (1.2) and, using the obtained μ_t, $t \geq 0$, we shall obtain a (probabilistically) weak solution to (1.1) with the time marginal laws of $X(t), t \geq 0$, given by these $\mu_t,\ t \geq 0$. It turns out that, once one has solved (1.2), which is in general a hard task, and if one can prove some mild integrability properties for the solutions, a recent version of the so-called *superposition principle* by Trevisan in [29] (generalizing earlier work by Figalli [14]), in connection with a classical result by Stroock and Varadhan (see, e.g., [27]) yields the desired weak solution of (1.1) (see Section 2 below for details).

We would like to mention at this point that, by the very same result from [29], one can also easily prove that, if (1.1) has a unique solution in law, then the solution to (1.2) does not only exist as described above, but is also unique. In this paper, however, we concentrate on existence of weak solutions to (1.1). We shall do this in the singular case, where the coefficients in (1.1) are of *Nemytskii-type*, that is, we consider the following situation: b_i, a_{ij} depend on μ in the following way:

$$b_i(t, x, \mu) := \bar{b}_i \left(t, x, \frac{d\mu}{dx}(x) \right), \quad a_{ij}(t, x, \mu) := \bar{a}_{ij} \left(t, x, \frac{d\mu}{dx}(x) \right), \tag{1.7}$$

for $t \geq 0,\ x \in \mathbb{R}^d,\ 1 \leq i, j \leq d$, where $\bar{b}_i, \bar{a}_{ij} : [0, \infty) \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$, are measurable functions. Then, under the conditions on $\bar{b}_i$ and $\bar{a}_{ij}$, $1 \leq i, j \leq d$, specified in Section 3, we shall construct solutions $(\mu_t)_{t \geq 0}$ to (1.1) which are absolutely continuous with respect to the Lebesgue dx, i.e., $\mu_t(dx) = u(t, x)dx$, $t \geq 0$. So, as indicated above, by the superposition principle, we obtain weak solutions to DDSDEs of type

$$\begin{aligned} dX(t) &= \bar{b} \left(t, X(t), \frac{d\mathcal{L}_{X(t)}}{dx}(X(t)) \right) dt \\ &\quad + \bar{\sigma} \left(t, X(t), \frac{d\mathcal{L}_{X(t)}}{dx}(X(t)) \right) dW(t), \\ X(0) &= \xi_0, \end{aligned} \tag{1.8}$$

with $(\bar{\sigma}\bar{\sigma}^T)_{i,j} = \bar{a}_{ij}$.

In particular, we obtain a probabilistic representation of the solution μ_t, $t \geq 0$, of the nonlinear FPE (1.2) (or (1.5)) as the time marginal laws of a stochastic process, namely the solution of the DDSDE (1.8).

We would like to emphasize that the coefficients as in (1.8), which we consider below, have no continuity properties with respect to their dependence on the law $\mathcal{L}_{X(t)}$ of $X(t)$, such as those imposed in the existing literature on the subject. Nevertheless, such *Nemytskii-type*-dependence is very natural and, of course, independent of the dx-version of the Lebesgue density of $\mathcal{L}_{X(t)}$ we choose in (1.8), since we are looking only for solutions of (1.8) in the class with $\mathcal{L}_{X(t)}$ being absolutely continuous with respect to dx. Precise conditions on the coefficients $\bar{b}_i, \bar{a}_{ij}$ are formulated in Section 3 (there, for simplicity, denoted by b_i, a_{ij}). Our main existence results for solutions of the nonlinear FPE (1.2) are Theorems 3.4 and 3.7 below. Our main result on solutions to (1.1) (more precisely, (1.8)) is Theorem 4.1. Subsequently, in Remark 4.2 we discuss connections with previous related, but much more special, results from [4]–[9]. A class of cases where we also have uniqueness in law results for solutions to (1.8) is described in Remark 4.3.

Notations. Given an open subset $\mathcal{O} \subset \mathbb{R}^d$, we denote by $L^p(\mathcal{O})$, $1 \leq p \leq \infty$, the standard Lebesgue p-integrable functions on $\mathcal{O}$, and by $H^1(\mathcal{O})$, the Sobolev space $\{u \in L^2(\mathcal{O}); \nabla u \in L^2(\mathcal{O})\}$.

We set $H^1_0(\mathcal{O}) = \{u \in H^1(\mathcal{O});\ u = 0 \text{ on } \partial\mathcal{O}\}$ and denote by $H^{-1}(\mathcal{O})$ the dual space of $H^1_0(\mathcal{O})$. By $C^\infty_0(\mathcal{O})$ we denote the space of infinitely differentiable functions with compact support in $\mathcal{O}$. We set $H^1 = H^1(\mathbb{R}^d)$, $H^{-1} = H^{-1}(\mathbb{R}^d)$ and denote by $H^1_{\rm loc}$ the corresponding local space.

We also set $L^p = L^p(\mathbb{R}^d)$ with the norm denoted $|\cdot|_p$ and $L^p_{\rm loc} = L^p_{\rm loc}(\mathbb{R}^d)$, $1 \leq p \leq \infty$. By $\mathcal{D}'(\mathbb{R}^d)$ and $\mathcal{D}'((0,\infty) \times \mathbb{R}^d)$, we denote the space of distributions on $\mathbb{R}^d$ and $(0,\infty) \times \mathbb{R}^d$, respectively.

We shall denote either by $\frac{\partial u}{\partial x_j}$ or by u_{x_j}, D_j, u the partial derivative of the function $u = u(x_1, ..., x_d)$ with respect to x_j, $1 \leq j \leq d$. By D^2_{ij}, $i, j = 1, ..., d$, we shall denote the second order derivatives $\frac{\partial^2 u}{\partial x_i \partial x_j}$.

We denote by $C(\mathbb{R}^d \times \mathbb{R})$ and $C(\mathbb{R}^d)$ the space of continuous functions on $\mathbb{R}^d \times \mathbb{R}$ and $\mathbb{R}$, respectively, and by $C_b(\mathbb{R}^d \times \mathbb{R})$ and $C_b(\mathbb{R})$ the corresponding subspaces of continuous and bounded functions.

By $C^1(\mathbb{R}^d \times \mathbb{R})$ and $C^1(\mathbb{R})$ we denote the spaces of continuously differentiable functions on $\mathbb{R}^d \times \mathbb{R}$ and $\mathbb{R}$, respectively. Finally, C^1_b is the space of bounded continuously differentiable functions with bounded derivatives.

If $\mathcal{X}$ is a real Banach space and $0 < T < \infty$, we denote by $L^p(0, T; \mathcal{X})$ the space of Bochner p-integrable functions $u:(0,T)\to\mathcal{X}$ and by $C([0,T];\mathcal{X})$ the space of of $\mathcal{X}$-valued continuous functions on $[0,T]$.

2 From nonlinear FPEs to DDSDEs: general scheme

Let $a_{ij}, b_i : [0,\infty)\times\mathbb{R}^d\times\mathcal{P}(\mathbb{R}^d)\to\mathbb{R}$, $1\le i,j\le d$, be measurable.

Hypothesis 2.1 *There exists a solution $(\mu_t)_{t\ge 0}$ to (1.2) such that*

(i) $\mu_t\in\mathcal{P}(\mathbb{R}^d)$ *for all* $t\ge 0$.

(ii) $(t,x)\mapsto a_{ij}(t,x,\mu_t)$ *and* $(t,x)\mapsto b_i(t,x,\mu_t)$ *are measurable and*

$$\int_0^T\int_{\mathbb{R}^d}[|a_{ij}(t,x,\mu_t)|+|b_i(t,x,\mu_t)|]\mu_t(dx)dt<\infty,$$

for all $T\in(0,\infty)$.

(iii) $[0,\infty)\ni t\mapsto\mu_t$ *is weakly continuous.*

Under Hypothesis 2.1, we can apply the superposition principle (see Theorem 2.5 in [29]) for *linear* FPEs applied to the (*linear*) Kolmogorov operator

$$L_{\mu_t}:=\frac{1}{2}\sum_{i,j=1}^d a_{ij}(t,x,\mu_t)\frac{\partial^2}{\partial x_i\partial x_j}+\sum_{i=1}^d b_i(t,x,\mu_t)\frac{\partial}{\partial x_i}, \tag{2.1}$$

with $(\mu_t)_{t\ge 0}$ from Hypothesis 2.1 fixed.

More precisely, by Theorem 2.5 in [29], there exists a probability measure P on $C([0,T];\mathbb{R}^d)$ equipped with its Borel σ-algebra and its natural normal filtration obtained by the evaluation maps π_t, $t\in[0,T]$, defined by

$$\pi_t(w):=w(t),\ \ w\in C([0,T],\mathbb{R}^d),$$

solving the martingale problem (see [29], Definition 2.4) for the time-dependent (*linear*) Kolmogorov operator $\frac{\partial}{\partial t}+L_{\mu_t}$ (with $(\mu_t)_{t\ge 0}$ as above fixed) with time marginal laws

$$P\circ\pi_t^{-1}=\mu_t,\ \ t\ge 0.$$

Then, a standarad result (see Theorem 4.5.2 in [27]) implies that there exists a d-dimensional $(\mathcal{F}_t)$-Brownian motion $W(t)$, $t \geq 0$, on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, Q)$ and a continuous $(\mathcal{F}_t)$-progressively measurable map $X : [0,\infty) \times \Omega \to \mathbb{R}^d$ satisfying the following (DD)SDE

$$dX(t) = b(t, X(t), \mu_t)dt + \sigma(t, X(t), \mu_t)dW(t), \tag{2.2}$$

with the law

$$Q \circ X^{-1} = P,$$

where $\sigma = ((a_{ij})_{1\leq i,j\leq d})^{\frac{1}{2}}$. In particular, we have, for the marginal laws,

$$\mathcal{L}_{X(t)} := Q \circ X(t)^{-1} = \mu_t,\ t \geq 0. \tag{2.3}$$

Remark 2.2 Because of (2.3), the process $X(t)$, $t \geq 0$, is also called a *probabilistic representation* of the solution $(\mu_t)_{t\geq 0}$ for the nonlinear FPE (1.2).

Remark 2.3 It is much harder to prove that the solution to SDE (2.2) for fixed $(\mu_t)_{t\geq 0}$ is unique in law, provided its initial distribution is μ_0, which would, of course, be very desirable. For this, one has to prove the uniqueness of the solutions to the *linear* Fokker-Planck equation

$$\partial_t \nu_t = L^*_{\mu_t} \nu_t,\ \nu_0 = \mu_0$$

for all initial condition of the type $\mu_0 = \delta_x$, $x \in \mathbb{R}^d$ (see [27], Theorem 6.2.3). For a large class of initial conditions μ_0, this was achieved in certain cases where $d = 1$ (see [5], [9], [25]).

Conclusion. To weakly solve DDSDE (1.1), we have to solve the correspon-ding nonlinear FPE (1.2) (hard!) and then check Hypothesis 2.1 above.

3 Existence of solutions to the nonlinear FPEs

Consider the following time-independent special case of (1.5) with Nemytskii-type dependence of the coefficients on $u(t,x)dx$, $t \geq 0$, i.e., the nonlinear Fokker-Planck equation

$$\begin{aligned} &\frac{\partial u}{\partial t} - \sum_{i,j=1}^{d} D^2_{ij}(a_{ij}(x,u)u) + \operatorname{div}(b(x,u)u) = 0 \text{ in } \mathcal{D}'((0,\infty)\times\mathbb{R}^d), \\ &u(0,x) = u_0(x),\ x \in \mathbb{R}^d, \end{aligned} \tag{3.1}$$

where $b(x,u)=\{b_i(x,u)\}_{i=1}^d$.

We shall study this equation under two different sets of hypotheses specified in the following.

(H1) $a_{ij}\in C^2(\mathbb{R}^d\times\mathbb{R})\cap C_b(\mathbb{R}^d\times\mathbb{R}^d),\ (a_{ij})_x\in C_b(\mathbb{R}^d\times\mathbb{R};\mathbb{R}^d),\ \forall a_{ij}=a_{ji}$, $i,j=1,...,d$.

(H2) $\sum\limits_{i,j=1}^{d}(a_{ij}(x,u)+(a_{ij}(x,u))_u u)\xi_i\xi_j\ \ge\ \gamma|\xi|^2,\ \forall\xi\in\mathbb{R}^d,\ x\in\mathbb{R}^d,\ u\in\mathbb{R}$, where $\gamma>0$.

(H3) $b_i\in C_b(\mathbb{R}^d\times\mathbb{R})\cap C^1(\mathbb{R}^d\times\mathbb{R})$, $b_i(x,0)\equiv 0$, $\forall x\in\mathbb{R}^d,\ i=1,...,d$.

(H1)$'$ $a_{ij}(x,u)\equiv a_{ij}(u),\ a_{ij}\in C^2(\mathbb{R})\cap C_b(\mathbb{R})$, $a_{ij}=a_{ji}$, $\forall i,j=1,...,d$.

(H2)$'$ $\sum\limits_{i,j=1}^{d}(a_{ij}(u)+u(a_{ij}(u))_u)\xi_i\xi_j\ge 0,\ \forall\xi\in\mathbb{R}^d,\ u\in\mathbb{R}$.

(H3)$'$ $b_i\in C_b(\mathbb{R})\cap C^1(\mathbb{R}^d),\ b_i(0)=0,\ i=1,...,d$.

Here $(a_{ij}(x,u))_u=\frac{\partial}{\partial u}a_{ij}(x,u),\ \forall u\in\mathbb{R}$, and $(a_{ij})_x(x,u)=(\nabla_x a_{ij})(x,u)$, $x=\{x_i\}_{i=1}^d$. The first set of hypotheses, that is (H1)–(H3), allows for nonlinear nondegenerate FPEs with x-dependent coefficients, while the second set (H1)$'$–(H3)$'$ allows for degenerate nonlinear FPEs, however, with x-independent coefficients.

Nonlinear FPEs of the form (3.1) describe in the mean field theory the dynamics of a set of interacting particles or many body systems. The function $u=u(t,x)$ is associated with the probability to find a certain subsystem or particle at time t in the state x. Equation (3.1) arises also as a closed loop system corresponding to a velocity field system

$$\frac{\partial v}{\partial t}=F(x,u)v=\sum_{i,j=1}^{d}D_{ij}^2(a_{ij}(x,u)v)-\mathrm{div}(b(x,u)v)$$

with coefficients depending on the probability density u. If $v=u$, one may view this system as a statistical feedback (see [15]).

The first part of this section is concerned with the existence of a weak (mild) solution to equation (3.1) in the space $L^1(\mathbb{R}^d)$. This result is obtained via the Crandall and Liggett existence theorem for the nonlinear Cauchy problem

$$
\begin{aligned}
&\frac{du}{dt}(t) + Au(t) = 0,\ t \geq 0, \\
&u(0) = u_0,
\end{aligned}
\tag{3.2}
$$

in a Banach space $\mathcal{X}$.

An operator $A : D(A) \subset \mathcal{X} \to \mathcal{X}$ (possibly multivalued) is said to be m-accretive if, for each $\lambda > 0$, the range $R(I + \lambda A)$ of the operator $I + \lambda A$ is all of $\mathcal{X}$ and

$$
\|(I + \lambda A)^{-1}u - (I + \lambda A)^{-1}v\|_{\mathcal{X}} \leq \|u - v\|_{\mathcal{X}},\ \forall u, v \in \mathcal{X},\ \lambda > 0. \tag{3.3}
$$

The continuous function $u : [0, \infty) \to \mathcal{X}$ is said to be a mild solution to (3.2) if, for each $0 < T < \infty$,

$$
u(t) = \lim_{h \to 0} u_h(t) \text{ strongly in } \mathcal{X}, \text{ uniformly in } t \in [0, T] \tag{3.4}
$$

where $u_h : [0, T] \to \mathcal{X}$ is defined by

$$
u_h(t) = u_h^i,\ t \in [ih, (i+1)h),\ i = 0, 1, ..., N = \left[\frac{T}{h}\right]. \tag{3.5}
$$

$$
u_h^i + hAu_h^i = u_h^{i-1},\ i = 1, ..., N;\ u_h^0 = u_0. \tag{3.6}
$$

By the Crandall and Liggett theorem (see, e.g., [1], p. 99), if A is m-accretive, then for each $u_0 \in \overline{D(A)}$ (the closure of $D(A)$ in $\mathcal{X}$) there is a unique mild solution $u \in C([0, \infty); \mathcal{X})$ to (3.2). Moreover, the map $u_0 \to u(t)$ is a continuous semigroup of contractions on $\overline{D(A)}$ equipped with $\| \cdot \|_{\mathcal{X}}$.

The first main existence result of this section, Theorem 3.4, is obtained by writing equation (3.1) in the form (3.2) with a suitable m-accretive operator A in the space $\mathcal{X} = L^1(\mathbb{R}^d)$.

It should be said that the space $L^1(\mathbb{R}^d)$ is not only appropriate to represent equation (3.1) in the form (3.2), but it is the unique $L^p(\mathbb{R}^d)$-space in which the operator defined by equation (3.1) is m-accretive, that is, which gives the parabolic character of this equation. Only in the particular case of porous media equations (i.e., (3.1) with $b \equiv 0$), an alternative is the Sobolev space $H^{-1}(\mathbb{R}^d)$, but this does not work for the more general case (3.1). On the other hand, taking into account the significance of the solution u as probability density, the space L^1 is very convenient for the treatment of equation (1.1).

Our work [4] contains the following special case of (3.1):

$$\frac{\partial u}{\partial t} - \Delta\beta(u) + \operatorname{div}(b(u)u) = 0 \text{ in } (0,T)\times\mathbb{R}^d, \tag{3.7}$$

where $\beta : \mathbb{R} \to 2^{\mathbb{R}}$ is a maximal monotone (multivalued) function with $\sup\{|s| : s \in \beta(r)\} \le C|r|^m,\ r \in \mathbb{R}$, for some $C, m \in [0,\infty)$. (See also [2].) In the special case $b \equiv 0$ and $d = 1$, related results were obtained in [5], [9]. However, the present case is much more difficult and the arguments of [4] are not applicable here.

3.1 Existence for FPEs in the nondegenerate, x-dependent case

Define in the space $\mathcal{X} = L^1$ the operator $A : D(A) \subset L^1 \to L^1$,

$$Au = -\sum_{i,j=1}^{d} D_{ij}^2(a_{ij}(x,u)u) + \operatorname{div}(b(x,u)u),\ \forall u \in D(A), \tag{3.8}$$

$$D(A) = \left\{u \in L^1; -\sum_{i,j=1}^{d} D_{ij}^2(a_{ij}(x,u)u) + \operatorname{div}(b(x,u)u) \in L^1\right\}, \tag{3.9}$$

where D_{ij}^2 and div are taken in the sense of Schwartz distributions on $\mathbb{R}^d$, i.e., in $\mathcal{D}'(\mathbb{R}^d)$. We note that, since by (H1), (H3), $a_{ij}(x,u)u,\ b_i(x,u)u \in L^1$, $\forall i,j = 1,...,d$, $\forall u \in L^1$, Au is well defined in $\mathcal{D}'(\mathbb{R}^d)$. Moreover, since $C_0^\infty(\mathbb{R}^d) \subset D(A)$, it follows that $D(A)$ is dense in L^1.

Since we are going to represent equation (3.1) as (3.2) with A defined by (3.8)–(3.9), we must prove that A is m-accretive, that is, $R(I+\lambda A) = L^1$ and (3.3) holds in $\mathcal{X} = L^1$ for all $\lambda > 0$. For this purpose, we shall prove the following result.

Proposition 3.1 *Let* (H1)–(H3) *hold. Then, for each $f \in L^1$ and $\lambda > 0$, the equation*

$$u - \lambda\sum_{i,j=1}^{d} D_{ij}^2(a_{ij}(x,u)u) + \lambda \operatorname{div}(b(x,u)u) = f \text{ in } \mathcal{D}'(\mathbb{R}^d) \tag{3.10}$$

has a unique solution $u = u(\lambda, f) \in D(A)$.

Moreover, we have, for all $\lambda > 0$,

$$(3.11) \qquad |u(\lambda, f_1) - u(\lambda, f_2)|_1 \le |f_1 - f_2|_1, \ \forall f_1, f_2 \in L^1,$$

$$(3.12) \qquad (I + \lambda A)^{-1} f \ge 0, \ \textit{a.e. in } \mathbb{R}^d, \ \textit{if } f \in L^1, \ f \ge 0, \quad \textit{a.e. in } \mathbb{R}^d,$$

$$(3.13) \qquad \int_{\mathbb{R}^d} (I + \lambda A)^{-1} f(x) dx = \int_{\mathbb{R}^d} f(x) dx, \ \forall f \in L^1, \ \lambda > 0.$$

Proof. In the following, we shall simply write

$$a_{ij}(u) = a_{ij}(x, u), \ x \in \mathbb{R}^d, \ u \in \mathbb{R}, \ i, j = 1, ..., d.$$

We set

$$a^*_{ij}(u) \equiv a_{ij}(x, u) u, \ x \in \mathbb{R}^d, \ u \in \mathbb{R}, \ \forall i, j = 1, ..., d,$$
$$b(x, u) = \{b_i(x, u)\}_{i=1}^d, \ b^*(x, u) = b(x, u) u, \ x \in \mathbb{R}^d, \ u \in \mathbb{R}.$$

We note that, by (H2), we have

$$(3.14) \qquad \sum_{i,j=1}^d (a^*_{ij})_u(x, u) \xi_i \xi_j \ge \gamma |\xi|^2, \ \forall \xi \in \mathbb{R}^d, \ x \in \mathbb{R}^d, \ u \in \mathbb{R},$$

where $\gamma > 0$. We shall first prove Proposition 3.1 under the additional hypotheses

(K) $(a^*_{ij})_u \in C_b(\mathbb{R}^d \times \mathbb{R}), b_i \in C^1_b(\mathbb{R}^d \times \mathbb{R})$, and

$$(3.15) \qquad \begin{aligned} |(a^*_{ij})_u(x, u) - (a^*_{ij})_u(x, \bar{u})| + |\nabla_x (a^*_{ij})(x, u) - \nabla_x (a^*_{ij})(x, \bar{u})|_d \\ \le C |u - \bar{u}|, \ \forall u, \bar{u} \in \mathbb{R}, \ x \in \mathbb{R}^d, \end{aligned}$$

for $i, j = 1, ..., d$.

We rewrite (3.10) as

$$(3.10)' \qquad u - \lambda \sum_{i,j=1}^d D^2_{ij}(a^*_{ij}(u)) + \lambda \operatorname{div}(b^*(x, u)) = f \text{ in } \mathcal{D}'(\mathbb{R}^d).$$

Equivalently, if $Du \in L^1_{\mathrm{loc}}$, then

$$(3.10)'' \qquad u - \lambda \sum_{i,j=1}^d D_i((a^*_{ij})_u(u) D_j u + (a^*_{ij})_{x_j}(x, u) u) + \lambda \operatorname{div}(b^*(x, u)) = f \text{ in } \mathcal{D}'(\mathbb{R}^d).$$

We also set

$$\begin{aligned} b_\infty &= \sup\{|b_i(x,u)|;\ (x,u)\in\mathbb{R}^d\times\mathbb{R},\ i=1,...,d\},\\ c_\infty &= \sup\{|(a_{ij})_{x_j}(x,u)|;\ (x,u)\in\mathbb{R}^d\times\mathbb{R},\ i,j=1,...,d\}. \end{aligned}$$

(By virtue of (K), the formulation $(3.10)''$ of $(3.10)'$ makes sense only if $D_j u \in L^1_{\rm loc}$.)

For each $N>0$, we set $B_N=\{\xi\in\mathbb{R}^d;\ |\xi|<N\}$. We have

Lemma 3.2 *Let $f\in L^2$ and $0<\lambda<\lambda_0=\gamma(b_\infty^2+c_\infty^2)^{-1}$. Then, for each N there is at least one solution $u_N\in H_0^1(B_N)$ to the equation*

$$\begin{aligned} &u-\lambda\sum_{i,j=1}^d D_{ij}^2(a_{ij}^*(u))+\lambda\ \mathrm{div}(b^*(x,u))=f \ \text{ in } B_N,\\ &u=0 \ \text{ on } \partial B_N, \end{aligned} \tag{3.16}$$

which satisfies the estimate

$$\|u_N\|^2_{L^2(B_N)}+\lambda\gamma\|\nabla u_N\|^2_{L^2(B_N)}\le C\|f\|^2_{L^2(B_N)}, \tag{3.17}$$

where C is independent of N and λ.

Proof. For $\rho>0$, we set $\mathcal{M}_\rho=\{v\in L^2(B_N);\ \|v\|_{L^2(B_N)}\le\rho\}$ and consider the operator $F:\mathcal{M}_\rho\to L^2(B_N)$ defined by $F(v)=u\in H_0^1(B_N)$, where u is the solution to the linear elliptic problem

$$\begin{aligned} &u-\lambda\sum_{i,j=1}^d D_i((a_{ij}^*)_v(x,v)D_ju+(a_{ij})_{x_j}(x,v)u)+\lambda\ \mathrm{div}(b(x,v)u)=f\\ &\hspace{8cm}\text{in } \mathcal{D}'(B_N),\\ &u=0 \text{ on } \partial B_N. \end{aligned} \tag{3.18}$$

By (3.14) and (H2), it follows via the Lax-Milgram lemma that, for each $v\in\mathcal{M}_\rho$ and $\lambda\in(0,\lambda_0)$, problem (3.18) has a unique solution $u=F(v)$. Moreover, by (3.18) and (H1), we see that

$$\begin{aligned} &\|u\|^2_{L^2(B_N)}+\gamma\lambda\|\nabla u\|^2_{L^2(B_N)}\\ &\quad\le\lambda b_\infty\|\nabla u\|_{L^2(B_N)}\|u\|_{L^2(B_N)}+c_\infty\lambda\|u\|_{L^2(B_N)}\|\nabla u\|_{L^2(B_N)}\\ &\quad+\|f\|_{L^2(B_N)}\|u\|_{L^2(B_N)}\\ &\quad\le\lambda b_\infty\|\nabla u\|_{L^2(B_N)}\rho+\rho\|f\|_{L^2(B_N)}+c_\infty\lambda\rho\|\nabla u\|_{L^2(B_N)}. \end{aligned} \tag{3.19}$$

Hence, for $\lambda \in (0, \lambda_0)$ and ρ suitably chosen, independent of N, $F(\mathcal{M}_\rho) \subset \mathcal{M}_\rho$.

Indeed, if $v_n \to v$ in $L^2(B_N)$ and $u_n = Fv_n$, we have $b(x, v_n) \to b(x, v)$,

$$(a^*_{ij})_v(x, v_n) \to (a^*_{ij})_v(x, v),\ (a_{ij})_{x_j}(x, v_n) \to (a_{ij})_{x_j}(x, v)$$

strongly in $L^2(B_N)$. Along a subsequence we have, by (3.19),

$$u_n \to u \text{ weakly in } H^1(B_N), \text{strongly in } L^2(B_N).$$

Now, letting $n \to \infty$ in equation (3.18), where $v = v_n$ and $u = u_n$, that is,

$$\int_{B_N} \Big(u_n\psi + \lambda \sum_{i,j=1}^{d} (a^*_{ij})_v(x, v_n) - D_j u_n + (a_{ij})_{x_j}(x, v_n), u_n\Big) D_i\psi dx$$
$$-\lambda \int_{B_N} u_n b^*(x, v_n) \cdot \nabla\psi dx = \int_{B_N} f\psi dx,\ \forall \psi \in C^\infty_0(B_N),$$

we see that $u = Fv$ and, therefore, F is continuous on $L^2(B_N)$.

Moreover, since the Sobolev space $H^1(B_N)$ is compactly embedded in $L^2(B_N)$, by (3.19) we see that $F(\mathcal{M}_\rho)$ is relatively compact in $L^2(B_N)$. Then, by the Schauder theorem, F has a fixed point $u_N \in \mathcal{M}_\rho$ which, clearly, is a solution to (3.16). Also, by (3.19), it follows that estimate (3.17) holds.

Lemma 3.3 *Let $f \in L^2(\mathbb{R}^d)$ and $\lambda < \lambda_0$. Then equation* (3.10) *has at least one solution $u \in H^1(\mathbb{R}^d)$ which satisfies the estimate*

$$|u|_2^2 + \gamma\lambda|\nabla u|_2^2 \le C(|f|_2^2 + 1). \tag{3.20}$$

Proof. Consider a sequence $\{N\} \to \infty$ and $u_N \in H^1_0(B_N)$ a solution to (3.16) given by Lemma 3.2. By (3.17), we have

$$\|u_N\|_{H^1_0(B_N)} \le C,\ \forall N,$$

and so, on a subsequence, again denoted $\{N\}$, we have

$$u_N \to u \text{ weakly in } H^1(\mathbb{R}^d),\ \text{strongly in } L^2_{\rm loc}(\mathbb{R}^d). \tag{3.21}$$

Then, letting $N \to \infty$ in the equation

$$u_N - \lambda \sum_{i,j=1}^{d} D_i((a^*_{ij})_u u_N) D_j u_N + ((a_{ij})_{x_j}(x, u_N) u_N) + \lambda\,\mathrm{div}(b(x, u_N) u_N) = f \text{ in } B_N,$$

or, more precisely, in its weak form

$$\int_{\mathbb{R}^d} u_N\psi\,dx + \lambda\sum_{i,j=1}^d \int_{\mathbb{R}^d} (a^*_{ij}(u_N)D_ju_N + (a_{ij})_{x_j}(x,u_N)u_N)D_i\psi\,dx$$
$$-\lambda\sum_{i=1}^d \int_{\mathbb{R}^d} b(x,u_N)u_N\cdot\nabla\psi\,dx = 0,\ \forall\psi\in C_0^\infty(\mathbb{R}^d),$$

we infer by (H1), (H3) and (3.21) that $u\in H^1(\mathbb{R}^d)$ is a solution to (3.10). Also, estimate (3.20) follows by (3.17). This completes the proof of Lemma 3.3.

Now, we come back to the proof of Proposition 3.1. We prove first that, for each $f\in L^2\cap L^1$ and $\lambda\in(0,\lambda_0)$, the solution $u=u(\lambda,f)\in H^1$ to equation (3.10) is unique and we have

$$|u(\lambda,f_1)-u(\lambda,f_2)|_1 \le |f_1-f_2|_1,\ \forall f_1,f_2\in L^2\cap L^1. \tag{3.22}$$

Here is the argument. We set $u_i=u(\lambda,f_i)$, $i=1,2$, and $f=f_1-f_2$, $u=u_1-u_2$. Then, we have

$$\begin{aligned} u-\lambda\sum_{i,j=1}^d D^2_{ij}(a^*_{ij}(x,u_1)-a^*_{ij}(x,u_2))& \\ +\lambda\,\mathrm{div}(b^*(x,u_1)-b^*(x,u_2)) &= f \text{ in } \mathcal{D}'(\mathbb{R}^d). \end{aligned}\tag{3.23}$$

More precisely, since $u_i\in H^1(\mathbb{R}^d)$, equation (3.23) is taken in its weak form

$$\begin{aligned}\int_{\mathbb{R}^d}\Big(u\psi+\lambda\sum_{i,j=1}^d D_i(a^*_{ij}(x,u_1)-a^*_{ij}(x,u_2))D_j\psi& \\ -\lambda(b^*(x,u_1)-b^*(x,u_2))\cdot\nabla\psi\Big)dx &= \int_{\mathbb{R}^d} f\psi\,dx,\ \forall\psi\in H^1(\mathbb{R}^d),\end{aligned}\tag{3.24}$$

In order to fix the idea of the proof, we invoke first a heuristic argument. Namely, if we multiply (3.23) by $\eta\in L^\infty(\mathcal{O})$, $\eta(x)\in\mathrm{sign}(u(x))$, a.e. $x\in\mathbb{R}^d$, and take into account that, by the monotonicity of the functions a^*_{ij},

$$\eta(x)\in\mathrm{sign}(a^*_{ij}(x,u_1(x))-a^*_{ij}(x,u_2(x))),\ \text{a.e. } x\in\mathbb{R}^d,$$

we get

$$|u|_1+\lambda\int_{\mathbb{R}^d}\sum_{i,j=1}^d D_i(a^*_{ij}(x,u_1(x))-a^*_{ij}(x,u_2(x)))D_j\eta(x)dx$$
$$+\lambda\int_{\mathbb{R}^d}\mathrm{div}(b^*(x,u_1)-b^*(x,u_2))\eta\,dx=\int_{\mathbb{R}^d} f\eta\,dx.$$

Taking into account that, by the monotonicity of $u \to a^*_{ij}(x,u)$, we have (formally)

$$D_i(a^*_{ij}(x,u_1(x)) - a^*_{ij}(x,u_2(x)))D_j\eta(x) \geq 0 \text{ in } \mathbb{R}^d,$$

while

$$((a^*_{ij})_{x_i}(x,u_1) - (a^*_{ij})_{x_i}(x,u_2))D_j\eta(x) = 0,$$

$$\int_{\mathbb{R}^d} \operatorname{div}(b^*(x,u_1) - b^*(x,u_2))\eta\, dx = \int_{[|u|=0]} (b^*(x,u_1) - b^*(x,u_2)) \cdot \nabla\eta\, dx = 0,$$

we get (3.22). This formal argument can be made rigorous by using a smooth approximation $\mathcal{X}_\delta$ of the signum graph. Namely, let $\mathcal{X}_\delta \in \mathrm{Lip}(\mathbb{R})$ be the function

$$\mathcal{X}_\delta(r) = \begin{cases} 1 & \text{for } r \geq \delta, \\ \dfrac{r}{\delta} & \text{for } |r| < \delta, \\ -1 & \text{for } r < -\delta, \end{cases}$$

where $\delta > 0$. We note that, since $u \in L^2$, it follows $\mathcal{X}_\delta(u) \in L^2$ and

$$\begin{aligned} G_\varepsilon = -\mathcal{X}_\delta(u)\Big(\sum_{i,j=1}^d D_iD_j(a^*_{ij}(x,u_1) - a^*_{ij}(x,u_2)) \\ + \operatorname{div}(b^*(x,u_1) - b^*(x,u_2))\Big) \in L^1, \end{aligned}$$

and, therefore,

$$\begin{aligned} \int_{\mathbb{R}^d} G_\varepsilon dx = \int_{\mathbb{R}^d} \Big(\sum_{i,j=1}^d D_j(a^*_{ij}(x,u_1) - a^*_{ij}(x,u_2))D_i\mathcal{X}_\delta(u) \\ + (b^*(x,u_1) - b^*(x,u_2)) \cdot \nabla\mathcal{X}_\delta(u)\Big)dx. \end{aligned}$$

Since $b^*(x,u_i) \in L^2$ and since, by (K), it follows that $a^*_{ij}(x,u_i) \in H^1(\mathbb{R}^d)$, $i = 1,2$, taking in (3.24) $\psi = \mathcal{X}_\delta(u)$ yields

$$\begin{aligned} &\int_{[|u(x)|\geq\delta]} |u(x)|dx + \frac{1}{\delta}\int_{[|u(x)|\leq\delta]} |u(x)|dx \\ &+ \lambda \sum_{i,j=1}^d \int_{\mathbb{R}^d} D_j(a^*_{ij}(x,u_1) - a^*_{ij}(x,u_2))D_i(\mathcal{X}_\delta(u))dx \\ &= \lambda \int_{\mathbb{R}^d} (b^*(x,u_1) - b^*(x,u_2)) \cdot \nabla(\mathcal{X}_\delta(u))dx + \int_{\mathbb{R}^d} f\mathcal{X}_\delta(u)dx. \end{aligned} \tag{3.25}$$

We set

$$
\begin{aligned}
I_\delta^1 &= \int_{\mathbb{R}^d} (b^*(x,u_1) - b^*(x,u_2)) \cdot \nabla(\mathcal{X}_\delta(u))dx \\
&= \int_{\mathbb{R}^d} (b^*(x,u_1) - b^*(x,u_2)) \cdot \nabla u \mathcal{X}'_\delta(u)dx \\
&= \frac{1}{\delta} \int_{[|u|\le\delta]} (b^*(x,u_1) - b^*(x,u_2)) \cdot \nabla u\, dx.
\end{aligned}
$$

Since, by hypothesis (K), $|b^*(x,u_1) - b^*(x,u_2)| \le C|u|(|u_1| + |u_2|)$, a.e. on $\mathbb{R}^d$, and $u_i \in L^2$, it follows that

$$
\lim_{\delta\to 0} \frac{1}{\delta} \int_{[|u|\le\delta]} |(b^*(x,u_1) - b^*(x,u_2)) \cdot \nabla u|dx \le C \lim_{\delta\to 0} \left(\int_{[|u|\le\delta]} |\nabla u|^2 dx \right)^{\frac{1}{2}} = 0,
$$

because $u \in H^1(\mathbb{R}^d)$ and $\nabla u = 0$ on $[x; u(x) = 0]$. This yields

$$
\lim_{\delta\to 0} I_\delta^1 = 0. \tag{3.26}
$$

On the other hand, taking into account that $u_i, a_{ij}^*(u_i) \in H^1(\mathbb{R}^d)$, for $i = 1, 2$, we have

$$
\begin{aligned}
I_\delta^2 &= \int_{\mathbb{R}^d} \sum_{i,j=1}^d D_j(a_{ij}^*(x,u_1) - a_{ij}^*(x,u_2)) D_i(\mathcal{X}_\delta(u))dx \\
&= \frac{1}{\delta} \int_{E_\delta} \sum_{i,j=1}^d ((a_{ij}^*)_u(x,u_1) D_j u_1 - (a_{ij}^*)_u(u_2) D_j u_2 \\
&\quad + (a_{ij}^*)_{x_j}(x,u_1) - (a_{ij}^*)_{x_j}(x,u_1)) D_i u\, dx \\
&= \frac{1}{\delta} \int_{E_\delta} \sum_{i,j=1}^d (a_{ij}^*)_u(x,u_1) D_j u D_i u\, dx \\
&\quad + \frac{1}{\delta} \int_{E_\delta} \sum_{i,j=1}^d ((a_{ij}^*)_u(x,u_1) - (a_{ij}^*)_u(x,u_2)) D_j u_2 D_i u\, dx \\
&\quad + \frac{1}{\delta} \int_{E_\delta} \sum_{i,j=1}^d ((a_{ij}^*)_{x_j}(x,u_1) - (a_{ij}^*)_{x_j}(x,u_2)) D_i u\, dx \\
&= K_1^\delta + K_2^\delta + K_3^\delta.
\end{aligned} \tag{3.27}
$$

Here, $E_\delta=\{x\in\mathbb{R}^d;\ |u(x)|\le\delta\}$. By (H2), it follows that $K_1^\delta\ge 0$, while by (3.15) we have

$$|(a_{ij}^*)_u(x,u_1)-(a_{ij}^*)_u(x,u_2)|+|(a_{ij}^*)_{x_j}(x,u_1)-(a_{ij}^*)_{x_j}(x,u_2)|\le C|u|.$$

Taking into account that $u_i\in H^1(\mathbb{R}^d),\ i=1,2$, and that

$$\lim_{\delta\to 0}\int_{[|u|\le\delta]}|\nabla u(x)|^2dx=0, \tag{3.28}$$

we infer that $\lim\limits_{\delta\to 0}K_i^\delta=0$, $i=2,3$, and so, by (3.27) it follows that

$$\liminf_{\delta\to 0} I_\delta^2\ge 0.$$

This yields

$$|u|_1\le|f|_1,\ \forall\lambda\in(0,\lambda_0).$$

To resume, we have shown so far that under assumptions (H1)–(H3) and (K), for each $f\in L^2$, equation (3.10) has, for $\lambda\in(0,\lambda_0)$, a unique solution $u(\lambda,f)\in H^1(\mathbb{R}^d)$ which satisfies (3.20) and (3.22).

Now, we assume that a_{ij},b_i satisfy (H1)–(H3) only and consider, for $\varepsilon>0$, the functions

$$(a_{ij}^*)_\varepsilon(x,u)=\int_{\mathbb{R}}a_{ij}^*(x-\varepsilon y,u-\varepsilon v)\rho(y,v)dv\,dy,\ i,j=1,...,d, \tag{3.29}$$

$$b_i^\varepsilon(x,u)=\int_{\mathbb{R}}b_i(x-\varepsilon y,u-\varepsilon v)\rho(y,v)dv\,dy,\ i,j=1,...,d, \tag{3.30}$$

where $\rho\in C_0^\infty(\mathbb{R}^d\times\mathbb{R})$, $\int_{\mathbb{R}^d\times\mathbb{R}}\rho(y,v)dy\,dv=1$, $\rho\ge 0$, is a standard mollifier. Clearly, $(a_{ij}^*)_\varepsilon$, b_i^ε satisfy condition (K). We set $b^\varepsilon=\{b_i^\varepsilon\}_{i=1}^d$. Then, as shown above, the equation

$$u_\varepsilon-\lambda\sum_{i,j=1}^d D_{ij}^2(a_{ij}^*)_\varepsilon(x,u_\varepsilon)+\lambda\ \mathrm{div}(b^\varepsilon(x,u_\varepsilon)u_\varepsilon)=f \tag{3.31}$$

has, for each $\lambda\in(0,\lambda_0)$ and $f\in L^2\cap L^1$, a unique solution $u_\varepsilon=u_\varepsilon(\lambda,f)\in H^1(\mathbb{R}^d)$ satisfying (3.20) and (3.22). Hence

$$|u_\varepsilon(\lambda,f_1)-u_\varepsilon(\lambda,f_2)|\le|f_1-f_2|_1,\ \forall\varepsilon>0,\ f_1,f_2\in L^2\cap L^1. \tag{3.32}$$

$$|u_\varepsilon(\lambda,f)|_2^2+\gamma\lambda|\nabla u_\varepsilon(\lambda,f)|_2^2\le C(|f|_2^2+1),\ \forall\varepsilon>0,\ f\in L^2. \tag{3.33}$$

(We note that, by Lemma 3.2, λ_0 is independent of ε, because

$$\sup_{\varepsilon}\{|(a^{\varepsilon}_{ij})_{x_j}|_\infty + |b_i|_\infty;\ i,j=1,...,d\} < \infty.)$$

Now, for $\varepsilon \to 0$, it follows by the compactness of $H^1(\mathbb{R}^d)$ in $L^2_{\rm loc}$ that along a subsequence, again denoted ε, we have

$$u_\varepsilon(\lambda, f) \longrightarrow u \text{ strongly in } L^2_{\rm loc}$$

and so, by (3.29), (3.30), we have

$$\begin{array}{rcll} (a^*_{ij})_\varepsilon(x, u_\varepsilon(x)) & \longrightarrow & a^*_{ij}(x, u(x)), & \text{a.e. } x \in \mathbb{R}^d, \\ b^\varepsilon_i(x, u_\varepsilon(x)) & \longrightarrow & b_i(x, u(x)), & \text{a.e. } x \in \mathbb{R}^d, \end{array}$$

as $\varepsilon \to 0$. Hence, for $\varepsilon \to 0$,

$$\begin{array}{rcll} D^2_{ij}((a^*_{ij})_\varepsilon(x, u_\varepsilon)) & \longrightarrow & D^2_{ij}(a^*_{ij}(x, u)) & \text{in } \mathcal{D}'(\mathbb{R}^d), \\ \operatorname{div}(b^\varepsilon(x, u_\varepsilon)u_\varepsilon) & \longrightarrow & \operatorname{div}(b(x, u)u) & \text{in } \mathcal{D}'(\mathbb{R}^d). \end{array}$$

and so $u = u(\lambda, f)$ is a solution to (3.10). Moreover, by (3.32) it follows that

$$|u(\lambda, f_1) - u(\lambda, f_2)|_1 \le |f_1 - f_2|_1,\ \forall f_1, f_2 \in L^2. \tag{3.34}$$

Now, we fix $f \in L^1$ and consider a sequence $\{f_n\} \subset L^2$ such that $f_n \to f$ in L^1 and consider the corresponding solution $u_n = u(\lambda, f_n)$ to (3.10). By (3.34), we see that

$$|u_n - u_m|_1 \le |f_n - f_m|_1,\ \forall n, m \in \mathbb{N}.$$

Hence, there is $u^* = \lim\limits_{n\to\infty} u_n$ in L^1. Moreover, by (H1), we see that, for $n \to \infty$,

$$a^*_{ij}(u_n) \to a_{ij}(u^*), \text{ a.e. in } \mathbb{R}^d$$

and, since $a_{ij} \in C_b(\mathbb{R}^d \times \mathbb{R})$, we have

$$D^2_{ij} a^*_{ij}(u_n) \to D^2_{ij} a^*_{ij}(u^*) \text{ in } \mathcal{D}'(\mathbb{R}^d),$$

for all $i, j = 1, 2, ..., d$. Similarly, by (H3) we see that

$$\operatorname{div}(b(x, u_n)u_n) \to \operatorname{div}(b(x, u^*)u^*) \text{ in } \mathcal{D}'(\mathbb{R}^d).$$

We have, therefore,

$$\sum_{i,j=1}^{d} D_{ij}^2 a_{ij}^*(u_n) - \operatorname{div}(b(x,u_n)u_n) \to \sum_{i,j=1}^{d} D_{ij}^2 a_{ij}^*(u^*) - \operatorname{div}(b(x,u^*)u^*) \quad \text{strongly in } L^1. \tag{3.35}$$

Then, letting $n \to \infty$ in equation (3.10), where $f = f_n$, $u = u_n$, we see that $u^* = u(\lambda, f)$ is the solution to (3.10). Moreover, by (3.34), the inequality (3.11) follows for all $\lambda \in (0, \lambda_0]$. This means that

$$|(I + \lambda A)^{-1} f_1 - (I + \lambda A)^{-1}|_1 \le |f_1 - f_2|_1,\ \forall f_1, f_2 \in L^1,$$

for all $\lambda \in [0, \lambda_0)$. By Proposition 3.1 in [1], this implies that the above inequality holds for all $\lambda > 0$. Hence, (3.11) follows for all $\lambda > 0$, as claimed.

As regards (3.13), it first follows by equation (3.10), where $f \in L^2$ and $u \in H^1(\mathbb{R}^d)$, by integrating over $\mathbb{R}^d$. Then, by density, it extends to all of $f \in L^1$. Finally, (3.12) for $f \in L^2$, $f \ge 0$, follows by multiplying (3.10) with $\operatorname{sign}(u^-)$ (or, more exactly, by $\mathcal{X}_\delta(u^-)$ and letting $\delta \to 0$) and integrating over $\mathbb{R}^d$. This completes the proof of Proposition 3.1 under hypotheses (H1)–(H3).

Now, we are ready to formulate the existence theorem for equation (3.1). As mentioned earlier, we shall represent equation (3.1) as the evolution equation (3.2) in $\mathcal{X} = L^1$, where the operator A is defined by (3.8)-(3.9). By a *weak solution to equation* (3.1), we mean *a mild solution to equation* (3.2), where $\mathcal{X} = L^1$ and A is the operator defined by (3.8), (3.9).

We have

Theorem 3.4 *Assume that hypotheses* (H1)–(H3) *hold. Then, for each* $u_0 \in L^1(\mathbb{R}^d)$*, there is a unique weak solution* $u = u(\cdot, u_0) \in C([0,\infty); L^1)$ *to equation* (3.1)*. Moreover,* u *has the following properties*

$$|u(t,u_0^1) - u(t,u_0^2)|_1 \le |u_0^1 - u_0^2|_1,\ \forall u_0^1, u_0^2 \in L^1,\ t \ge 0, \tag{3.36}$$

$$u \ge 0 \ \textit{a.e. in } (0,\infty) \times \mathbb{R}^d \ \textit{if } u_0 \ge 0 \ \textit{a.e. in } \mathbb{R}^d, \tag{3.37}$$

$$\int_{\mathbb{R}^d} u(t,x)dx = \int_{\mathbb{R}^d} u_0(x)dx, \forall u_0 \in L^1,\ t \ge 0, \tag{3.38}$$

and u *is a solution to* (3.1) *in the sense of Schwartz distributions on* $(0,\infty)\times\mathbb{R}^d$, (*see* (1.2))*, that is,*

$$\begin{aligned}\int_0^\infty \int_{\mathbb{R}^d} (u(t,x)\varphi_t(t,x) + \sum_{i,j=1}^{d} a_{ij}(x, u(t,x))u(t,x) D_{ij}^2 \varphi(t,x) \\ + b(x,u) \cdot \nabla_x \varphi(t,x) u(t,x)) dt\, dx = 0,\ \forall \varphi \in C_0^\infty((0,\infty) \times \mathbb{R}^d).\end{aligned} \tag{3.39}$$

Proof. As mentioned above, the existence of a mild solution u for (3.2), which by our definition is a weak solution to (3.1), follows by the Crandall and Liggett theorem by virtue of Proposition 3.1, which implies the m-accretivity of the operator A defined by (3.8)–(3.9). The solution can be equivalently expressed by the exponential formula

$$u(t,u_0)=\lim_{n\to\infty}\left(I+\frac{t}{n}\,A\right)^{-n}u_0,\ \ \forall t\ge 0,\ \ u_0\in\overline{D(A)}=L^1. \tag{3.40}$$

Then, by (3.11)-(3.13), we get for $u=u(t,u_0)$ the corresponding properties (3.36)-(3.38) and this completes the proof. In particular, it follows that, if u_0 is a probability density, that is, $u_0dx\in\mathcal{P}(\mathbb{R}^d)$, then so is $u(t,u_0)$ for all $t\ge 0$. Note also that $t\to u(t,u_0)$ is a continuous semigroup of nonexpansive operators in the space L^1. As regards (3.39), it follows by letting $h\to 0$ in the equation

$$\begin{aligned}\int_0^\infty\int_{\mathbb{R}^d}\Big(u_h(t,x)(\varphi(t+h,x)-\varphi(t,x))+\sum_{i,j=1}^d a_{ij}(x,u_h(t,x))u_h(t,x)\\ D^2_{ij}\varphi(t,x)+b(x,u_h(t,x))\cdot\nabla_x\varphi(t,x)u_h(t,x)\Big)dt\,dx=0,\\ \forall\varphi\in C_0^\infty((0,\infty)\times\mathbb{R}^d).\end{aligned}$$

Remark 3.5 Assumptions $a_{ij}\in C^2(\mathbb{R}^d\times\mathbb{R})$ and $b_i\in C^1(\mathbb{R}^d\times\mathbb{R})$ in (H1)–(H3) were necessary for the density of $D(A)$ in L^1. Otherwise, it suffices to take only $a_{ij}\in C^1(\mathbb{R}^d\times\mathbb{R})$. In the special case $a_{ij}\equiv\beta\delta_{ij}$, the density of $D(A)$ follows, however, under the weaker condition $\beta\in C^1(\mathbb{R}^d)$ (see [7]).

3.2 Existence for degenerate FPEs

We consider here the equation

$$\begin{aligned}&u_t-\sum_{i,j=1}^d D^2_{ij}(a_{ij}(u)u)+\sum_{i=1}^d D_i(b_i(u)u)=0\text{ in }\mathcal{D}'((0,\infty)\times\mathbb{R}^d),\\ &u(0,x)=u_0(x),x\in\mathbb{R}^d,\end{aligned}\tag{3.41}$$

where a_{ij} and b_i satisfy hypotheses (H1)$'$–(H3)$'$.

Consider the operator $A_1 : D(A_1) \subset L^1 \to L^1$ defined by

$$
(3.42) \qquad \begin{aligned}
&A_1 u = -\sum_{i,j=1}^{d} D_{ij}^2(a_{ij}(u)u) + \sum_{i=1}^{d} D_i(b_i(u)u) \text{ in } \mathcal{D}'(\mathbb{R}^d),\\
&D(A_1) = \left\{ u \in L^1; -\sum_{i,j=1}^{d} D_{ij}^2(a_{ij}(u)u) + \sum_{i=1}^{d} D_i(b_i(u)u) \in L^1 \right\}.
\end{aligned}
$$

We have

Lemma 3.6 *Assume that* (H1)′–(H3)′ *hold. Then the operator A_1 is m-accretive in L^1.*

Proof. One should prove that, for each $\lambda \in (0, \lambda_0)$ and $f \in L^1$, the equation

$$
(3.43) \qquad u - \lambda \sum_{i,j=1}^{d} D_{ij}^2(a_{ij}(u)u) + \lambda \sum_{i=1}^{d} D_i(b_i(u)u) = f \text{ in } \mathcal{D}'(\mathbb{R}^d)
$$

has a unique solution $u = u(\lambda, f)$ which satisfies the estimate

$$
(3.44) \qquad |u(\lambda, f_1) - u(\lambda, f_2)|_1 \le |f_1 - f_2|_1, \ \forall f_1, f_2 \in L^1.
$$

We set, for each $\varepsilon > 0$,

$$
(3.45) \qquad a_{ij}^{\varepsilon}(r) = a_{ij}(r) + \varepsilon\delta_{ij}, \ i, j = 1, ..., d, \ r \in \mathbb{R},
$$

where δ_{ij} is the Kronecker symbol. Then, we approximate (3.43) by

$$
(3.46) \qquad u - \lambda \sum_{i,j=1}^{d} D_{ij}^2(a_{ij}^{\varepsilon}(u)u) + \lambda \sum_{i=1}^{d} D_i(b_i(u)u) = f \text{ in } \mathcal{D}'(\mathbb{R}^d).
$$

Equivalently,

$$
(3.47) \qquad u + \lambda A_1^{\varepsilon}(u) = f,
$$

where

$$
\begin{aligned}
A_1^{\varepsilon}(u) &= -\sum_{i,j=1}^{d} D_{ij}^2(a_{ij}^{\varepsilon}(u)u) + \sum_{i=1}^{d} D_i(b_i(u)u), \ \forall u \in D(A_1^{\varepsilon}),\\
D(A_1^{\varepsilon}) &= \left\{ u \in L^1; \ -\sum_{i,j=1}^{d} D_{ij}^2(a_{ij}^{\varepsilon}(u)u) + \sum_{i=1}^{d} D_i(b_i(u)u) \in L^1 \right\}.
\end{aligned}
$$

We shall prove that, for each $f \in L^1$, there is a solution $u = u_\varepsilon(\lambda, f)$ satisfying (3.44) for $0 < \lambda < \lambda_0$.

Since a_{ij}^ε and b_i satisfy, for each $\varepsilon > 0$, hypotheses (H1)–(H3), Proposition 3.1 implies the existence of a solution $u_\varepsilon = u_\varepsilon(\lambda, f)$ to (3.46) in $L^1(\mathbb{R}^d)$ for each $f \in L^2$ if $0 < \lambda \le \lambda_0^\varepsilon = \frac{C}{\varepsilon}$, with C independent of ε.

Moreover, one has

$$|u_\varepsilon(\lambda, f_1) - u_\varepsilon(\lambda, f_2)|_1 \le |f_1 - f_2|_1,\ \forall f_1, f_2 \in L^2,\ \lambda \in (0, \lambda_0^\varepsilon). \tag{3.48}$$

Then, by density, $u_\varepsilon(\lambda, f)$ extends as solution to (3.46) for all $f \in L^1$.

Note also that, by (3.11)–(3.13), we have, for all $\varepsilon > 0$ and $\lambda \in (0, \lambda_0^\varepsilon)$,

$$\int_{\mathbb{R}^d} (I + \lambda A_1^\varepsilon)^{-1} f\, dx = \int_{\mathbb{R}^d} f\, dx, \quad \forall f \in L^\infty, \tag{3.49}$$

$$(I + \lambda A_1^\varepsilon)^{-1} f \ge 0,\ \text{a.e. in } \mathbb{R}^d \text{ if } f \ge 0,\ \text{a.e. in } \mathbb{R}^d, \tag{3.50}$$

while (3.48) yields

$$|I+\lambda A_1^\varepsilon)^{-1} f_1 - (I+\lambda A_1^\varepsilon)^{-1} f_2|_1 \le |f_1 - f_2|_1, \forall f_1, f_2 \in L^1, \varepsilon > 0. \tag{3.51}$$

Though (3.49)–(3.51) were proved only for $0 < \lambda \le \lambda_0^\varepsilon$, it can be shown, however, as mentioned earlier, that $(I + \lambda A_1^\varepsilon)^{-1}$ extends to all $\lambda > 0$ by a well known argument based on the resolvent equation

$$(I + \lambda A_1^\varepsilon)^{-1} f = (I + \lambda_0 A_1^\varepsilon)^{-1} \left(\frac{\lambda_0^\varepsilon}{\lambda} f + \left(1 - \frac{\lambda_0^\varepsilon}{\lambda}\right) (I + \lambda A_1^\varepsilon)^{-1} f \right),\ \lambda > \lambda_0^\varepsilon.$$

(See [1], Proposition 3.3.)

Now, we are going to let $\varepsilon \to 0$ in (3.46). We set, for $f \in L^1$ and the solution u_ε to (3.46),

$$u_h^\varepsilon(x) = u_\varepsilon(x + h) - u_\varepsilon(x),\ f_h(x) = f(x + h) - f(x),\ x, h \in \mathbb{R}^d.$$

Since a_{ij}^ε and b_i^ε are independent of x, we see that $x \to u^\varepsilon(x + h)$ is the solution to (3.46) for $f(x) = f(x + h)$. Then, by (3.51), it follows that

$$|u_h^\varepsilon|_1 \le |f_h|_1,\ \forall h \in \mathbb{R}^d,\ \varepsilon > 0.$$

By the Kolmogorov compactness theorem (see, e.g., [11], p. 111), it follows that $\{u^\varepsilon\}$ is compact in $L^1_{\rm loc}(\mathbb{R}^d)$ and so, along a subsequence,

$$u_\varepsilon \to u \text{ strongly in } L^1_{\rm loc}(\mathbb{R}^d) \text{ for } \varepsilon \to 0.$$

Since $|u_\varepsilon|_1 \le C$, $\forall \varepsilon > 0$, it follows via Fatou's lemma that $u \in L^1$. Letting $\varepsilon \to 0$ in (3.46), where $u = u_\varepsilon$, and taking into account that

$$a^\varepsilon_{ij}(u_\varepsilon)u_\varepsilon \to a_{ij}(u)u, \ \ b_i(u_\varepsilon)u_\varepsilon \to b_i(u)u, \text{ a.e. in } \mathbb{R}^d,$$

while by (H1)$'$, (H3)$'$,

$$|a^\varepsilon_{ij}(u_\varepsilon)| + |b_{ij}(u_\varepsilon)| \le C, \text{ a.e. in } \mathbb{R}^d,$$

where C is independent of ε, we see that u is a solution to (3.46) and so $u = (I + \lambda A_1)^{-1} f$. Moreover, letting $\varepsilon \to 0$ in (3.49)–(3.51), we see that

$$|(I+\lambda A_1)^{-1} f_1 - (I+\lambda A_1)^{-1} f_2|_1 \le |f_1 - f_2|_1, \ \forall \lambda > 0, \ f_1, f_2 \in L^1, \tag{3.52}$$

$$\int_{\mathbb{R}^d} (I + \lambda A_1)^{-1} f \, dx = \int_{\mathbb{R}^d} f \, dx, \ \forall f \in L^1, \ \lambda > 0, \tag{3.53}$$

$$(I + \lambda A_1)^{-1} f \ge 0, \text{ a.e. in } \mathbb{R}^d \text{ if } f \ge 0, \text{ a.e. in } \mathbb{R}^d. \tag{3.54}$$

Then, by the Crandall and Liggett existence theorem, for each $u_0 \in \overline{D(A_1)} = L^1$, the differential equation

$$\begin{aligned} &\frac{du}{dt} + A_1 u = 0, \ t > 0, \\ &u(0) = u_0, \end{aligned} \tag{3.55}$$

has a unique mild solution $u \in C([0,\infty); L^1)$ in the sense of (3.4)–(3.6).

As in the previous case, this mild solution is, in fact, a solution to the Fokker-Planck equation (3.41) in the sense of Schwartz distributions (cf. (3.39)). We have, therefore, the following existence result.

Theorem 3.7 *Under hypotheses* (H1)$'$–(H3)$'$, *for each $u_0 \in L^1$, there is a unique weak solution $u = u(t, u_0) \in C([0,\infty); L^1)$ to equation* (3.41). *Moreover, this solution satisfies* (3.36)–(3.38) *and is a solution to* (3.41) *in the sense of Schwartz distributions on $(0,\infty) \times \mathbb{R}^d$, i.e., in the sense of* (3.39) *or* (1.2).

Remark 3.8 In particular, Theorems 3.4 and 3.7 imply the existence of a solution u in the sense of Schwartz distributions on $(0,\infty) \times \mathbb{R}^d$ for equation (3.1). Moreover, $u : [0,\infty) \to L^1$ is continuous. In some special cases, these two properties are sufficient to characterize such solutions to (3.1). In fact, this is the case if (see [12]) $b \equiv 0$ and

$$a_{ij}(x,u)u = \delta_{ij}\beta(u)u,\ \forall u \in \mathbb{R},\ i,j = 1,...,d,$$

where β is a continuous monotonically nondecreasing function because, in this case, one has the uniqueness of distributional solutions $u \in L^\infty((0,\infty)\times\mathbb{R}^d) \cap C([0,\infty);L^1)$. Such a result remains, however, open for general Fokker-Planck equations as in (3.1).

Remark 3.9 In the special case $a_{ij} = \delta_{ij}$, the weak solution u given by Theorem 3.7 is an entropic solution in sense of S. Kruzkov [20] for equation (3.1). In the present case, the solution u given by Theorem 3.7 is a *mild* solution to (3.1) defined, as in the previous case, by the finite difference scheme (3.4)–(3.6). It is, of course, a continuous in t distributional solution to (3.1), but we do not know if it is unique within this class. In fact, we should mention that the solutions u given by Theorems 3.4 and 3.7 are unique in the class of mild solutions generated by the operator A and not in the class of distributional or entropic solutions in the sense of Kruzkov.

4 Solution of the McKean-Vlasov SDE

Consider the following McKean-Vlasov SDE for $T \in (0,\infty)$

$$\begin{aligned} dX(t) &= b\left(X(t), \frac{d\mathcal{L}_{X(t)}}{dx}(X(t))\right)dt \\ &\quad +\sqrt{2}\,\sigma\left(X(t), \frac{d\mathcal{L}_{X(t)}}{dx}(X(t))\right)dW(t),\ 0 \le t \le T, \\ X(0) &= \xi_0, \end{aligned} \tag{4.1}$$

on $\mathbb{R}^d$, where $W(t),\ t \geq 0$, is an $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion on a probability space $(\Omega,\mathcal{F},P)$ with normal filtration $(\mathcal{F}_t)_{t\geq 0}$ and $\xi_0 : \Omega \to \mathbb{R}^d$ is $\mathcal{F}_0$-measurable such that

$$P \circ \xi_0^{-1}(dx) = u_0(x)dx.$$

Furthermore, $b = (b_1,...,b_d) : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \mathbb{R} \to L(\mathbb{R}^d;\mathbb{R}^d)$ are measurable.

Let $a_{ij} := 2(\sigma\sigma^T)_{ij},\ 1 \le i,j \le d$. Then, as an immediate consequence of Section 2 and Theorems 3.4 and 3.7, respectively, we obtain the following.

Theorem 4.1 *Suppose that a_{ij}, b_i, $1 \le i, j \le d$, satisfy either* (H1)–(H3) *or* (H1)$'$–(H3)$'$. *Then there exists a* (*in the probabilistic sense*) *weak solution to DDSDE* (4.1). *Furthermore, for the solution u in Theorem* 3.4 *and* 3.7, *respectively, with $u(0, \cdot) = u_0$, we have the "probabilistic representation"*

$$u(t, x)dx = P \circ X(t)^{-1}(dx), \ t \ge 0.$$

Remark 4.2

(i) In the case where in (4.1) we have $a_{ij}(x, u) = \delta_{ij}\beta(u)$, $1 \le i, j \le d$, and $\beta : \mathbb{R} \to 2^{\mathbb{R}}$ is maximal monotone with $\sup\{|s| : s \in \beta(r)\} \le C|r|^m$, $r \in \mathbb{R}$, for some $C, m \in [0, \infty)$ and b satisfies (H3)$'$, then the above theorem was already proved in [4]. The special case where, in addition, $b \equiv 0$, $d = 1$ and $m = 4$, was proved in [9] if $\beta(r)/r$ is nondegenerate at $r = 0$ and in [10] including the degenerate case.

(ii) The special case $d = 1$, $b \equiv 0$, $a_{ij}(x, u) = \delta_{ij}\beta(u)$, $1 \le i, j \le d$, with $\beta(r) := r|r|^{m-1}$, $r \in \mathbb{R}$, for some $m \in (1, \infty)$, was proved in [8].

(iii) [6] contains an analogous result as in [9], [10] in the case where a linear multiplicative noise is added to the nonlinear FPE, which thus becomes a stochastic porous media equation.

Our final remark concerns the uniqueness of the time marginal of solutions to (4.1).

Remark 4.3 If $b \equiv 0$ and $a_{ij}(x, u) = \delta_{ij}\beta(u)$, $1 \le i, j \le d$, and $\beta : \mathbb{R} \to \mathbb{R}$ is continuous, nondecreasing and $\beta(0) = 0$, then (3.1) has a unique solution among all the solutions in $(L^\infty \cap L^1)((0, T) \times \mathbb{R}^d)$ by the main result in [12]. Hence, obviously, we have uniqueness of the time marginals for weak solutions to (4.1) among all the solutions of (4.1) whose time marginals have densities in $L^\infty((0, T) \times \mathbb{R}^d)$.

Acknowledgements. This work was supported by the DFG through CRC 1283. V. Barbu was partially supported by the CNCS-UEFISCDI project PN-III-P4-ID-PCE-2015-0011.

References

[1] Barbu, V., *Nonlinear Differential Equations of Monotone Type in Banach Spaces*, Springer, 2010.

[2] Barbu, V., Generalized solutions to nonlinear Fokker-Planck equations, *J. Diff. Equations*, 261 (2016), 2446-2471.

[3] Barbu, V., Röckner, M., Nonlinear Fokker-Planck equations driven by Gaussian linear multiplicative noise, *J. Differential Equations* (2018).

[4] Barbu, V., Röckner, M., Probabilistic representation for solutions to nonlinear Fokker-Planck equation, *SIAM J. Math. Anal.*, 50 (2018), 2588-2607.

[5] Barbu, V., Röckner, M., Russo, F., Probabilistic representation for solutions of an irregular porous media type equation: the degenerate case, *Probab. Theory Rel. Fields,* 15 (2011), 1-43.

[6] Barbu, V., Röckner, M., Russo, F., Doubly probabilistic representation for the stochastic porous media equation, *Ann. Inst. Henri Poinfcaré*, 53 (4) (2017), 1061-1090.

[7] Barbu, V., Röckner, M., The evolution to equilibrium of solutions to nonlinear Fokker-Planck equations, arXiv:1904.08291.

[8] Benachour, S., Chassaing, P., Roynette, B., Vallois, P., Processus associés à l'équation des milieux poreux, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.*, (4) (1997), 23 (4) (1996), 793–832.

[9] Blanchard, Ph., Röckner, M., Russo, F., Probabilistic representation for solutions of an irregular porous media equation, *Ann. Probab.*, 38 (2010), 1870-1900.

[10] Bogachev, V.I., Krylov, N.V., Röckner, M., Stanislav V. Shaposhnikov, S.V., *Fokker-Planck-Kolmogorov equations*, Mathematical Surveys and Monographs, vol. 207, American Mathematical Society, Providence, RI, 2015.

[11] Brezis, H., *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, New York. Dordrecht. Heidelberg. London, 2011.

[12] Brezis, H., Crandall, M.G., Uniqueness of solutions of the initial-value problem for $u_t - \Delta\beta(u) = 0$, *J. Math. Pures et Appl.*, 58 (1979), 153-163.

[13] dos Reis, G., Smith, G., Tankov, P., Importance sampling for McKean-Vlasov SDEs, arXiv:1803.09320, 2018.

[14] Figalli, A., Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients, *J. Funct. Anal.*, 254 (1) (2008), 109–153.

[15] Franck, T.D., *Nonlinear Fokker-Planck Equations. Fundamentals and Applications*, Springer, Berlin. Heidelberg. New York, 2005.

[16] Funaki, T., A certain class of diffusion processes associated with nonlinear parabolic equations, *Z. Wahrsch. Verw. Gebiete*, 67 (3) (1984), 331-348.

[17] Hammersley, W., Šiška, D., Szpruch, L., McKean-Vlasov SDEs under measure mependent Lyapunov conditions, arXiv:1802.03974, 2018.

[18] Huang, X., Röckner, M., Wang, F.-Y., Nonlinear Fokker-Planck equations for probability measures on path space and path-distribution dependent SDEs, arXiv:1709.00556, 2017.

[19] Huang, X., Wang, F.-Y., Distribution dependent SDEs with singular coefficients, arXiv:1805.01682, 2018.

[20] Kruzkov, S., First order quasilinear equations with several independent variables, *Sbornic: Mathematics*, 10 (2) (1970), 217-243.

[21] McKean, Jr., H.P., A class of Markov processes associated with nonlinear parabolic equations, *Proc. Nat. Acad. Sci. U.S.A.*, 56 (1966), 1907-1911.

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

[23] Mehri, S., Scheutzow, M., Stannat, W., Zangeneh, B.Z., Propagation of Chaos for Stochastic Spatially Structured Neuronal Networks with Fully Path Dependent Delays and Monotone Coefficients driven by Jump Diffusion Noise, arXiv:1805.01654, 2018.

[24] Mishura, Yu.S., Veretennikov, A.Yu., Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations, arXiv:1603.02212, 2016.

[25] Röckner, M., Russo, F., Uniqueness of a class of stochastic Fokker-Planck and porous media equations, *J. Evol. Equ.*, 17 (3) (2017), 1049-1062.

[26] Scheutzow, M., Uniqueness and nonuniqueness of solutions of Vlasov-McKean equations, *J. Austral. Math. Soc. Ser. A*, 43 (2) (1987), 246-256.

[27] Stroock D.W., Srinivasa Varadhan, S.R., *Multidimensional diffusion processes*, Springer, Berlin. New York, 1997.

[28] Sznitman, A.-S., Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated, *J. Funct. Anal.*, 56 (3) (1984), 311-336.

[29] Trevisan, D., Well-posedness of multidimensional diffusion processes, *Electronic J. Probab.*, 21 (2016), 1-41.

[30] Wang, F.-Y., Distribution dependent SDEs for Landau type equations, *Stochastic Process. Appl.*, 128 (2) (2018), 595-621.