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

Viorel Barbu∗ Michael Röckner†‡

Abstract

We construct weak solutions to a class of distribution dependent SDE, of type

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

on $\mathbb{R}^d$ for possibly degenerate diffusion matrices $\sigma$ 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

$$dX(t) = b(t, X(t), \mathcal{L}_{X(t)})dt + \sigma(t, X(t), \mathcal{L}_{X(t)})dW(t)$$

$$X(0) = \xi_0,$$

∗Octav Mayer Institute of Mathematics of the Romanian Academy, Iași, 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 $b, \sigma$ 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 $\xi_0$ 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], [20], [21], [25], [27], and, e.g., the more recent papers [13], [17], [18], [19], [22], [23] and [29].

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 $\mu_0 := \text{law of } \xi_0$, 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 $t \geq 0$,

\begin{equation}
\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,
\end{equation}

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

\begin{equation}
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),
\end{equation}

is the corresponding Kolmogorov operator. For equations of type (1.2), we refer the reader, e.g., to [9]. We note that (1.2) is also shortly written as

\begin{equation}
\partial_t \mu_t = L_{\mu_t}^* \mu_t \text{ with } \mu_0 \text{ given.}
\end{equation}

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{equation}
\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)] - \sum_{i=1}^d \frac{\partial}{\partial x_i} [b_i(t, x, u(t, \cdot))dx]u(t, x)]
\end{equation}

\begin{equation}
u(0, x) = u_0(x) \left(= \frac{d\mu_0}{dx}(x)\right),
\end{equation}
\[ x \in \mathbb{R}^d, \ t \geq 0, \text{ or, shortly,} \]

\[
\partial_t u = \frac{1}{2} \partial_i \partial_j (a_{ij}(u) u) - \partial_i (b_i(u) u), \quad u(0, \cdot) = u_0.
\]

In this paper, we want to go in the opposite direction, that is, we first want to solve (1.2) and, using the obtained \( \mu_t, t \geq 0 \), 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 [28] (generalizing earlier work by Figalli [14]), in connection with a classical result by Stroock and Varadhan (see, e.g., [26]) 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 [28], one can also easily prove that, if (1.1) has a unique solution in law, then the solution to (1.1) 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 \( \mu \) in the following way:

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

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

\[
dX(t) = \bar{b} \left( t, X(t), \frac{dL_X(t)}{dx} (X(t)) \right) dt
\]

\[
+ \bar{\sigma} \left( t, X(t), \frac{dL_X(t)}{dx} (X(t)) \right) dW(t),
\]

\[
X(0) = \xi_0,
\]

with \( (\bar{\sigma}\bar{\sigma}^T)_{ij} = \bar{a}_{ij} \).
In particular, we obtain a probabilistic representation of the solution μ_t, t ≥ 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 its dependence on the law 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 L_{X(t)} we choose in (1.8), since we are looking only for solutions of (1.8) in the class with L_{X(t)} being absolutely continuous with respect to dx. Precise conditions on the coefficients b_i, 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 3.6 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]–[8]. 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 O ⊂ R^d, we denote by L^p(O), 1 ≤ p ≤ ∞, the standard Lebesgue p-integrable functions on O, and by H^1(O), the Sobolev space \{u ∈ L^2(O); \nabla u ∈ L^2(O)\}.

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

We also set L^p = L^p(R^d) with the norm denoted |·|_p and L^p_{loc} = L^p_{loc}(R^d), 1 ≤ p ≤ ∞. By \mathcal{D}'(R^d) and \mathcal{D}'((0, ∞) × R^d), we denote the space of distributions on R^d and (0, ∞) × R^d, respectively.

We denote by C(R^d × R) and C(R^d) the space of continuous functions on R^d × R and R, respectively, and by C^b(R^d × R) and C^b(R) the corresponding subspaces of continuous and bounded functions.

By C^1(R^d × R) and C^1(R) we denote the spaces of continuously differentiable functions on R^d × R and 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 < ∞, we denote by L^p(0, T; \mathcal{X}) the space of Bochner p-integrable functions u : (0, T) → \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 \( a_{ij}, b_i : [0, \infty) \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}, 1 \leq i, j \leq d \), be measurable.

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

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

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

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

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 [28]) for linear FPEs applied to the (linear) Kolmogorov operator

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

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

More precisely, by Theorem 2.5 in [28], there exists a probability measure \( P \) on \( C([0, T]; \mathbb{R}^d) \) equipped with its Borel \( \sigma \)-algebra and its natural normal filtration obtained by the evaluation maps \( \pi_t, t \in [0, T] \), defined by

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

solving the martingale problem (see [28], Definition 2.4) for the time-dependent (linear) Kolmogorov operator \( \frac{d}{dt} + L_{\mu_t} \) (with \((\mu_t)_{t \geq 0}\) as above fixed) with time marginal laws

\[
P \circ \pi_t^{-1} = \mu_t, \ t \geq 0.
\]

Then, a standard result (see [26]) 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

\[
dX(t) = b(t, X(t), \mu_t)dt + \sigma(t, X(t), \mu_t)dW(t),
\]
with the law
\[ Q \circ X^{-1} = P, \]
where \( \sigma = \left( (a_{ij})_{1 \leq i,j \leq d} \right)^{1/2} \). In particular, we have, for the marginals,
(2.3) \[ \mathcal{L}_{X(t)} := Q \circ X(t)^{-1} = \mu_t, \quad t \geq 0. \]

**Remark 2.2** Because of (2.3), the process \( X(t), \quad t \geq 0, \) 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 \( \mu_0 \), which would, of course, be very desirable. For this, one has to prove the uniqueness of the solutions to the *linear* Fokker-Planck equation
\[ \partial_t \nu_t = L_{\mu_t}^* \nu_t, \quad \nu_0 = \mu_0 \]
(see (1.4)). However, this was achieved in certain cases where \( d = 1 \) (see [5], [8], [24]).

**Conclusion.** To weakly solve DDSDE (1.1), we have to solve the corresponding 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 \( u(t, x)dx, \quad t \geq 0, \) i.e., the nonlinear Fokker-Planck equation

\[ \frac{\partial u}{\partial t} - \sum_{i,j=1}^{d} D^2_{ij}(a_{ij}(x, u)u) + \text{div}(b(x, u)u) = 0 \quad \text{in} \quad \mathcal{D}'((0, \infty) \times \mathbb{R}^d), \]

\[ u(0, x) = u_0(x), \quad x \in \mathbb{R}^d, \]

where \( b(x, u) = \{b_i(x, u)\}_{i=1}^{d} \) and \( D^2_{ij} := \frac{\partial^2}{\partial x_i \partial x_j}. \)

We shall study this equation under two different sets of hypotheses specified in the following.
(H1) \( a_{ij} \in C^1(\mathbb{R}^d \times \mathbb{R}), u(a_{ij})_u, (a_{ij})_x \in C_b(\mathbb{R}^d \times \mathbb{R}), \forall a_{ij} = a_{ji}, i, j = 1, ..., d. \) Moreover, the functions \((a_{ij}(x, u))_u\) and \((a_{ij}(x, u))_x\) are Lipschitz in \( u \) uniformly with respect to \( x \).

(H2) \[
\sum_{i,j=1}^d (a_{ij}(x, u) + (a_{ij}(x, u))_u) \xi_i \xi_j \geq \gamma |\xi|^2, \quad x \in \mathbb{R}^d, \ u \in \mathbb{R},
\]
where \( \gamma > 0 \).

(H3) \( b_i, ub_i \in C_b(\mathbb{R}^d \times \mathbb{R}), (b_i)_u \in L^\infty(\mathbb{R}^d \times \mathbb{R}), b_i(x, 0) \equiv 0, i = 1, 2, ..., d. \)

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

(H2)^' \[
\sum_{i,j=1}^d (a_{ij}(u) + (a_{ij}(u))_u) \xi_i \xi_j \geq 0, \quad \forall \xi \in \mathbb{R}^d, \ u \in \mathbb{R}.
\]

(H3)^' \( b_i, ub_i \in C_b(\mathbb{R}), i = 1, ..., d. \)

Here \((a_{ij}(x, u))_u = \frac{\partial}{\partial u} a_{ij}(x, u), \forall u \in \mathbb{R},\) and \((a_{ij}(x, u))_x = \nabla_x a_{ij}(x, u).\) 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

\[
\frac{\partial v}{\partial t} = F(x, u)v = \sum_{i,j=1}^d D^2_{ij}(a_{ij}(x, u)v) - \text{div}(b(x, u)u)
\]

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

\[
\begin{align*}
\frac{du}{dt}(t) + Au(t) &= 0, \ t \geq 0, \\
\quad u(0) &= u_0,
\end{align*}
\]

7
in a Banach space $\mathcal{X}$.

The 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

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

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

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

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

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

(3.6) \[ u_h^i + hAu_h^i = u_h^{i-1}, i = 1, ..., N; u_0^0 = u_0. \]

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 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).

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

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

where $\beta : \mathbb{R} \to [0, \infty)$ is a maximal monotone (multivalued) function with $\sup\{|s| : s \in \beta(r)\} \leq C|r|^m$, $r \in \mathbb{R}$, 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], [8]. However, the present case is much more difficult and the arguments of [4] are not applicable here. We note that, for $b \equiv 0$, (3.7) is just the generalized porous media/fast diffusion equation.
3.1 Existence for FPEs in the degenerate, \(x\)-dependent case

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

\[
(3.8) \quad Au = -\sum_{i,j=1}^{d} D^2_{ij}(a_{ij}(x,u))u + \text{div}(b(x,u)u), \quad \forall \in D(A),
\]

\[
(3.9) \quad D(A) = \left\{ u \in L^1; -\sum_{i,j=1}^{d} D^2_{ij}(a_{ij}(x,u))u + \text{div}(b(x,u)u) \in L^1 \right\},
\]

where \(D^2_{ij} = \frac{\partial^2}{\partial x_i \partial x_j}\) and \(\text{div}\) are taken in sense of Schwartz distributions on \(\mathbb{R}^d\), i.e., in \(D'(\mathbb{R}^d)\).

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, (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

\[
(3.10) \quad u - \lambda \sum_{i,j=1}^{d} D^2_{ij}(a_{ij}(x,u))u + \lambda \text{div}(b(x,u)u) = f \quad \text{in} \quad D'(\mathbb{R}^d)
\]

has a unique solution \(u = u(\lambda, f) \in L^1\).

Moreover, if \(f \geq 0\) a.e. in \(\mathbb{R}^d\), then \(u \geq 0\) a.e. in \(\mathbb{R}^d\) and

\[
(3.11) \quad \int_{\mathbb{R}^d} u(x)dx = \int_{\mathbb{R}^d} f(x)dx.
\]

Finally, we have, for all \(\lambda > 0\),

\[
(3.12) \quad |u(\lambda, f_1) - u(\lambda, f_2)| \leq |f_1 - f_2|_1, \quad \forall f_1, f_2 \in L^1,
\]

\[
(3.13) \quad (I + \lambda A)^{-1}f \geq 0, \quad \text{a.e. in} \ \mathbb{R}^d, \quad \text{if} \ f \in L^1, \ f \geq 0, \quad \text{a.e. in} \ \mathbb{R}^d,
\]

\[
(3.14) \quad \int_{\mathbb{R}^d} (I + \lambda A)^{-1}f(x)dx = \int_{\mathbb{R}^d} f(x)dx, \quad \forall f \in L^1, \ \lambda > 0.
\]

**Proof.** In the following, we shall simply write

\[a_{ij}(u) = a_{ij}(x,u), \ x \in \mathbb{R}^d, \ u \in \mathbb{R}.\]
We set

\[ a^*_{ij}(u) \equiv a_{ij}(x, u)u, \quad \forall i, j = 1, \ldots, d, \]

and note that, by hypotheses (H1), (H2), we also have

\[ a^*_{ij}, (a^*_{ij})_u \in C(\mathbb{R}^d \times \mathbb{R}), \quad (a^*_{ij})_u \in C_b(\mathbb{R}^d \times \mathbb{R}), \quad i, j = 1, \ldots, d, \]

\[ (3.15) \quad \sum_{i,j=1}^{d} (a^*_{ij}(x, u))u \xi_i \xi_j \geq \gamma |\xi|^2, \quad \forall \xi \in \mathbb{R}^d, \ x \in \mathbb{R}^d. \]

Then we rewrite (3.10) as

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

Equivalently,

\[ (3.10)'' \quad u - \lambda \sum_{i,j=1}^{d} D_i(a^*_{ij}(u))_u D_j u + (a^*_{ij}(x, u))_{x_j} + \lambda \text{div}(b(x, u)u) = f, \]

where \((a^*_{ij}(x, u))_{x_j} = \frac{\partial}{\partial x_j} a^*_{ij}(x, u)\).

We also set

\[ b_\infty = \sup \{|b_i(x, u)|; \quad (x, u) \in \mathbb{R}^d \times \mathbb{R}, \ i = 1, \ldots, d\}, \]

\[ c_\infty = \sup \{|(a_{ij}(x, u))_{x_j}|; \quad (x, u) \in \mathbb{R}^d \times \mathbb{R}, \ i, j = 1, \ldots, d\}. \]

(The latter formulation of (3.10)' makes sense only if \(D_j u \in L^1_{\text{loc}}\).)

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

**Lemma 3.2** Let \(f \in L^2\) and \(0 < \lambda \leq \lambda_0 = \gamma (b^2_\infty + c^2_\infty)^{-1}\). Then, for each \(N\) there is at least one solution \(u_N \in H^1_0(B_N)\) to the equation

\[ u - \lambda \sum_{i,j=1}^{d} D^2_{ij}(a^*(u)) + \lambda \text{div}(b(x, u)u) = f \text{ in } B_N, \]

\[ u = 0 \text{ on } \partial B_N, \]

which satisfies the estimate

\[ \|u_N\|_{L^2(B_N)} + \lambda \|\nabla u_N\|_{L^2(B_N)} \leq C(\|f\|_{L^2(B_N)}^2 + 1), \]

where \(C\) is independent of \(N\) and \(\lambda\).
Proof. For $\rho > 0$, we set $\mathcal{M}_\rho = \{ v \in L^2(B_N); \| v \|_{L^2(B_N)} \leq \rho \}$ and consider the operator $F : \mathcal{M}_\rho \to L^2(B_N)$ defined by $F(v) = u \in H^1_0(B_N)$, where $u$ is the solution to the linear elliptic problem

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

in $B_N$, $u = 0$ on $\partial B_N$.

By (3.15), (3.16) and (H2), it follows via the Lax-Millgram lemma that, for each $v \in \mathcal{M}_\rho$ and $\lambda \in (0, \lambda_0)$, problem (3.19) has a unique solution $u = F(v)$ and that $F$ is continuous in $L^2(B_N)$. Moreover, by (3.19) and (H1), we see that

$$\|u\|_{L^2(B_N)}^2 + \gamma \lambda \|\nabla u\|_{L^2(B_N)}^2 \leq \lambda b_\infty \|\nabla u\|_{L^2(B_N)}^2 + \gamma \lambda \|\nabla u\|_{L^2(B_N)}^2 + c_\infty \lambda \rho \|u\|_{L^2(B_N)}^2 + c_\infty \lambda \rho \|\nabla u\|_{L^2(B_N)}^2.$$

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

Moreover, since the Sobolev space $H^1(B_N)$ is compactly embedded in $L^2(B_N)$, by (3.20) 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 (3.17). Also, by (3.18), it follows that estimate (3.18) holds.

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

$$|u|^2_2 + \gamma \lambda \|\nabla u\|^2_2 \leq C(|f|^2_2 + 1).$$

Proof. Consider a sequence $\{N\} \to \infty$ and $u_N \in H^1_0(B_N)$ a solution to (3.17) given by Lemma 3.2. By (3.18), we have

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

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

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

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

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

\[
\int_{\mathbb{R}^d} u_N \psi \, dx + \lambda \sum_{i,j=1}^d \int_{\mathbb{R}^d} a_{ij}^*(u_N) D_j u_N D_i \psi \, dx \\
- \lambda \sum_{i=1}^d \int_{\mathbb{R}^d} b(x, u_N) u_N \cdot \nabla \psi \, dx = 0, \quad \forall \psi \in H^1_0(B_N),
\]

we infer by (H1), (H3) and (3.22) that \( u \in H^1(\mathbb{R}^d) \) is a solution to (3.10). Also, estimate (3.21) follows by (3.18).

This completes the proof of Lemma 3.3.

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

\[
|u(\lambda, f_1) - u(\lambda, f_2)|_1 \leq |f_1 - f_2|_1, \quad \forall f_1, f_2 \in L^2.
\]

Here is the proof. 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

\[
\begin{align*}
&u - \lambda \sum_{i,j=1}^d D^2_{ij}(a_{ij}^*(x, u_1) - a_{ij}^*(x, u_2)) \\
&\quad + \lambda \text{div}(b^*(x, u_1) - b^*(x, u_2)) = f \text{ in } \mathcal{D}'(\mathbb{R}^d),
\end{align*}
\]

where \( u_i = u(\lambda, f_i) \), \( b^*(x, u_i) = b(x, u_i) u_i, \ i = 1, 2 \).

More precisely, since \( u_i \in H^1(\mathbb{R}^d) \), equation (3.24) is taken in its weak form

\[
\int_{\mathbb{R}^d} (u \psi + \lambda \sum_{i,j=1}^d D_i(a_{ij}^*(x, u_1) - a_{ij}^*(x, u_2))D_j \psi \, dx \\
- \lambda \sum_{i=1}^d (b^*(x, u_1) - b^*(x, u_2)) \cdot \nabla \psi \, dx = \int_{\mathbb{R}^d} f \psi \, dx, \quad \forall \psi \in H^1(\mathbb{R}^d),
\]

In order to fix the idea of the proof, we invoke first an heuristic argument. Namely, if multiply (3.24) by \( \eta \in L^\infty(\mathcal{O}) \), \( \eta(x) \in \text{sign}(u(x)) \), a.e. \( x \in \mathbb{R}^d \),
and take into account that, by the monotonicity of functions $a^*_ij$, 
\[ \eta(x) \in \text{sign}(a^*_ij(x, u_1(x)) - a^*_ij(x, u_2(x))), \text{ a.e. } x \in \mathbb{R}^d, \]
we get by (3.15)
\[ |u|_1 + \lambda \int_{\mathbb{R}^d} \sum_{i,j=1}^d D_j(a^*_ij(x, u_1(x)) - a^*_ij(x, u_2(x)))D_j\eta(x)dx \]
\[ + \lambda \int_{\mathbb{R}^d} \text{div}(b^*(x, u_1) - b^*(x, u_2))\eta dx = \int_{\mathbb{R}^d} f\eta dx. \]

Taking into account that, by monotonicity of $u \rightarrow a^*_ij(x, u)$, we have (formally)
\[ D_j(a^*_ij(x, u_1(x)) - a^*_ij(x, u_2(x)))D_j\eta(x) \geq 0 \text{ in } \mathbb{R}^d \]
and
\[ \int_{\mathbb{R}^d} \text{div}(b^*(x, u_1) - b^*(x, u_2))\eta dx = \int_{\|u\| = 0} (b^*(x, u_1) - b^*(x, u_2)) \cdot \nabla(u)dx = 0, \]
we get (3.23). This formal argument can be made rigorous by using a smooth approximation $X_\delta$ of signum graph. Namely, let $X_\delta \in \text{Lip}(\mathbb{R})$ be the function
\[ X_\delta(r) = \begin{cases} 
1 & \text{for } r \geq \delta, \\
\frac{r}{\delta} & \text{for } |r| < \delta, \\
-1 & \text{for } r < -\delta,
\end{cases} \]
where $\delta > 0$ and let $\varphi \in C_0^\infty(\mathbb{R}^d)$ be the cut-off function
\[ \varphi(x) = \varphi_\delta(x) = \eta(\delta|x|), \text{ } x \in \mathbb{R}^d, \delta > 0, \]
where $\eta \in C_0^\infty(\mathbb{R}^+)$, is such that $\eta(r) = 1$ for $x \in [0, 1]$, $\eta(r) = 0$ for $r > 2$ and $0 \leq \eta \leq 1$ on $\mathbb{R}^+$. Then
\[ |\nabla \varphi_\delta(x)| \leq \delta |\eta'(\delta|x|)|, \forall x \in \mathbb{R}^\delta, \delta > 0. \]
If multiply (3.24) by $\varphi \mathcal{X}_\delta(u)$ and integrate on $\mathbb{R}^d$, we obtain (we omit $x$ in $a^*_ij(x,u)$)

\begin{align}
\int_{|u(x)| \geq \delta} \varphi(x)|u(x)|\,dx + \frac{1}{\delta} \int_{|u(x)| \leq \delta} \varphi(x)|u(x)|\,dx \\
+ \lambda \sum_{i,j=1}^d \int_{\mathbb{R}^d} D_j(a^*_{ij}(u_1) - a^*_{ij}(u_2)) D_i(\varphi \mathcal{X}_\delta(u))\,dx
\end{align}

(3.27)

$$= \lambda \int_{\mathbb{R}^d} (b^*(x,u_1) - b^*(x,u_2)) \cdot \nabla(\varphi \mathcal{X}_\delta(u))\,dx + \int_{\mathbb{R}^d} \varphi \mathcal{X}_\delta(u)\,dx.$$ 

We set

$$I^1_\delta = \int_{\mathbb{R}^d} (b^*(x,u_1) - b^*(x,u_2)) \cdot \nabla(\varphi \mathcal{X}_\delta(u))\,dx$$

$$= \int_{\mathbb{R}^d} (b^*(x,u_1) - b^*(x,u_2)) \cdot \nabla u \mathcal{X}_\delta'(u)\varphi\,dx + I^2_\delta$$

$$= \frac{1}{\delta} \int_{|u| \leq \delta} \varphi(b^*(x,u_1) - b^*(x,u_2)) \cdot \nabla u \,dx + I^2_\delta,$$

where

$$I^2_\delta = \int_{\mathbb{R}^d} (b^*(x,u_1) - b^*(x,u_2)) \cdot \nabla \varphi \mathcal{X}_\delta(u)\,dx.$$ 

Since, by (H3), $|b^*(x,u_1) - b^*(x,u_2)| \leq C|u|(|u_1| + |u_2|)$ and $u_i \in L^2$, it follows that

$$\lim_{\delta \to 0} \frac{1}{\delta} \int_{|u| \leq \delta} \varphi(b^*(x,u_1) - b^*(x,u_2)) \cdot \nabla u \,dx \leq \lim_{\delta \to 0} \left( \int_{|u| \leq \delta} \varphi |\nabla u|^2\,dx \right)^{\frac{1}{2}}.$$ 

(See (3.30) below.) Moreover, since by (H3) the function $u \to b(x,u)$ is Lipschitz (uniformly) in $x$ and $b(x,0) \equiv 0$, we infer that $b^*(x,u_i) \in L^1$, $i = 1, 2$, and so, by (3.26), $\lim_{\delta \to 0} I^2_\delta = 0$. This yields

(3.28) $$\lim_{\delta \to 0} I^1_\delta = 0.$$ 

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
\[ I^3_\delta = \int_{E_\delta} \sum_{i,j=1}^d D_j(a_{ij}^*(u_1) - a_{ij}^*(u_2))D_i(\varphi \chi_\delta(u))dx \]

\[ = \frac{1}{\delta} \int_{E_\delta} \varphi \sum_{i,j=1}^d ((a_{ij}^*(x, u_1))_u D_j u_1 - (a_{ij}^*(u_2))_u D_j u_2 \]

\[ + (a_{ij}^*(x, u_1))_{x_j} - (a_{ij}^*(x, u_1))_{x_j}) D_i u dx \]

\[ + \int_{\mathbb{R}^d} \sum_{i,j=1}^d D_j(a_{ij}^*(u_1) - a_{ij}^*(u_2))D_i \varphi \chi_\delta(u)dx \]

\( (3.29) \)

\[ = \frac{1}{\delta} \int_{E_\delta} \varphi \sum_{i,j=1}^d (a_{ij}^*(x, u_1))_u D_j u D_i u dx \]

\[ + \frac{1}{\delta} \int_{E_\delta} \varphi \sum_{i,j=1}^d ((a_{ij}^*(x, u_1))_u - (a_{ij}^*(x, u_2))_u) D_j u_2 D_i u dx \]

\[ + \frac{1}{\delta} \int_{E_\delta} \varphi \sum_{i,j=1}^d ((a_{ij}^*(x, u_1))_{x_j} - (a_{ij}^*(x, u_2))_{x_j}) D_i u dx \]

\[ + \int_{\mathbb{R}^d} \sum_{i,j=1}^d D_j(a_{ij}^*(u_1) - a_{ij}^*(u_2))D_i \varphi \chi_\delta(u)dx \]

\[ = K_1^\delta + K_2^\delta + K_3^\delta + K_4^\delta. \]

Here, \( E_\delta = \{ x \in \mathbb{R}^d; |u(x)| \leq \delta \}. \) By (H2), it follows that \( K_1^\delta \geq 0. \) By (H1), we have

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

Taking into account that \( u_i \in H^1(\mathbb{R}^d), \ i = 1, 2, \) and, for each \( v \in H^1(\mathbb{R}^d) \) and \( \varphi \in C_0^\infty(\mathbb{R}^d), \varphi \geq 0, \)

\( (3.30) \)

\[ \lim_{\delta \to 0} \int_{|v| \leq \delta} \varphi(x)|\nabla v(x)|^2 dx = 0, \]

we infer that \( \lim_{\delta \to 0} K_i^\delta = 0, \) for \( i = 2, 3, \) and so it follows by (3.29) that

\[ \lim_{\delta \to 0} \inf I^3_\delta \geq \lim_{\delta \to 0} \inf K_4^\delta. \]
On the other hand, since \( u_i \in H^1(\mathbb{R}^d) \), by (H3) we see that
\[
D_j(a_{ij}^*(u_1)), D_j(a_{ij}^*(u_2)) \in L^1.
\]
Then, by (3.26), we see that
\[
\lim_{\delta \to 0} K_4^\delta = 0.
\]
Together with (3.27) and (3.28), the latter implies that
\[
\int_{\mathbb{R}^d} \phi_\delta(x)|u(x)|dx \leq \int_{\mathbb{R}^d} \phi_\delta(x)|f(x)|dx, \ \forall \delta > 0.
\]
Then, letting \( \delta \to 0 \), yields
\[
|u|_1 \leq |f|_1, \ \forall \lambda \in (0, \lambda_0), \ f \in L^1.
\]
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 (3.31), we see that
\[
|u_n - u_m|_1 \leq |f_n - f_m|_1, \ \forall n, m \in \mathbb{N}.
\]
Hence, there is \( u^* = \lim_{n \to \infty} u_n \) in \( L^1 \). Moreover, by (H1), we see that
\[
a_{ij}^*(u_n) \to a_{ij}(u^*), \ \text{a.e. in } \mathbb{R}^d
\]
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 \( i, j = 1, 2, ..., d \). We have, therefore,
\[
\sum_{i,j=1}^d D^2_{ij}a_{ij}^*(u_n) \to \sum_{i,j=1}^d D^2_{ij}a_{ij}^*(u) \ \text{strongly in } L^1.
\]
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 (3.23), it follows (3.12) for all \( \lambda \in (0, \lambda_0] \). However, arguing as in [1] (Proposition 3.1), it follows that (3.12) extends to all \( \lambda > 0 \).
As regards (3.14), it simply follows by equation (3.10) (where $f \in L^2$)
by integrating on $\mathbb{R}^d$. Then, by density, it extends to all of $f \in L^1$.
Finally, (3.13) for $f \in L^2$, $f \geq 0$, follows by multiplying (3.10) with $\text{sign}(u^-)$ (or,
more exactly, by $\mathcal{X}_\delta(u^-)$ and letting $\delta \to 0$) and integrating on $\mathbb{R}^d$.
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 weak
solution to equation (3.1), we mean
a mild solution to equation (3.2).

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

\begin{align}
(3.33) \quad |u(t, u_0^1) - u(t, u_0^2)|_1 & \leq |u_0^1 - u_0^2|_1, \quad \forall u_0^1, u_0^2 \in L^1, \ t \geq 0, \\
(3.34) \quad u & \geq 0 \text{ a.e. in } (0, \infty) \times \mathbb{R}^d \text{ if } u_0 \geq 0 \text{ a.e. in } \mathbb{R}^d, \\
(3.35) \quad \int_{\mathbb{R}^d} u(t, x) dx & = \int_{\mathbb{R}^d} u_0(x) dx, \forall u_0 \in L^1, \ t \geq 0,
\end{align}

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{align}
(3.36) \quad \int_0^\infty \int_{\mathbb{R}^d} (u(t, x) \varphi(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^\infty_0((0, \infty) \times \mathbb{R}^d).
\end{align}

Proof. As mentioned above, the existence of a mild solution $u$ for (3.4),
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). We note that the finite difference
scheme (3.4)–(3.5) implies (3.36) and it can be equivalently expressed by the
exponential formula

\begin{align}
(3.37) \quad u(t, u_0) = \lim_{n \to \infty} \left( I + \frac{t}{n} A \right)^{-n} u_0, \forall t \geq 0, \ u_0 \in \overline{D(A)} = L^1.
\end{align}

Then, by (3.12)–(3.14), we get for $u = u(t, u_0)$ the corresponding properties
(3.33)–(3.35) and this completes the proof. In particular, it follows that, if
\( u_0 \) is a density probability, then so is \( u(t, u_0) \) for all \( t \geq 0 \). Note also that \( t \to u(t, u_0) \) is a continuous semigroup of nonexpansive operators in the space \( L^1 \).

### 3.2 Existence for degenerate FPEs

We consider here the equation

\[
  u_t - \sum_{i,j=1}^{d} D_{ij}^2(a_{ij}(u)u) + \sum_{i=1}^{d} D_i(b_i(u)u) = 0 \text{ in } \mathcal{D}'((0, \infty) \times \mathbb{R}^d),
\]

\[
  u(0, x) = u_0(x), \quad x \in \mathbb{R}^d,
\]

where \( a_{ij} \equiv a_{ij}(u) \) and \( b_i \) satisfy hypotheses \((H1)'-(H3)'\).

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

\[
  A_1u = -\sum_{i,j=1}^{d} D_{ij}^2(a_{ij}(u)u) + \sum_{i=1}^{d} D_i(b_i(u)u) \text{ in } \mathcal{D}'(\mathbb{R}^d),
\]

\[
  D(A_1) = \left\{ u \in L^1; -\sum_{i,j=1}^{d} D_{ij}^2(a_{ij}(u)u) + \sum_{i=1}^{d} D_i(b_i(u)u) \in L^1 \right\}.
\]

We have

**Lemma 3.5** Assume that \((H1)'-(H3)'\) hold. Then the operator \( A_1 \) is m-accretive in \( L^1 \).

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

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

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

\[
  |u(\lambda, f_1) - u(\lambda, f_2)|_1 \leq |f_1 - f_2|_1, \quad \forall f_1, f_2 \in L^1.
\]

We consider a smooth approximation \( b^\varepsilon_i \in C^1_b(\mathbb{R}^d) \) to \( b_i \) such that

\[
  b^\varepsilon_i, u(b^\varepsilon_i)u \in C_b(\mathbb{R}), \quad i = 1, 2, \ldots, d,
\]

\[
  \lim_{\varepsilon \to 0} b^\varepsilon_i(u) = b_i(u) \text{ uniformly in } \mathbb{R}, \quad i = 1, \ldots, d.
\]
and \( a_{ij}^\varepsilon \equiv a_{ij} \ast \rho_\varepsilon + \varepsilon \delta_{ij} \) where \( \rho_\varepsilon \) is a mollifier and \( \delta_{ij} \) is the Kronecker symbol. (We may also take \( b_i^\varepsilon \) as \( b_i \ast \rho_\varepsilon \).) Then, we approximate (3.40) by

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

Equivalently,

\[
(3.45) \quad u + \lambda A_1^\varepsilon(u) = f,
\]

where

\[
A_1^\varepsilon(u) = - \sum_{i,j=1}^d D_{ij}^2(a_{ij}(u)u) + \sum_{i=1}^d D_i(b_i^\varepsilon(u)u), \quad \forall u \in D(A_1^\varepsilon),
\]

\[
D(A_1^\varepsilon) = \left\{ u \in L^1; - \sum_{i,j=1}^d D_{ij}^2(a_{ij}(u)u) + \sum_{i=1}^d D_i(b_i^\varepsilon(u)u) \in L^1 \right\}.
\]

We shall prove that, for each \( f \in L^1 \), there is a solution \( u = u_\varepsilon(\lambda, f) \) satisfying (3.41) for \( 0 < \lambda < \lambda_0 \).

Since \( a_{ij}^\varepsilon \) and \( b_i^\varepsilon \) satisfy, for each \( \varepsilon > 0 \), hypotheses (H1)–(H3), Proposition 3.1 implies existence for (3.44) of a solution \( u_\varepsilon = u_\varepsilon(\lambda, f) \) in \( L^1(\mathbb{R}^d) \) for each \( f \in L^2 \) if \( 0 < \lambda \leq \lambda_0^\varepsilon = \frac{C}{\varepsilon} \), \( C \) independent of \( \varepsilon \).

Moreover, one has

\[
|u_\varepsilon(\lambda, f_1) - u_\varepsilon(\lambda, f_2)|_1 \leq |f_1 - f_2|_1, \quad \forall f_1, f_2 \in L^2, \quad \lambda \in (0, \lambda_0^\varepsilon).
\]

Then, by density, \( u_\varepsilon(\lambda, f) \) extends as solution to (3.44) for all \( f \in L^1 \).

Note also that, by (3.12)–(3.14), we have, for all \( \varepsilon > 0 \) and \( \lambda \in (0, \lambda_0^\varepsilon) \),

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

\[
(3.48) \quad (I + \lambda A_1^\varepsilon)^{-1} f \geq 0, \text{ a.e. in } \mathbb{R}^d \text{ if } f \geq 0, \text{ a.e. in } \mathbb{R}^d,
\]

while (3.46) yields

\[
(3.49) \quad |(I+\lambda A_1^\varepsilon)^{-1} f_1 - (I+\lambda A_1^\varepsilon)^{-1} f_2|_1 \leq |f_1 - f_2|_1, \quad \forall f_1, f_2 \in L^1, \varepsilon > 0.
\]
Though (3.47)–(3.49) were proved only for $0 < \lambda \leq \lambda_0^\varepsilon$, 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

$$(I + \lambda A_2)^{-1}f = (I + \lambda_0 A_2)^{-1}\left(\frac{\lambda_0^\varepsilon}{\lambda} f + \left(1 - \frac{\lambda_0^\varepsilon}{\lambda}\right)(I + \lambda A_2)^{-1}f\right), \lambda > \lambda_0^\varepsilon.$$  

(See [1], Proposition 3.3.)

Now, we are going to let $\varepsilon \to 0$ in (3.44). We set, for $f \in L^1$ and $u_\varepsilon$, the solution to (3.44),

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

Since $a_{ij}^\varepsilon$ and $b_i^\varepsilon$ are independent of $x$, we see that $x \to u_\varepsilon^h(x + h)$ is the solution to (3.44) for $f(x) = f(x + h)$. Then, by (3.49), it follows that

$$|u_\varepsilon^h|_1 \leq |f_\varepsilon^h|_1, \ \forall h \in \mathbb{R}^d, \ \varepsilon > 0.$$  

By the Kolmogorov compactness theorem (see, e.g., [10], p. 111), it follows that $\{u_\varepsilon\}$ is compact in $L^1_{loc}(\mathbb{R}^d)$ and so, on a subsequence,

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

Since $|u_\varepsilon|_1 \leq C, \ \forall \varepsilon > 0$, it follows via Fatou’s lemma that $u \in L^1$. Letting $\varepsilon \to 0$ in (3.44), where $u = u_\varepsilon$, and taking into account that

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

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

$$|a_{ij}^\varepsilon(u_\varepsilon) u_\varepsilon| + |b_i^\varepsilon(u_\varepsilon) u_\varepsilon| \leq C, \text{ a.e. in } \mathbb{R}^d,$$

where $C$ is independent of $\varepsilon$, we see that $u$ is a solution to (3.44) and so $u = (I + \lambda A_1)^{-1}f$. Moreover, letting $\varepsilon \to 0$ in (3.47)–(3.49), we see that

$$\begin{align*}
(I + \lambda A_1)^{-1}f_1 - (I + \lambda A_1)^{-1}f_2 &\leq |f_1 - f_2|_1, \ \forall \lambda > 0, \ f_1, f_2 \in L^1, \\
\int_{\mathbb{R}^d} (I + \lambda A_1)^{-1} f \ dx = \int_{\mathbb{R}^d} f \ dx, \ \forall f \in L^1, \ \lambda > 0, \\
(I + \lambda A_1)^{-1}f &\geq 0, \text{ a.e. in } \mathbb{R}^d \text{ if } f \geq 0, \text{ a.e. in } \mathbb{R}^d.
\end{align*}$$  

20
Then, by the Crandall and Liggett existence theorem, for each \( u_0 \in \overline{D(A_1)} = L^1 \), the differential equation

\[
\begin{align*}
\frac{du}{dt} + A_1 u &= 0, \quad t > 0, \\
u(0) &= u_0,
\end{align*}
\]

has a unique mild solution \( u \in C([0, \infty); L^1) \) in sense of (3.41)–(3.43).

As in the previous case, this mild solution is by definition the weak solution to the Fokker-Planck equation (3.38).

We have, therefore, the following existence result.

**Theorem 3.6** 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 (3.38). Moreover, this solution satisfies (3.33)–(3.35) and is a solution to (3.38) in sense of Schwartz distributions on \((0, \infty) \times \mathbb{R}^d\), i.e., in the sense of (3.36) or (1.2).

**Remark 3.7** In particular, Theorems 3.4 and 3.6 implies the existence of a solution \( u \) in sense of distributions on \((0, \infty) \times \mathbb{R}^d\) for equation (3.1). Moreover, \( u : [0, \infty) \rightarrow L^1 \) is continuous. In some special cases, these two properties are sufficient to characterize the weak solution to (3.1). In fact, this is the case if (see [11]) \( b \equiv 0 \) and

\[
a_{ij}(x, u)u = \delta_{ij} \beta(u)u, \quad \forall u \in \mathbb{R}, \quad i, j = 1, \ldots, d,
\]

where \( \beta \) 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 open for the Fokker-Planck equation (3.1).

**Remark 3.8** An important case, which was not treated here, is that where \( a_{ij} \equiv a_{ij}(t, x, u) \). In this case, under hypotheses (H1)–(H3) the operator \( A(t) \) defined as in (3.8)–(3.9) is, of course, \( m \)-accretive in \( L^1 \), but for the existence of a mild solution \( u \) on \([0, T]\) to the corresponding equation (3.8) a condition of the form

\[
|(I + \lambda A(t))^{-1}u_0 - (I + \lambda A(s))^{-1}u_0|_1 \leq \lambda|t - s|C_T(u_0|_1)(1 + |A(t)u_0|_1),
\]

\[
\forall u_0 \in D(A(t)) \equiv D(A(0)), \quad s, t \in [0, T],
\]

where \( \lambda > 0 \) and \( C_T \) is a continuous function, is needed (see [12]).

However, at this time it is not clear if such a condition holds for sufficiently smooth functions \( t \rightarrow a_{ij}(t, \cdot) \).
Remark 3.9 In the special case \( a_{ij} = \delta_{ij} \), the weak solution \( u \) given by Theorem 3.6 is an entropic solution in sense of S. Kruzkov for equation (3.1). In the present case, the solution \( u \) given by Theorem 3.6 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.

4 Solution of the DDSDE

Consider the following DDSDE for \( T \in (0, \infty) \)
\[
dX(t) = b \left( X(t), \frac{dL}{dx}(X(t)) \right) dt + \sqrt{2} \sigma \left( X(t), \frac{dL}{dx}(X(t)) \right) dW(t), \ 0 \leq t \leq T, 
\]
\( X(0) = \xi_0, \)
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 \( a_{ij} := 2(\sigma^T)_{ij}, \ 1 \leq i, j \leq d. \) Then, as an immediate consequence of Section 2 and Theorems 3.4 and 3.6, respectively, we obtain the following.

Theorem 4.1 Suppose that \( a_{ij}, b_i, \ 1 \leq i, j \leq d, \) 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 3.6, respectively, with \( u(0, \cdot) = u_0, \) we have the "probabilistic representation"
\[
u(t, x)dx = P \circ X(t)^{-1}(dx), \ t \geq 0.
\]

Remark 4.2

(i) In the case where in (4.1) we have \( a_{ij}(x, u) = \delta_{ij} \beta(u), \ 1 \leq i, j \leq d, \) and \( \beta : \mathbb{R} \to 2^\mathbb{R} \) is maximal monotone with \( \sup \{|s| : s \in \beta(r)\} \leq C|r|^m, \ r \in \mathbb{R}, \) 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 [8] if \( \beta(r)/r \) is nondegenerate at \( r = 0 \) and in [9] including the degenerate case.

(ii) The special case \( d = 1, b \equiv 0, \) \( a_{ij}(x, u) = \delta_{ij}\beta(u), \) \( 1 \leq i, j \leq d, \) with \( \beta(r) := r|r|^{m-1}, r \in \mathbb{R}, \) for some \( m \in (1, \infty) \), was proved in [7].

(iii) [6] contains an analogous result as in [8], [9] 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 \( a_{ij}(x, u) = \delta_{ij}\beta(u), \) \( 1 \leq i, j \leq d, \) 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 [11]. Hence, obviously, we have uniqueness of the time marginal 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**


