Uniqueness for nonlinear Fokker–Planck equations and for McKean–Vlasov SDEs: The degenerate case

Viorel Barbu* Michael Röckner†

Abstract

This work is concerned with the existence and uniqueness of generalized (mild or distributional) solutions to (possibly degenerate) Fokker–Planck equations (possibly degenerate) $\rho_t - \Delta\beta(\rho) + \operatorname{div}(Db(\rho)\rho) = 0$ in $(0,\infty) \times \mathbb{R}^d$, $\rho(0,x) \equiv \rho_0(x)$. Under suitable assumptions on $\beta : \mathbb{R} \to \mathbb{R}$, $b : \mathbb{R} \to \mathbb{R}$ and $D : \mathbb{R}^d \to \mathbb{R}^d$, $d \geq 1$, this equation generates a unique flow $\rho(t) = S(t)\rho_0 : [0,\infty) \to L^1(\mathbb{R}^d)$ as a mild solution in the sense of nonlinear semigroup theory. This flow is also unique in the class of $L^\infty((0,T) \times \mathbb{R}^d) \cap L^1((0,T) \times \mathbb{R}^d)$, $\forall T > 0$, Schwartz distributional solutions on $(0,\infty) \times \mathbb{R}^d$. Moreover, for $\rho_0 \in L^1(\mathbb{R}^d) \cap H^{-1}(\mathbb{R}^d)$, $t \to S(t)\rho_0$ is differentiable from the right on $[0,\infty)$ in $H^{-1}(\mathbb{R}^d)$-norm. As a main application, the weak uniqueness of the corresponding McKean–Vlasov SDEs is proven.

MSC: 60H15, 47H05, 47J05.
Keywords: Fokker–Planck equation, McKean–Vlasov equation, mild solution, nonlinear semigroups, accretive.

1 Introduction

We shall treat here the nonlinear Fokker–Planck equation (NFPE)

*Al.I. Cuza University and Octav Mayer Institute of Mathematics of Romanian Academy, Iaşi, Romania. Email: vbarbu41@gmail.com
†Fakultät für Mathematik, Universität Bielefeld, D-33501 Bielefeld, Germany. Email: roeckner@math.uni-bielefeld.de

$$\begin{aligned}&\rho_t(t,x)-\Delta\beta(\rho(t,x))+\mathrm{div}(D(x)b(\rho(t,x))\rho(t,x))=0,\\&\qquad\qquad\qquad\qquad\qquad\qquad (t,x)\in(0,\infty)\times\mathbb{R}^d,\\&\rho(0,x)=\rho_0(x),\ \ x\in\mathbb{R}^d,\end{aligned}\tag{1.1}$$

where $\beta:\mathbb{R}\to\mathbb{R}$ is monotonically nonincreasing and $D:\mathbb{R}^d\to\mathbb{R}^d$, $b:\mathbb{R}\to\mathbb{R}$ are given functions to be made precise below in Hypotheses (i)–(iv).

The Cauchy problem (1.1) with the conditions

$$\rho(t,x)\ge 0,\ \ \forall\, t\in[0,\infty) \text{ and a.e. } x\in\mathbb{R}^d,\tag{1.2}$$

$$\int_{\mathbb{R}^d}\rho(t,x)dx=\int_{\mathbb{R}^d}\rho(0,x)dx=1,\ \ \forall\, t\ge 0,\tag{1.3}$$

is relevant in statistical mechanics (see, e.g., [15], [26]), mean field game theory ([21], [22]), as well as in stochastic analysis, where it is used to reproduce the microscopic dynamics of the solution $X(t)$ to the McKean–Vlasov stochastic differential equation

$$\begin{aligned}&dX(t)=D(X(t))b(\rho(t,X(t))dt+\sqrt{\frac{2\beta(\rho(t,X(t)))}{\rho(t,X(t))}}\,dW(t),\\&X(0)=X_0,\end{aligned}\tag{1.4}$$

by the macroscopic dynamics of its time marginal law.

In fact, if $\rho:[0,\infty)\to L^1(\mathbb{R}^d)$ is a distributional solution to (1.1), then (1.4) has a probabilistically weak solution X on some probability space $(\Omega,\mathcal{F},\mathbb{P})$ with normal filtration $(\mathcal{F}_t)_{t\ge0}$ and $(\mathcal{F}_t)$-Brownian motion W with values in $\mathbb{R}^d$ such that $\rho(t,x)dx=\mathbb{P}\circ(X(t))^{-1}(dx)$, $\rho_0(x)dx=\mathbb{P}\circ(X_0)^{-1}(dx)$. (See [2]–[6].)

Our hypotheses on β,b and D are the following:

(i) $\beta\in C^2(\mathbb{R}),\ \beta'(r)>0,\ \forall\, r\ne 0,\ \beta(0)=0$ and

$$|\beta(r)|\le\alpha_1|r|,\ \forall\, r\in\mathbb{R},\tag{1.5}$$

where $\alpha_1>0$.

(ii) $D\in L^\infty(\mathbb{R}^d;\mathbb{R}^d),\mathrm{div}\,D\in L^m_{\mathrm{loc}}(\mathbb{R}^d)$, $(\mathrm{div}\,D)^-\in L^\infty$, where $m>\frac{d}{2}$ if $d\ge 2$, $m=1$ if $d=1$.

(iii) $b\in C^1(\mathbb{R})\cap C_b(\mathbb{R})$, $b(r)\ge 0,\ \forall r\in\mathbb{R}$.

(iv) $|b^*(r)-b^*(\bar r)|\le\alpha_2|\beta(r)-\beta(\bar r)|,\ \forall\, r,\bar r\in\mathbb{R}$, where $b^*(r)\equiv b(r)r$ and $\alpha_2>0$.

A typical example is $\beta(r) \equiv \operatorname{sign} r \log(1+|r|)$ (Bose–Einstein statistic).

In general, for $\rho_0 \in L^1$, NFPE (1.1) has not a classical (strong) solution and the best one can expect is a generalized solution in the sense of the next definition.

Definition 1.1. A function $\rho : [0,\infty) \to L^1(\mathbb{R}^d)$ is said to be a *mild solution* to (1.1) if $\rho \in C([0,\infty); L^1(\mathbb{R}^d))$ and we have

$$\rho(t) = \lim_{h\to 0} \rho_h(t) \text{ in } L^1(\mathbb{R}^d),\ \forall\, t \in [0,\infty), \tag{1.6}$$

uniformly on compacts in $[0,\infty)$, where $\rho_h : [0,T] \to L^1(\mathbb{R}^d)$ is the step function,

$$\rho_h(t) = \rho_h^j,\ \forall\, t \in [jh,(j+1)h),\ j = 0,1,...,N = \left[\tfrac{T}{h}\right], \tag{1.7}$$

$$\rho_h^{j+1} - h\Delta\beta(\rho_h^{j+1}) + h\,\operatorname{div}(Db(\rho_h^{j+1})\rho_h^{j+1}) = \rho_h^j \text{ in } \mathcal{D}'(\mathbb{R}^d), \tag{1.8}$$

$$\rho_h^j \in L^1(\mathbb{R}^d),\ \forall\, j = 0,...,N;\ \rho_h^0 = \rho_0. \tag{1.9}$$

We note that (1.7)–(1.9) can be equivalently written as

$$\begin{aligned} &\frac{1}{h}\,(\rho_h(t)) - \rho_h(t-h) - \Delta\beta(\rho_h(t)) + \operatorname{div}(Db(\rho_h(t))\rho_h(t)) = 0,\\ &\hspace{20em}\forall\, t \in [h,T),\\ &\rho_h(t) = \rho_0,\ \forall\, t \in [0,h), \end{aligned} \tag{1.10}$$

where (1.10) is meant in the sense of Schwartz distributions. If we denote by $A_0 : L^1(\mathbb{R}^d) \to L^1(\mathbb{R}^d)$ the operator

$$\begin{aligned} A_0(u) &= -\Delta\beta(u) + \operatorname{div}(Db(u)u),\ \forall\, u \in D(A_0),\\ D(A_0) &= \left\{u \in L^1(\mathbb{R}^d); -\Delta\beta(u) + \operatorname{div}(Db(u)u) \in L^1(\mathbb{R}^d)\right\}, \end{aligned} \tag{1.11}$$

where Δ and div are taken in the sense of Schwartz distributions on $\mathbb{R}^d$, then the system (1.8)–(1.9) is equivalent to

$$\begin{aligned} &\rho_h^{j+1} + hA_0(\rho_h^{j+1}) = \rho_h^j,\ j = 0,1,...,N,\\ &\rho_h^0 = \rho_0. \end{aligned} \tag{1.12}$$

In other terms, this means (see, e.g., [1], p. 129) that ρ is a *mild solution* to the Cauchy problem

$$\begin{aligned} &\frac{d\rho}{dt} + A_0(\rho) = 0,\ \forall\, t \geq 0,\\ &\rho(0) = \rho_0. \end{aligned} \tag{1.13}$$

By the general existence theory for the nonlinear Cauchy problem in Banach spaces (the Crandall & Liggett existence theorem), for each $\rho_0 \in \overline{D(A_0)}$, problem (1.13) has a unique mild solution $\rho \in C([0,\infty); L^1(\mathbb{R}^d))$, if A_0 is m-accretive in $L^1(\mathbb{R}^d)$, that is,

$$R(I + \lambda A_0) = L^1(\mathbb{R}^d),\ \forall\, \lambda > 0, \tag{1.14}$$

$$\|(I + \lambda A_0)^{-1} u - (I + \lambda A_0)^{-1} v\|_{L^1(\mathbb{R}^d)} \le \|u - v\|_{L^1(\mathbb{R}^d)}, \quad \forall u, v \in L^1(\mathbb{R}^d). \tag{1.15}$$

Moreover, in this case (see, e.g., [1], p. 154), the solution $\rho = \rho(t, \rho_0) \equiv S(t)\rho_0$ defines a semigroup of contractions on $\overline{D(A_0)}$, that is, $\rho(t, \rho_0) \in \overline{D(A_0)}$, $\forall\, t \ge 0$, and

$$\rho(t + s, \rho_0) = \rho(t, \rho(s, \rho_0)),\ \forall\, t, s, \ge 0,\ \rho_0 \in \overline{D(A_0)}, \tag{1.16}$$

$$\|\rho(t, \rho_0) - \rho(t, \bar{\rho}_0)\|_{L^1(\mathbb{R}^d)} \le \|\rho_0 - \bar{\rho}_0\|_{L^1(\mathbb{R}^d)},\ \forall\, t \ge 0,\ \rho_0, \bar{\rho}_0 \in \overline{D(A_0)}, \tag{1.17}$$

and, as seen by (1.6)–(1.9), ρ also denoted e^{-tA_0}, is given by the exponential formula

$$\rho(t, \rho_0) = \lim_{n \to \infty} \left(I + \frac{t}{n}\, A_0 \right)^{-n} \rho_0 \quad \text{in } L^1(\mathbb{R}^d),\ \forall\, t \ge 0, \tag{1.18}$$

uniformly on compacts in $[0, \infty)$. In the works [4]–[8], under even weaker hypotheses than (i)–(iii), the range condition (1.14) and the existence of a family $\{J_\lambda\}$ of nonlinear contractions in $L^1(\mathbb{R}^d)$ with $J_\lambda(L^1(\mathbb{R}^d)) \subset D(A_0)$ was proven, which satisfy the resolvent equation

$$J_{\lambda_2}(f) = J_{\lambda_1}\left(\frac{\lambda_1}{\lambda_2}\, f + \left(1 - \frac{\lambda_1}{\lambda_2}\right) J_{\lambda_2}(f) \right),\ \forall\, \lambda_1, \lambda_2 > 0,\ f \in L^1,$$

and $J_\lambda(f) \in (I + \lambda A_0)^{-1} f$, $\forall f \in L^1(\mathbb{R}^d)$, $\lambda > 0$.

Then, the operator $A : D(A) \subset D(A_0) \subset L^1(\mathbb{R}^d) \to L^1(\mathbb{R}^d)$, defined by

$$\begin{aligned} A(u) &= A_0(u), \quad u = J_\lambda(f), \quad f \in L^1(\mathbb{R}^d), \\ D(A) &= J_\lambda(L^1(\mathbb{R}^d)), \end{aligned} \tag{1.19}$$

is independent of λ, m-accretive and

$$(I + \lambda A)^{-1}(f) = J_\lambda(f) \in (I + \lambda A_0)^{-1} f,\ \forall f \in L^1(\mathbb{R}^d),\ \lambda > 0. \tag{1.20}$$

This means that, for $\rho_0 \in \overline{D(A)}$, the mild solution ρ to the Cauchy problem

$$\begin{aligned} &\frac{d\rho}{dt} + A(\rho) = 0, \quad t \geq 0, \\ &\rho(0) = \rho_0, \end{aligned} \tag{1.21}$$

is just a mild solution to (1.1) in the sense of Definition 1.1. However, this does not imply the uniqueness of the mild solution to (1.1), because the family $\{J_\lambda\}$, and so the operator A, which is defined by $\{J_\lambda\}$, are not unique. In fact, $J_\lambda(f)$ is given by

$$J_\lambda(f) = \lim_{\varepsilon \to 0} y_\varepsilon \text{ in } L^1(\mathbb{R}^d), \tag{1.22}$$

where $y_\varepsilon \in L^1(\mathbb{R}^d) \cap H^2(\mathbb{R}^d)$ is the solution to an approximating equation of the form

$$y_\varepsilon - \lambda(\varepsilon I + \Delta)\beta(y_\varepsilon) + \lambda \operatorname{div}(Db_\varepsilon(y_\varepsilon)y_\varepsilon) = f \text{ in } \mathcal{D}'(\mathbb{R}^d), \tag{1.23}$$

where b_ε is a smooth approximation of b. (Other approximating equations of the form (1.23) could be considered as well.) However, since the limit (1.22) depends on the sequence $\{y_\varepsilon\}$ and hence might be not unique, one obtains in this way a family of mappings $\{J_\lambda\}_{\lambda>0}$ and each one defines via (1.19) an m-accretive operator A. So the uniqueness is not in the class of all mild solutions ρ to (1.21) and for uniqueness one should request a further condition to restrict the class of mild solutions. Such a situation is encountered for the conservation law equation as well, i.e., for $\beta \equiv 0$, where the corresponding mild solution is unique in the narrow class of Kruzkov entropic solutions [18].

The aim of this paper is twofold. The first is to prove that, if $1 \leq d$, then *under the additional Hypothesis* (iv) *the operator A_0 is itself m-accretive* (that is, $I + \lambda A_0$ is invertible for all $\lambda > 0$, and also (1.15) holds) and so *the Crandall & Liggett existence theorem is applicable to the Cauchy problem* (1.13) *to derive not only the existence, but also the uniqueness of a mild solution ρ to* (1.1) (Theorem 2.1). The second is to prove the uniqueness of solutions ρ to (1.1) in the class of L^1-valued distributional solutions ρ. This means $\beta(\rho) \in L^1_{\rm loc}((0,\infty) \times \mathbb{R}^d)$ and

$$\begin{aligned} \int_0^\infty \int_{\mathbb{R}^d} (\rho\varphi_t + \beta(\rho)\Delta\varphi + b(\rho)\rho(D \cdot \nabla\varphi))dt\,dx + \int_{\mathbb{R}^d} \varphi(0,x)\rho_0(dx) = 0, \\ \forall \varphi \in C_0^\infty([0,\infty) \times \mathbb{R}^d), \end{aligned} \tag{1.24}$$

where ρ_0 is a signed Radon measure on $\mathbb{R}^d$ of bounded variation. As seen later on in Theorem 2.1, under the above assumptions the *mild solution ρ is also a distributional solution to* (1.1). However, the uniqueness of distributional solutions to (1.1) is still an open problem under the general Hypotheses (i)–(iii). For the porous media equation (that is, $D \equiv 0$) such a uniqueness result was established by H. Brezis and M.G. Crandall [14] and M. Pierre [23] in the class $L^\infty((0,\infty)\times\mathbb{R}^d)$ for distributional solutions and a similar result was established in [6], [7] for NFPE (1.1) if β is strictly monotone. Here, one proves such a result if one merely assumes that β, D and b satisfy Hypotheses (j)–(jv) in Section 3 (Theorem 3.2 and Corollary 3.4). In particular, degenerate cases where $\{\beta'=0\}\neq\emptyset$ are covered and our results generalize [14] to the case $D\not\equiv 0$.

In the limit case $\beta\equiv 0$, $Db(r)\equiv\{a_i(r)\}_{i=1}^d\equiv a(r)$, equation (1.1) reduces to

$$\begin{aligned}&\rho_t+\operatorname{div}(a(\rho))=0,\ \ \mathcal{D}'((0,\infty)\times\mathbb{R}^d),\\ &\rho(0)=\rho_0,\end{aligned}\tag{1.25}$$

which has a unique Kružkov's solution $\rho(t)=S_0(t)\rho_0$ defined as

$$|\rho-k|_t+\operatorname{div}((a(\rho)-a(k))\operatorname{sign}(\rho-k))\le 0 \text{ in } \mathcal{D}'((0,\infty)\times\mathbb{R}^d),\ k\in\mathbb{Z}^+,\tag{1.26}$$

$\rho\in C([0,\infty);L^1)$, $\rho(0)=\rho_0$.

So, contrary to the situation encountered in Theorem 3.2, the uniqueness for (1.25) is not in the class of distributional solutions but in that of entropic solutions and this fact emphasizes the role of the diffusion term $\Delta\beta(u)$ in distributional uniqueness.

By virtue of the equivalence of equations (1.1) and (1.4), the above results have implications on the uniqueness of probabilistically weak solutions to the McKean-Vlasov equation (1.4). The key additional analytical result is to prove "linearized uniqueness" for (1.24) (see Theorem 4.1 and Corollary 4.2), which together with Corollary 3.4 imply weak uniqueness for SDE (1.4) (see Theorem 4.3 below).

As regards the literature on generalized (mild) solutions to nonlinear Fokker–Planck equations via the nonlinear semigroup theory, we refer to [4]–[5], [8] and the related papers [17], [19].)

The time-dependent case was treated in [22] under time-regularity hypotheses on the diffusion and drift coefficients invoking the general existence theory for the Cauchy problem in a Banach space with time-dependent accretive operators.

Notation. $L^p(\mathbb{R}^d)$, $1 \le p \le \infty$ (denoted L^p) is the space of all Lebesgue measurable and p-integrable functions on $\mathbb{R}^d$, with the standard norm $|\cdot|_p$. $(\cdot,\cdot)_2$ denotes the inner product in L^2. By $L^p_{\rm loc}$ we denote the corresponding local space. For any open set $\mathcal{O} \subset \mathbb{R}^d$ let $W^{k,p}(\mathcal{O})$, $k \ge 1$, denote the standard Sobolev space on $\mathcal{O}$ and by $W^{k,p}_{\rm loc}(\mathcal{O})$ the corresponding local space. We set $W^{1,2}(\mathcal{O}) = H^1(\mathcal{O})$, $W^{2,2}(\mathcal{O}) = H^2(\mathcal{O})$, $H^1_0(\mathcal{O}) = \{u \in H^1(\mathcal{O}),$ $u = 0$ on $\partial\mathcal{O}\}$, where $\partial\mathcal{O}$ is the boundary of $\mathcal{O}$. By $H^{-1}(\mathcal{O})$ we denote the dual space of $H^1_0(\mathcal{O})$ (of $H^1(\mathbb{R}^d)$, respectively, if $\mathcal{O} = \mathbb{R}^d$). $C^\infty_0(\mathcal{O})$ is the space of infinitely differentiable real-valued functions with compact support in $\mathcal{O}$ and $\mathcal{D}'(\mathcal{O})$ is the dual of $C^\infty_0(\mathcal{O})$, that is, the space of Schwartz distributions on $\mathcal{O}$. Denote by $C^k(\mathbb{R})$ the space of all continuously differentiable real-valued functions on $\mathbb{R}$ up to order k, by $C_b(\mathbb{R})$ the space of continuous and bounded real-valued functions on $\mathbb{R}$, and by ${\rm Lip}(\mathbb{R})$ the space of real-valued Lipschitz functions on $\mathbb{R}$ with the norm denoted by $\|\cdot\|_{\rm Lip}$. $C([0,\infty); L^1(\mathbb{R}^d))$ is the space of continuous functions $y : [0,\infty) \to L^1(\mathbb{R}^d)$. Denote also by $C^\infty_0([0,\infty) \times \mathbb{R}^d)$ the space of all $\varphi \in C^\infty([0,\infty) \times \mathbb{R}^d)$ such that *support* $\varphi \subset K$, where K is compact in $[0,\infty) \times \mathbb{R}^d$. We shall also use the following notations:

$$\beta'(r) \equiv \frac{d}{dr}\,\beta(r),\ b'(r) \equiv \frac{d}{dr}\,b(r),$$

$$y_t = \frac{\partial}{\partial t}\,y,\ \Delta y = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}\,y,\ \nabla y = \left\{\frac{\partial y}{\partial x_i}\right\}_{i=1}^d,\ {\rm div}\,u = \sum_{i=1}^d \frac{\partial u_i}{\partial x_i},\ u = \{u_i\}_{i=1}^d$$

We also denote by $\mathcal{P}$ the set of all probability densities ρ on $\mathbb{R}^d$, that is,

$$\mathcal{P} = \left\{ y \in L^1(\mathbb{R}^d);\ y \ge 0,\ \text{a.e. in } \mathbb{R}^d,\ \int_{\mathbb{R}^d} y(x)dx = 1 \right\}.$$

Denote by $\mathcal{M}(\mathbb{R}^d)$ the space of all signed Radon measures on $\mathbb{R}^d$of bounded variation. A sequence $\{\mu_n\} \subset \mathcal{M}(\mathbb{R}^d)$ is said to be converging to μ in $\sigma(\mathcal{M}(\mathbb{R}^d), C_b(\mathbb{R}^d))$ topology if

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} \psi\,d\mu_n = \int_{\mathbb{R}^d} \psi\,d\mu,\quad \forall \psi \in C_b(\mathbb{R}^d). \tag{1.27}$$

The function $t \to \mu(t) \in \mathcal{M}(\mathbb{R}^d)$ is said to be *narrowly continuous* on $[0,\infty)$ if, for every $\psi \in C_b(\mathbb{R}^d)$, the function $t \to \int_{\mathbb{R}^d} \psi\,d\mu(t)$ is continuous on on $[0,\infty)$.

2 The existence and uniqueness of a mild solution to NFPE

Theorem 2.1 is one of the main results of this work.

Theorem 2.1. *Assume that* $1 \leq d$. *Then, under Hypotheses* (i)–(iv), *for each* $\rho_0 \in L^1$ *there is a unique mild solution* $\rho = \rho(t, \rho_0)$ *to equation* (1.1) *which is also a distributional solution. Moreover,* (1.16)–(1.18) *hold and, if* $\rho_0 \in \mathcal{P}$, *then* $\rho(t) \in \mathcal{P}$, $\forall\, t \geq 0$, *that is,* (1.2)–(1.3) *hold.*

Finally, if $\rho_0 \in L^1 \cap L^\infty$, *then* $\rho \in L^\infty((0,T) \times \mathbb{R}^d)$ *for all* $T > 0$.

By definition, the continuous semigroup $t \to \rho = \rho(t, \rho_0)$ given by Theorem 2.1 is the *nonlinear Fokker–Planck flow* associated with the nonlinear diffusion $\beta = \beta(\rho)$ and the drift term $\rho \to Db(\rho)\rho$. This means that the operator A_0 defined by (1.11) is the infinitesimal generator of the continuous semigroup of contractions $S(t)\rho_0 = \rho(t, \rho_0)$, that is, of the Fokker–Planck flow $S(t)$. Theorem 2.1 can be rephrased in terms of semigroup theory as follows: *Under Hypotheses* (i)–(iv), *the operator* A_0 *generates in a weak-mild sense a continuous semigroup of contractions in* $L^1(\mathbb{R}^d)$ *which leaves invariant the set* $\mathcal{P}$ *of all probability densities.* Moreover, by Theorem 5.1 in [5] it follows that, if $d = 3$ and $\beta'(r) \geq a > 0$, then the flow $S(t)$ has a smoothing effect on initial data and extends to all $\rho_0 \in \mathcal{M}_b$.

It should be emphasized that such a generation result for the Fokker–Planck flow $S(t)$ is specific to the space $L^1(\mathbb{R}^d)$, because only in this space the operator A_0 is accretive. Moreover, the semigroup $S(t)$ is not t-differentiable in this space and so the solution ρ exists in the above mild sense only. However, for $\rho_0 \in L^1 \cap L^2$, ρ is a strong solution to (1.1) in $H^{-1}(\mathbb{R}^d)$. In fact, we have

Theorem 2.2. *Let Hypotheses* (i)–(iv) *hold and assume in addition that*

(v) $\beta(r) \geq \alpha_0 |r|^2$, $\forall r \in \mathbb{R}$, *where* $\alpha_0 > 0$.

Assume that $\rho_0 \in L^1 \cap L^2$. *Then, every mild solution* ρ *to* (1.1) *satisfies*

$$\beta(\rho) \in L^2(0,T; H^1(\mathbb{R}^d)) \cap L^\infty(0,T; L^2), \quad \forall\, T > 0. \tag{2.1}$$

Moreover, $\rho : [0,T] \to H^{-1}(\mathbb{R}^d)$ *is absolutely continuous and*

$$\frac{d\rho}{dt} \in L^2(0,T; H^{-1}(\mathbb{R}^d)), \tag{2.2}$$

$$\frac{d\rho}{dt}(t) + A_0(\rho(t)) = 0, \quad \text{a.e. } t \in (0,T),\ \forall\, T > 0. \tag{2.3}$$

Theorem 2.2 amounts to saying that, a.e. $t > 0$, the semigroup $S(t)$ maps the space $L^1 \cap L^2$ into $\{\rho_0 \in L^1 \cap L^2; \beta(\rho_0) \in H^1(\mathbb{R}^d)\}$ and it is a.e. $H^{-1}(\mathbb{R}^d)$-valued differentiable on $(0, \infty)$. More will be said about this in Section 5 (Theorem 5.2).

Proof of Theorem 2.1.

As mentioned earlier, Theorem 2.1 is implied by the following key result.

Proposition 2.3. *Let* $1 \le d$. *Then, under Hypotheses* (i)–(iv)*, the operator* A_0 *is* m*-accretive in* L^1. *Moreover, one has*

$$\overline{D(A_0)} = L^1, \tag{2.4}$$

$$(I + \lambda A_0)^{-1}(\mathcal{P}) \subset \mathcal{P}, \ \forall\, \lambda > 0. \tag{2.5}$$

(*Here,* $\overline{D(A_0)}$ *is the closure of* $\overline{D(A_0)}$ *in* L^1.)

We shall prove Proposition 2.3 following several steps and the first one is the following uniqueness result for the stationary (resolvent) equation associated with the operator A_0.

Lemma 2.4. *For all* $f \in L^1$ *and* $\lambda > 0$, *there is at most one solution* $y \in D(A_0)$ *to the equation*

$$y + \lambda A_0(y) = f. \tag{2.6}$$

Proof. We shall prove first that each solution $y \in D(A_0)$ to (2.6) is regular. Namely,

$$\beta(y) \in W^{1,q}_{\rm loc}(\mathbb{R}^d),\ \forall q \in \left[1, \tfrac{d}{d-1}\right), \qquad \text{if } d \ge 2, \tag{2.7}$$

$$\beta(y) \in W^{1,\infty}_{\rm loc}(\mathbb{R}^d), \qquad \text{if } d = 1, \tag{2.8}$$

and setting $\frac{d}{d-2} := \infty$ if $d = 2$

$$\begin{aligned} \|\nabla\beta(y)\|_{L^q(B_R)} + \|\beta(y)\|_{L^p(B_R)} \le C_R(|f|_1 + |y|_1),\ 1 < p < \tfrac{d}{d-2}, \\ \forall q \in \left[1, \tfrac{d}{d-1}\right), \end{aligned} \tag{2.9}$$

for all $B_R = \{x \in \mathbb{R}^d;\ |x| < \mathbb{R}\}$.

We note that (2.9) follows by (2.7) via the Sobolev–Gagliardo–Nirenberg inequality invoked below (see, e.g., [13], p. 278).

Consider first the case $d = 1$. We have $\beta(y) \in L^1$ by Hypothesis (i) and

$$(\beta(y))'' - (Db^*(y))' = \frac{1}{\lambda}(y - f) \text{ in } \mathcal{D}'(\mathbb{R}). \tag{2.10}$$

Since $Db^*(y) \in L^1(\mathbb{R})$, it follows that $(\beta(y))' \in L^1_{\rm loc}(\mathbb{R})$ and so $\beta(y) \in W^{1,1}_{\rm loc}(\mathbb{R})$. Then, by (iv), $b^*(y) \in W^{1,1}_{\rm loc}(\mathbb{R}^d)$ and so, by (2.10) and (ii) we infer that $Db^*(y) \in W^{1,1}_{\rm loc}(\mathbb{R}^d)$ and $(\beta(y))'' \in L^1_{\rm loc}(\mathbb{R}^d)$. Hence, $\beta(y) \in W^{1,\infty}_{\rm loc}(\mathbb{R}^d)$, as claimed.

Consider now the case $2 \le d$. By (2.6), we have, for all $\varphi \in C^2_0(\mathbb{R}^d)$,

$$\begin{aligned}
\frac{1}{\lambda}(y-f)\varphi &= \varphi\Delta\beta(y) - \varphi\,{\rm div}(Db^*(y))\\
&= \Delta(\varphi\beta(y)) - \beta(y)\Delta\varphi - {\rm div}(D\varphi b^*(y)) - 2\nabla\beta(y)\cdot\nabla\varphi + (D\cdot\nabla\varphi)b^*(y)\\
&= \Delta(\varphi\beta(y)) + \beta(y)\Delta\varphi + (D\cdot\nabla\varphi)b^*(y) - {\rm div}(2\beta(y)\nabla\varphi + D\varphi b^*(y)),
\end{aligned}$$

and, therefore,

$$\Delta(\varphi\beta(y)) = f_1 + {\rm div}\, f_2 \ \text{ in } \mathcal{D}'(\mathbb{R}^d), \tag{2.11}$$

where

$$\begin{aligned}
f_1 &= \tfrac{1}{\lambda}(y - f)\varphi - \beta(y)\Delta\varphi - (D\cdot\nabla\varphi)b^*(y)\\
f_2 &= 2\beta(y)\nabla\varphi + D\varphi b^*(y).
\end{aligned} \tag{2.12}$$

We set $u = \varphi\beta(y)$, $u_\varepsilon = u * \Psi_\varepsilon$, $f^\varepsilon_i = f_i * \Psi_\varepsilon$, $i = 1, 2$, where Ψ_ε is a standard mollifier, that is,

$$\Psi_\varepsilon(x) = \frac{1}{\varepsilon^d}\,\Psi\left(\frac{x}{\varepsilon}\right),\ \Psi \in C^\infty_0(\mathbb{R}^d), \text{support}\,\Psi \subset \{x; |x| \le 1\},\ \int_{\mathbb{R}^d} \Psi(x)dx = 1.$$

Let $\mathcal{O}, \mathcal{O}'$ be open balls in $\mathbb{R}^d$ centered at zero such that $\overline{\mathcal{O}}' \subset \mathcal{O}$ and choose $\varphi \in C^2_0(\mathbb{R}^d)$ such that $\varphi = 1$ on $\mathcal{O}'$ and $(\text{supp } \varphi)_\varepsilon \subset \mathcal{O}$, $\varepsilon \in (0, 1]$, where $(\text{supp}\,\varphi)_\varepsilon$ denotes the closed ε-neighbourhood of $\text{supp}\,\varphi$. Then, by (2.11) we have

$$\Delta u_\varepsilon = f^\varepsilon_1 + {\rm div}\, f^\varepsilon_2 \ \text{ in } \mathcal{O}, \quad u_\varepsilon \in C^\infty_0(\mathcal{O}). \tag{2.13}$$

Hence, by the uniqueness of the solution u_ε to (2.13), we have $u_\varepsilon = u^1_\varepsilon + u^2_\varepsilon$, where $u^1_\varepsilon, u^2_\varepsilon \in C^\infty(\mathcal{O}) \cap C(\overline{\mathcal{O}})$ are the solutions to the boundary value problems

$$\Delta u^1_\varepsilon = f^\varepsilon_1 \text{ in } \mathcal{O}, \qquad u^1_\varepsilon = 0 \text{ on } \partial\mathcal{O}, \tag{2.14}$$

$$\Delta u^2_\varepsilon = {\rm div}\, f^\varepsilon_2 \text{ in } \mathcal{O}, \qquad u^2_\varepsilon = 0 \text{ on } \partial\mathcal{O}. \tag{2.15}$$

By the standard existence theory for elliptic equations, we know that (see, e.g., [15], Corollary 12)

$$\|u_\varepsilon^1\|_{W_0^{1,q}(\mathcal{O})} \le C\|f_1^\varepsilon\|_{L^1(\mathcal{O})} \le C(|y|_1+|f|_1),\ \forall\,\varepsilon>0, \tag{2.16}$$

where $1\le q<\frac{d}{d-1}$ and where we used Hypotheses (i)–(iii) for the last inequality. By the Sobolev–Galiardo–Nirenberg theorem (see, e.g., [13], p. 278 and p. 281), it follows by (2.16) that we have

$$\|u_\varepsilon^1\|_p \le C(|f|_1+|y|_1),\quad \forall\,p\in\left[1,\tfrac{d}{d-2}\right)\ \text{ if } d>2, \tag{2.17}$$

$$\|u_\varepsilon^1\|_p \le C(|f|_1+|y|_1),\quad \forall\,p\ge 1\ \text{ if } d=2. \tag{2.18}$$

(In the following, we shall denote by the same symbol C several positive constants independent of ε and $\|\cdot\|_p$ is the norm of $L^p(\mathcal{O})$.)

Consider now the solution u_ε^2 to equation (2.15).

If $\psi\in L^m(\mathcal{O}),\ m>d$, and $\theta\in W^{2,m}(\mathcal{O})\cap W_0^{1,m}(\mathcal{O})$ is the solution to the Dirichlet problem

$$-\Delta\theta=\psi \text{ in } \mathcal{O};\quad \theta=0 \text{ on } \partial\mathcal{O}, \tag{2.19}$$

we see by (2.15) and by the Morrey embedding theorem (see [13], p. 282) that $\nabla\theta\in L^\infty(\mathcal{O})$ and, therefore, by Green's formula, since $u_\varepsilon^2=\theta=0$ on $\partial\mathcal{O}$,

$$\int_{\mathcal{O}} u_\varepsilon^2\Delta\theta dx=-\int_{\mathcal{O}} f_2^\varepsilon\cdot\nabla\theta dx\le |f_2^\varepsilon|_1\|\nabla\theta\|_\infty\le C(|f|_1+|y|_1)\|\psi\|_m. \tag{2.20}$$

This yields

$$\left|\int_{\mathcal{O}} u_\varepsilon^2\psi dx\right|\le C(|f|_1+|y|_1)\|\psi\|_m,\ \forall\psi\in L^m(\mathcal{O}). \tag{2.21}$$

Then, if $\frac{1}{m'}=1-\frac{1}{m}$, by (2.21) it follows by duality that $u_\varepsilon^2\in L^{m'}(\mathcal{O})\subset L^q(\mathcal{O})$ for all $q\in\left[1,\frac{d}{d-1}\right)$ and

$$\|u_\varepsilon^2\|_q\le C(|f|_1+|y|_1),\quad \forall\,\varepsilon>0,$$

and so, by (2.16), it follows also that $u_\varepsilon^i\in L^q(\mathcal{O})$, $i=1,2$, and

$$\|u_\varepsilon^i\|_q\le C(|f|_1+|y|_1),\ \forall\,q\in\left[1,\tfrac{d}{d-1}\right),\ i=1,2.$$

Hence,

$$\|u_\varepsilon\|_q \le C(|f|_1 + |y|_1),\ \forall\, \varepsilon > 0,\ q \in \big[1, \tfrac{d}{d-1}\big). \tag{2.22}$$

Finally, taking into account that $u_\varepsilon = (\varphi\beta(y)) * \Psi_\varepsilon$, by letting $\varepsilon \to 0$ we see by (2.22) and (1.5) that

$$\|\varphi\beta(y)\|_q \le C(|f|_1 + |y|_1),\quad \forall\, q \in \big[1, \tfrac{d}{d-1}\big).$$

Because φ and the corresponding ball $\mathcal{O}$ are arbitrary, we conclude that y, $\beta(y) \in L^q_{\rm loc}(\mathbb{R}^d)$ and that (for a possible larger C, still independent of ε)

$$\|\beta(y)\|_q \le C(|f|_1 + |y|_1),\ \forall q \in \big[1, \tfrac{d}{d-1}\big).$$

In particular, by Hypothesis (iv), this implies that

$$\|f_2\|_q \le C(|f|_1 + |y|_1)$$

and, therefore,

$$\|f_2^\varepsilon\|_q \le C(|f|_1 + |y|_1),\ \forall \varepsilon > 0, q \in \big[1, \tfrac{d}{d-1}\big). \tag{2.23}$$

Now, we shall improve the last estimate by invoking a bootstrap argument. Namely, we take in (2.19) $\psi \in L^\ell(\mathcal{O})$, where $\frac{d}{2} < \ell$. This yields as above that

$$\left|\int_{\mathcal{O}} u_\varepsilon^2 \psi\, dx\right| \le \int_{\mathcal{O}} |f_2^\varepsilon|\, |\nabla\theta| dx \le C\|f_2^\varepsilon\|_q \|\nabla\theta\|_{q'}$$

for all $q \in \big[1, \frac{d}{d-1}\big)$ and $q' = \frac{q}{q-1} > d$. Again by the Sobolev inequality we have, for all $\ell \in \big(\frac{d}{2}, d\big)$ that $q' := d\ell/(d-\ell) > d$ and

$$\|\nabla\theta\|_{q'} \le C\|\theta\|_{W^{2,\ell}(\mathcal{O})} \le C\|\psi\|_\ell. \tag{2.24}$$

This yields, for all $\ell \in \big(\frac{d}{2}, d\big)$,

$$\left|\int_{\mathcal{O}} u_\varepsilon^2 \psi\, dx\right| \le C\|f_2^\varepsilon\|_q \|\psi\|_\ell \le C(|f|_1 + |y|_1)\|\psi\|_\ell,\ \forall \psi \in L^\ell(\mathcal{O}), \tag{2.25}$$

and, therefore, putting $\frac{d}{d-2} := \infty$, if $d = 2$,

$$\|u_\varepsilon^2\|_r \le C(|f|_1 + |y|_1),\ \forall r \in \big[1, \tfrac{d}{d-2}\big), \tag{2.26}$$

Then, by (2.17), (2.18), we get

$$\|u_\varepsilon\|_r \le C(|f|_1 + |y|_1),\ \forall\, r \in \big[1, \tfrac{d}{d-2}\big). \tag{2.27}$$

Letting $\varepsilon \to 0$, this yields the bound for $\|y\|_{L^p(B_R)}$ in (2.9), since $\varphi \in C_0^2(\mathbb{R}^d)$ was arbitrary. Furthermore, (2.27) implies that (2.23) is strengthened to

$$\|f_2^\varepsilon\|_\nu \le C(|f|_1 + |y|_1),\ \forall \nu \in \left[1, \tfrac{d}{d-2}\right),\ \varepsilon > 0. \tag{2.28}$$

For $\varepsilon \to 0$, this yields, since f_2 has compact support in $\mathcal{O}$,

$$|f_2|_\nu \le C(|f|_1 + |y|_1), \quad \forall \nu \in \left[1, \tfrac{d}{d-2}\right). \tag{2.29}$$

Hence, for $d \in [2, 3]$, we get

$$\|f_2^\varepsilon\|_2 \le C(|f|_1 + |y|_1),\ \forall \varepsilon > 0, \tag{2.30}$$

and so

$$\|\mathrm{div}(f_2^\varepsilon)\|_{H^{-1}(\mathcal{O})} \le C(|f|_1 + |y|_1),\ \forall \varepsilon > 0.$$

Then, by equation (2.15) and, since $\frac{d}{d-1} \le 2$, it follows that $u_2^\varepsilon \in H_0^1(\mathcal{O})$ and

$$\|u_\varepsilon^2\|_{W^{1,q}(\mathcal{O})} \le \|u_\varepsilon^2\|_{H_0^1(\mathcal{O})} \le C(|f|_1 + |y|_1),\ \forall \varepsilon > 0. \tag{2.31}$$

Hence, by (2.16), we have

$$\|u_\varepsilon\|_{W_0^{1,q}(\mathcal{O})} \le C(|y|_1 + |f|_1),\ \forall\, \varepsilon > 0,\ \text{if } d = 2, 3. \tag{2.32}$$

Letting $\varepsilon \to 0$ we get the estimate

$$\|u\|_{W^{1,q}(\mathcal{O})} \le C(|f|_1 + |y|_1), \quad d = 2, 3. \tag{2.33}$$

Recalling that $u = \varphi\beta(y)$, since φ and the corresponding ball $\mathcal{O}$ are arbitrary, this implies (2.7), for $d \in [2, 3]$, as claimed.

We shall consider now the case $d \ge 3$. To this end, we come back to equation (2.11) and note that $\varphi\beta(y) = u_1 + u_2$, where u_1, u_2 are solutions to the equations

$$\Delta u_1 = f_1 \text{ in } \mathcal{D}'(\mathbb{R}^d), \tag{2.34}$$

$$\Delta u_2 = \mathrm{div}(f_2) \text{ in } \mathcal{D}'(\mathbb{R}^d), \tag{2.35}$$

where f_1, f_2 are defined by (2.12).

Since $f_1 \in L^1$, it follows by [12, Lemma A.5] that u_1 is given by the representation formula

$$u_1 = -E * f_1 \ \text{ in } \mathbb{R}^d,$$

where $E(x) \equiv \frac{1}{(d-2)\omega_d |x|^{d-2}}$ is the fundamental solution to Δ.

Hence (see, e.g, [12]), $u_1 \in M^{\frac{d}{d-2}}(\mathbb{R}^d) \subset L^p_{\rm loc}(\mathbb{R}^d)$, $\forall p \in \left[1, \frac{d}{d-2}\right)$ and $|\nabla u_1| = |\nabla E * f_1| \in M^{\frac{d}{d-1}}(\mathbb{R}^d) \subset L^p_{\rm loc}(\mathbb{R}^d)$, $\forall p \in \left[1, \frac{d}{d-1}\right)$ with

$$\|\nabla u_1\|_{L^p(B_R)} \le C(|y|_1 + |f|_1),\ \forall R > 0,\ p \in \left[1, \tfrac{d}{d-1}\right). \tag{2.36}$$

(Here, M^ℓ is the Marcinkievicz space of order ℓ.)

As regards the solution u_2 to equation (2.35), we note that from (2.29) it follows that $f_2 \in L^p(\mathbb{R}^d)$, $\forall p \in \left[1, \frac{d}{d-1}\right)$. Let $u_{2,\varepsilon} = u_2 * \Psi_\varepsilon$. Then

$$u_{2,\varepsilon} = -\nabla E * f_2^\varepsilon \text{ in } \mathbb{R}^d.$$

Taking into account that $|\nabla^2 E(x)| \le C|x|^d$, $\forall x \ne 0$, it follows by the Calderon–Zygmund theorem (see, e.g, [16] and estimate (2.29)) that

$$|\nabla u_{2,\varepsilon}|_p \le C|f_2^\varepsilon|_p \le C(|f|_1 + |y|_1),\ \forall p \in \left[1, \tfrac{d}{d-1}\right)$$

and, after letting $\varepsilon \to 0$, together with (2.36) this yields

$$\|\nabla u\|_{L^p(B_R)} \le C_R(|f|_1 + |y|_1),\ \forall p \in \left[1, \tfrac{d}{d-1}\right),$$

and so (2.7) and (2.9) hold for all $d \ge 2$.

Now, let us prove the uniqueness of the solution to (2.6).

If $y_1, y_2 \in D(A_0)$ are two solutions, we have

$$y_1 - y_2 - \lambda\Delta(\beta(y_1) - \beta(y_2)) + \lambda\,{\rm div}(D(b^*(y_1) - b^*(y_2))) = 0. \tag{2.37}$$

Let $\eta \in C^2([0,\infty))$ be such that

$$\eta(r) \ge 0,\ \eta(r) = 1,\ \forall r \in [0,1];\ \eta(r) = 0,\ \forall r\ \in [2,\infty).$$

We set

$$\varphi_n(x) = \eta\left(\frac{|x|^2}{n}\right),\ \forall x \in \mathbb{R}^d,\ n \in \mathbb{N},$$

and note that

$$|\nabla\varphi_n(x)| \ \le\ \frac{4}{\sqrt{n}}\,|\eta'|_\infty,\ \forall\, x \in \mathbb{R}^d, \tag{2.38}$$

$$|\Delta\varphi_n(x)| \ \le\ \frac{1}{n}\,(2d|\eta'|_\infty + 8|\eta''|_\infty),\ \forall\, x \in \mathbb{R}^d. \tag{2.39}$$

By (2.37), we have

$$
\begin{aligned}
&\varphi_n(y_1-y_2)-\lambda\Delta(\varphi_n(\beta(y_1)-\beta(y_2)))+\lambda\,\mathrm{div}(D\varphi_n(b^*(y_1)-b^*(y_2)))\\
&\quad=\lambda(\nabla\varphi_n\cdot D)(b^*(y_1)-b^*(y_2))-\lambda(\beta(y_1)-\beta(y_2))\Delta\varphi_n\\
&\quad-2\lambda\nabla\varphi_n\cdot\nabla(\beta(y_1)-\beta(y_2))\text{ in }\mathcal{D}'(\mathbb{R}^d).
\end{aligned}
\tag{2.40}
$$

Let $\mathcal{X}_\delta:\mathbb{R}\to\mathbb{R}$ be the function

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

and let

$$
j_\delta(s)=\int_0^s\mathcal{X}_\delta(r)dr,\ \forall\,s\in\mathbb{R}.
$$

We know by (2.7)–(2.9) that

$$
\beta(y_i)\in L^p_{\mathrm{loc}},\ \forall p\in\left[1,\tfrac{d}{d-2}\right),\ \nabla\beta(y_i)\in L^q_{\mathrm{loc}},\ \forall q\in\left[1,\tfrac{d}{d-1}\right).
$$

Moreover, by Hypothesis (iv) it follows that

$$
|b^*(y_i)|\le\alpha_2|\beta(y_i)|,\quad|\nabla b^*(y_i)|\le\alpha_2|\nabla\beta(y_i)|,\quad\text{a.e. on }\mathbb{R}^d,\ i=1,2,
$$

and, therefore,

$$
b^*(y_i)\in L^p_{\mathrm{loc}},\ |\nabla b^*(y_i)|\in L^q_{\mathrm{loc}},\ i=1,2,\ \forall\,p\in\left[1,\tfrac{d}{d-2}\right),q\in\left[1,\tfrac{d}{d-1}\right).
$$

This implies that

$$
\mathrm{div}(Db^*(y_i))=D\cdot\nabla b^*(y_i)+b^*(y_i)\mathrm{div}\,D\in L^1_{\mathrm{loc}},\ i=1,2,
$$

because, by (ii), $\mathrm{div}(D)\in L^m_{\mathrm{loc}}$ for some $m>\frac{d}{2}$. Since $y_i\in D(A_0)$, $i=1,2$, we have therefore that $\Delta(\varphi_n(\beta(y_1)-\beta(y_2)))$ and $\varphi_n(\beta(y_1)-\beta(y_2))$ are in $L^1(\mathbb{R}^d)$. This yields

$$
\begin{aligned}
&-\int_{\mathbb{R}^d}\Delta(\varphi_n(\beta(y_1)-\beta(y_2)))\mathcal{X}_\delta(\beta(y_1)-\beta(y_2))dx\\
&\quad=\int_{\mathbb{R}^d}\nabla(\varphi_n(\beta(y_1)-\beta(y_2)))\cdot\nabla(\beta(y_1)-\beta(y_2))\mathcal{X}'_\delta(\beta(y_1)-\beta(y_2))dx\\
&\quad\ge\frac{1}{\delta}\int_{[|\beta(y_1)-\beta(y_2)|\le\delta]}(\nabla\varphi_n\cdot\nabla(\beta(y_1)-\beta(y_2)))(\beta(y_1)-\beta(y_2))dx,
\end{aligned}
$$

and so, by (2.38)–(2.40), we have

$$\begin{aligned}&\int_{\mathbb{R}^d}\varphi_n(y_1-y_2)\mathcal{X}_\delta(\beta(y_1)-\beta(y_2))dx\\&\qquad+\frac{\lambda}{\delta}\int_{[|\beta(y_1)-\beta(y_2)|\le\delta]}\varphi_n(b^*(y_1)-b^*(y_2))(D\cdot\nabla(\beta(y_1)-\beta(y_2)))dx\\&\qquad\le\frac{C\lambda}{\sqrt{n}}+I^\delta_{\lambda,n},\end{aligned}\tag{2.41}$$

where

$$I^\delta_{\lambda,n}\le\lambda|D|_\infty\int_{[|\beta(y_1)-\beta(y_2)|\le\delta]}|\varphi_n|\cdot|\nabla(\beta(y_1)-\beta(y_2))|dx\to 0\ \text{ as }\ \delta\to 0,$$

because

$$|\nabla\beta(y_1)-\nabla\beta(y_2)|=0,\ \text{a.e. on }\{x;\,|\beta(y_1)(x)-\beta(y_2)(x)|=0\}.$$

To obtain (2.41), we have used the relation

$$\begin{aligned}&-2\lambda\int_{\mathbb{R}^d}(\nabla\varphi_n\cdot\nabla(\beta(y_1)-\beta(y_2)))\mathcal{X}_\delta(\beta(y_1)-\beta(y_2))dx\\&\quad=-2\lambda\int_{\mathbb{R}^d}\nabla\varphi_n\cdot\nabla j_\delta(\beta(y_1)-\beta(y_2))dx\\&\quad=2\lambda\int_{\mathbb{R}^d}\Delta\varphi_n j_\delta(\beta(y_1)-\beta(y_2))dx\\&\quad\le\frac{2\lambda}{n}(2d|\eta'|_\infty+8|\eta''|_\infty)\int_{\mathbb{R}^d}(|\beta(y_1(x))|+|\beta(y_2(x))|)dx\\&\quad\le\frac{4\lambda}{n}\,\alpha_1(d|\eta'|_\infty+4|\eta''|_\infty)(|y_1|_1+|y_2|_1)\le\frac{C\lambda}{\sqrt{n}}\,|f|_1,\end{aligned}$$

where C is independent of n.

On the other hand, recalling that

$$D\in L^\infty(\mathbb{R}^d),\ \nabla\beta(y_i)\in L^q_{\rm loc},\ 1\le q<\frac{d}{d-1},$$

while by Hypothesis (iv), we have

$$|b^*(y_1)-b^*(y_2)|\le\alpha_2|\beta(y_1)-\beta(y_2)|,\ \text{a.e. in }\mathbb{R}^d,$$

and so, by (2.41), it follows for $\delta \to 0$ that

$$\int_{\mathbb{R}^d} \varphi_n(y_1 - y_2)\operatorname{sign}(\beta(y_1) - \beta(y_2))dx \le \frac{C\lambda}{\sqrt{n}}\,|f|_1.$$

This yields for $n \to \infty$ that

$$|y_1 - y_2|_1 = 0,$$

as claimed. $\square$

Lemma 2.5. *Assume that $d \ge 1$. Then, for each $f \in L^1(\mathbb{R}^d)$ and all $\lambda > 0$, equation* (2.6) *has a unique solution $y = J_\lambda(f)$. Moreover, one has*

$$|J_\lambda(f_1) - J_\lambda(f_2)|_1 \le |f_1 - f_2|_1,\ \forall f_1, f_2 \in L^1,\ \lambda > 0, \tag{2.42}$$

$$J_\lambda(f) \in \mathcal{P},\ \forall\, f \in \mathcal{P},\ \forall \lambda > 0, \tag{2.43}$$

and $\overline{D(A_0)} = L^1$.

Proof. The proof of Lemma 2.5 was given under the assumptions (i)–(iii) in [5] (see also [8]), so here its proof under our assumptions (i)–(iii) will be outlined only.

We assume first that $f \in L^1 \cap L^2$ and approximate equation (2.6) by

$$y + \lambda(\varepsilon I - \Delta)(\beta(y) + \varepsilon y) + \lambda \operatorname{div}(D_\varepsilon b^*_\varepsilon(y)) = f, \tag{2.44}$$

where $b^*_\varepsilon \in C^1_b(\mathbb{R}) \cap C_b(\mathbb{R})$ is a smooth approximation of b^* such that

$$|b^*_\varepsilon(r)| \le C|r|,\ \lim_{\varepsilon\to 0} b^*_\varepsilon(r) = b(r)r \text{ uniformly on compacts,}$$

and

$$D_\varepsilon = \eta_\varepsilon D,\ \eta_\varepsilon \in C^1_0(\mathbb{R}^d),\ 0 \le \eta_\varepsilon \le 1,\ |\nabla \eta_\varepsilon| \le 1,\ \eta_\varepsilon(x) = 1 \text{ if } |x| < \frac{1}{\varepsilon}.$$

Clearly, we have

$$\begin{gathered} |D_\varepsilon| \in L^\infty \cap L^2,\ |D_\varepsilon| \le |D|,\ \lim_{\varepsilon\to\infty} D_\varepsilon(x) = D(x),\ \text{a.e. } x \in \mathbb{R}^d. \\ \operatorname{div} D_\varepsilon \in L^1,\ (\operatorname{div} D_\varepsilon)^- \le (\operatorname{div} D)^- + 1\!\!1_{\left[|x|>\frac{1}{\varepsilon}\right]}|D|. \end{gathered} \tag{2.45}$$

A typical example for b_ε is

$$b_\varepsilon \equiv b * \varphi_\varepsilon, \quad b^*_\varepsilon(r) \equiv \frac{b_\varepsilon(r)r}{1+\varepsilon|r|}, \quad r \in \mathbb{R},$$
$$\varphi_\varepsilon(r) = \frac{1}{\varepsilon}\,\varphi\left(\frac{r}{\varepsilon}\right),\ \varphi \in C^\infty_0(\mathbb{R}),\ \int_{\mathbb{R}} \varphi(x)dx = 1.$$

We can rewrite (2.44) equivalently as the following equation on L^2:

$$\lambda(\beta(y)+\varepsilon y)+(\varepsilon I-\Delta)^{-1}y+\lambda(\varepsilon I-\Delta)^{-1}\mathrm{div}(D_\varepsilon b^*_\varepsilon(y)) = (\varepsilon I-\Delta)^{-1}f. \quad (2.46)$$

We set

$$F(y) = \lambda(\beta(y)+\varepsilon y)+(\varepsilon I-\Delta)^{-1}y+\lambda(\varepsilon I-\Delta)^{-1}\mathrm{div}(D_\varepsilon b^*_\varepsilon(y))$$

and note that

$$\begin{aligned}
&(F(y_1)-F(y_2), y_1-y_2)_2 = \varepsilon|(\varepsilon I-\Delta)^{-1}(y_1-y_2)|^2_2\\
&\quad +\lambda(\beta(y_1)-\beta(y_2), y_1-y_2)_2 + \varepsilon\lambda|y_1-y_2|^2_2\\
&\quad +|\nabla(\varepsilon I-\Delta)^{-1}(y_1-y_2)|^2_2 - \lambda(D_\varepsilon(b^*_\varepsilon(y_1)-b^*_\varepsilon(y_2)), \nabla(\varepsilon I-\Delta)^{-1}(y_1-y_2))_2\\
&\quad \ge \varepsilon\lambda|y_1-y_2|^2_2 + |(\varepsilon I-\Delta)^{-1}(y_1-y_2)|^2_2 + |\nabla(\varepsilon I-\Delta)^{-1}(y_1-y_2)|^2_2\\
&\quad -\lambda|D_\varepsilon|_\infty |b^*_\varepsilon|_{\mathrm{Lip}}|y_1-y_2|_2|\nabla(\varepsilon I-\Delta)^{-1}(y_1-y_2)|_2 \ge 0,\ \text{for } 0<\lambda<\lambda_\varepsilon.
\end{aligned}$$

It is also clear that $(F(y),y)_2 \ge \alpha_\varepsilon\lambda|y|^2_2$ and so, for $0<\lambda<\lambda_\varepsilon$, F is monotone, continuous and coercive on $L^2(\mathbb{R}^d)$. Hence, it is surjective and so equation (2.46) has a solution $y_\varepsilon \in L^2(\mathbb{R}^d)$ and the latter is then true for all $\lambda > 0$ (see Propositions 3.1 and 3.2 in [2]). By (2.44), it follows also that $y_\varepsilon, \beta(y_\varepsilon) \in H^1(\mathbb{R}^d)$.

If $f \in L^1 \cap L^\infty$, we have

$$|y_\varepsilon|_\infty \le (1+||D|+(\mathrm{div}\,D)^-|^{\frac{1}{2}}_\infty)|f|_\infty, \quad 0<\lambda<\lambda_0<1. \quad (2.47)$$

Indeed, by (2.44) we see that, for $M = |(\mathrm{div}\,D_\varepsilon)^-|^{\frac{1}{2}}_\infty|f|_\infty$ and $\lambda<\lambda_0$,

$$\begin{aligned}
&(y_\varepsilon-|f|_\infty-M)-\lambda\Delta(\beta_\varepsilon(y_\varepsilon)-\beta_\varepsilon(|f|_\infty+M))\\
&\quad +\lambda\varepsilon(\beta_\varepsilon(u_\varepsilon)-\beta_\varepsilon(|f|_\infty+M))+\lambda\,\mathrm{div}(D_\varepsilon(b^*_\varepsilon(u_\varepsilon)-b^*_\varepsilon(|f|_\infty+M)))\\
&\quad \le f-|f|_\infty-M-\lambda b^*_\varepsilon(M+|f|_\infty)\mathrm{div}\,D_\varepsilon \le 0,
\end{aligned}$$

where $\beta_\varepsilon(r) = \beta(r)+\varepsilon r$.

Multiplying the above equation by $\mathcal{X}_\delta((y_\varepsilon-(|f|_\infty+M))^+)$ and integrating over $\mathbb{R}^d$, we get as above, for $\delta\to 0$, $|(y_\varepsilon-|f|_\infty-M)^+|_1\le 0$ and, therefore, by (2.45) and since $b\ge 0$,

$$y_\varepsilon \le (1+||D|+(\operatorname{div} D)^-|_\infty^{\frac12})|f|_\infty,\ \text{a.e. in}\mathbb{R}^d.$$

Similarly, one gets that

$$y_\varepsilon \ge -(1+||D|+(\operatorname{div} D)^-|_\infty^{\frac12})|f|_\infty,\ \text{a.e. in } \mathbb{R}^d,$$

and so (2.47) follows.

Let us denote the solution to (2.44) by $y^\varepsilon_\lambda(f)$ and define $\beta_\varepsilon(r):=\beta(r)+\varepsilon r$, $r\in\mathbb{R}$. Then, we multiply the equation

$$\begin{aligned}(y^\varepsilon_\lambda(f_1)-y^\varepsilon_\lambda(f_2))+\lambda(\varepsilon I-\Delta)(\beta_\varepsilon(y^\varepsilon_\lambda(f_1))-\beta_\varepsilon(y^\varepsilon_\lambda(f_2)))\\ +\lambda\operatorname{div}(D_\varepsilon(b^*_\varepsilon(y^\varepsilon_\lambda(f_1))-b^*_\varepsilon(y^\varepsilon_\lambda(f_2)))=f_1-f_2\end{aligned}$$

by $\mathcal{X}_\delta(\beta_\varepsilon(y^\varepsilon_\lambda(f_1))-\beta_\varepsilon(y^\varepsilon_\lambda(f_2)))$ and integrate over $\mathbb{R}^d$. We set for $\lambda\in(0,\lambda_1)$ and $\delta>0$

$$E^\varepsilon_{\lambda,\delta}=\{x\in\mathbb{R}^d;\ |\beta_\varepsilon(y^\varepsilon_\lambda(f_1)(x))-\beta_\varepsilon(y^\varepsilon_\lambda(f_2)(x))|\le\delta\}.$$

Since $|\beta_\varepsilon(r)-\beta_\varepsilon(\bar r)|\ge\varepsilon|r-\bar r|$, $r,\bar r\in\mathbb{R}$, and b^*_ε is Lipschitz, we have

$$\begin{aligned}&\int_{\mathbb{R}^d}(y^\varepsilon_\lambda(f_1)-y^\varepsilon_\lambda(f_2))\mathcal{X}_\delta(\beta_\varepsilon(y^\varepsilon_\lambda(f_1))-\beta_\varepsilon(y^\varepsilon_\lambda(f_2)))dx \qquad (2.48)\\ &\le|f_1-f_2|_1+\lambda\int_{\mathbb{R}^d}(b^*_\varepsilon(y^\varepsilon_\lambda(f_1))-b^*_\varepsilon(y^\varepsilon_\lambda(f_2)))\\ &\qquad D_\varepsilon\cdot\nabla(\beta_\varepsilon(y^\varepsilon_\lambda(f_1)-\beta_\varepsilon(y^\varepsilon_\lambda(f_2)))\mathcal{X}'_\delta(\beta_\varepsilon(y^\varepsilon_\lambda(f_1))-\beta_\varepsilon(y^\varepsilon_\lambda(f_2)))dx\\ &\le|f_1-f_2|_1+\frac{C_\varepsilon\lambda}{\delta}\int_{E^\varepsilon_{\lambda,\delta}}|\beta_\varepsilon(y^\varepsilon_\lambda(f_1))-\beta_\varepsilon(y^\varepsilon_\lambda(f_2))||\nabla(\beta_\varepsilon(y^\varepsilon_\lambda(f_1)))-\beta_\varepsilon(y^\varepsilon_\lambda(f_2))||D_\varepsilon|dx\\ &\le|f_1-f_2|_1+C_\varepsilon\lambda|D_\varepsilon|_2\left(\int_{E^\varepsilon_{\lambda,\delta}}|\nabla(\beta_\varepsilon(y^\varepsilon_\lambda(f_1))-\beta_\varepsilon(y^\varepsilon_\lambda(f_2)))|^2dx\right)^{\frac12}.\end{aligned}$$

Then, letting $\delta\to 0$ in (2.48) and recalling that

$$\operatorname{sign}(\beta_\varepsilon(y^\varepsilon_\lambda(f_1))-\beta_\varepsilon(y^\varepsilon_\lambda(f_2)))=\operatorname{sign}(y^\varepsilon_\lambda(f_1)-y^\varepsilon_\lambda(f_2)),\ \text{a.e. in } \mathbb{R}^d,$$

we get by monotone convergence

$$|y_\lambda^\varepsilon(f_1) - y_\lambda^\varepsilon(f_2)|_1 \le |f_1 - f_2|_1. \tag{2.49}$$

In particular, for all $f \in L^1 \cap L^\infty$,

$$\sup_{\lambda,\varepsilon>0} |y_\lambda^\varepsilon(f)|_1 \le |f|_1. \tag{2.50}$$

Recall that by (2.47) we have for some $C_D \in (0,\infty)$

$$\sup_{\substack{\lambda\in(0,\lambda_0)\\ \varepsilon>0}} |y_\lambda^\varepsilon(t)|_\infty \le C_D|f|_\infty. \tag{2.51}$$

Hence, multiplying (2.44) by $\beta(y_\varepsilon)$ and integrating over $\mathbb{R}^d$, we see that, for some $C \in (0,\infty)$ and all $\lambda \in (0,\lambda_0)$, $\varepsilon \in (0,1)$,

$$\lambda|\nabla\beta_\varepsilon(y_\lambda^\varepsilon(f))|_2^2 \le C\left(|f|_\infty + \sup_{|r|\le C_D|f|_\infty} |\beta(r)|\right)|f|_1. \tag{2.52}$$

Now, fix $\lambda \in (0,\lambda_0)$. Set $y_\varepsilon := y_\lambda^\varepsilon(f)$. Then, by (2.50)–(2.52), $\{\beta_\varepsilon(y_\varepsilon)\}$ is bounded in H^1 and $\{y_\varepsilon\}$ is bounded in L^2.

This implies that $\{\beta_\varepsilon(y_\varepsilon)\}$ is compact in $L^2_{\rm loc}$ and, therefore, along a subsequence $\{\varepsilon\} \to 0$, we have

$$\begin{array}{rcll} y_\varepsilon & \to & y & \text{weakly in } L^2,\\ \beta(y_\varepsilon) & \to & v & \text{strongly in } L^2_{\rm loc},\\ \nabla\beta_\varepsilon(y_\varepsilon) & \to & \nabla v & \text{weakly in } L^2. \end{array}$$

Since the map $y \to \beta(y)$ is maximal monotone in $L^2(\mathcal{O})$ for every bounded, open $\mathcal{O} \subset \mathbb{R}^d$, it follows that $v(x) = \beta(y(x))$, a.e. $x \in \mathbb{R}^d$. Moreover, as β is, by Hypothesis (i), continuous and strictly monotone, it follows that

$$y_\varepsilon \to y \text{ a.e. on } \mathbb{R}^d, \tag{2.53}$$

and, therefore, we have

$$b_\varepsilon(y_\varepsilon)y_\varepsilon \to b(y)y, \text{ a.e. in } \mathbb{R}^d, \quad b_\varepsilon(y_\varepsilon)y_\varepsilon \to b(y)y \text{ weakly in } L^2,$$

selecting another subsequence $\{\varepsilon\} \to 0$, if necessary. Then, letting $\varepsilon \to 0$ in (2.44), we see that

$$y - \lambda\Delta\beta(y) + \lambda\,{\rm div}(Db(y)y) = f \text{ in } H^{-1}(\mathbb{R}^d). \tag{2.54}$$

Moreover, we see that $\beta(y) \in H^1(\mathbb{R}^d)$ and (2.50)–(2.52) hold for $y, \beta(y)$ replacing $y_\lambda^\varepsilon(f)$ and $\beta(y_\lambda^\varepsilon(f))$, respectively. We denote this solution to equation (2.54) by $J_\lambda(f)$. We shall now prove

$$|J_\lambda(f_1) - J_\lambda(f_2)|_1 \le |f_1 - f_2|_1, \ \forall f_1, f_2 \in L^1 \cap L^\infty, \ \lambda \in (0, \lambda_0), \qquad (2.55)$$

which implies, in particular, that $J_\lambda(f) \in L^1$, $\forall f \in L^1 \cap L^\infty$.

Let $f \in L^1$ be arbitrary but fixed and let $\{f_n\} \subset L^1 \cap L^\infty$ be such that $f_n \to f$ in L^1 as $n \to \infty$. We set $y_n = y(\lambda, f_n)$, that is,

$$y_n - \lambda\Delta\beta(y_n) + \lambda \operatorname{div}(Db(y_n)y_n) = f_n \text{ in } H^{-1}(\mathbb{R}^d). \qquad (2.56)$$

Then, by (2.49), we have

$$|y_n - y_m|_1 \le |f_n - f_m|_1, \ \forall n, m \in \mathbb{N},$$

and so there is $y = \lim\limits_{n\to\infty} y_n$ in L^1. Since A_0 is closed on L^1, it follows by (2.56) that $y \in D(A_0)$ and $y + \lambda A_0 y = f$ and so $R(I + \lambda A_0) = L^1$, as claimed. The fact that y is the unique solution to $y + \lambda A_0 y = f$ follows by Lemma 2.4. Denoting this solution y by $J_\lambda(f)$, we obtain by (2.55) that (2.42) holds.

If $f \in \mathcal{P} \cap L^\infty$, it follows by (2.44) that $y_\varepsilon \in \mathcal{P}$. By (2.51), (2.53), it follows that $y_\varepsilon \to J_\lambda(f)$ in $L^1_{\rm loc}$ as $\varepsilon \to 0$. Then, the argument from the proof of Lemma 3.3 in [5] implies that $y_\varepsilon \to J_\lambda(f)$ in L^1, so $J_\lambda(f) \in \mathcal{P}$. Finally, by density, (2.43) follows.

To prove that $\overline{D(A_0)} = L^1$, it suffices to note that, by (ii), (iii), $C_0^\infty(\mathbb{R}^d) \subset D(A_0)$ (because $\operatorname{div} D \in L^1_{\rm loc}$ and $\beta \in C^2(\mathbb{R}^d)$, $b \in C^1 \cap C_b$). This completes the proof of Lemma 2.5. □

Proof of Theorem 2.1 (continued). By Lemmas 2.4 and 2.5, it follows that A_0 is m-accretive in $L^1(\mathbb{R}^d)$ and $\overline{D(A_0)} = L^1$. Then, as mentioned earlier, the existence and uniqueness of a mild solution ρ to (1.1) follows by the Crandall & Liggett theorem (see [1], p. 154). Moreover, by (2.43) and (1.18) it follows that $\rho(t) \in \mathcal{P}$, $\forall t \ge 0$, if $\rho_0 \in \mathcal{P}$.

We shall show now that ρ is a distributional solution to (1.1). Since $\rho_h \to \rho$ in $L^1((0,T) \times \mathbb{R}^d)$ as $h \to 0$, we have along a subsequence $\{h\} \to 0$

$$\beta(\rho_h) \to \beta(\rho), \quad b^*(\rho_h) \to b^*(\rho), \text{ a.e. in } (0,\infty) \times \mathbb{R}^d.$$

Then, taking into account that $|\beta(\rho_h)| \le \alpha_1|\rho_h|$, a.e. in $(0,T) \times \mathbb{R}^d$, it follows by a standard argument and by Hypothesis (iv) that

$$\beta(\rho_h) \to \beta(\rho), \quad b^*(\rho) \to b^*(\rho) \text{ in } L^1((0,\infty) \times \mathbb{R}^d)$$

as $h \to 0$. By (1.10), we have

$$
\begin{aligned}
\int_h^\infty\int_{\mathbb{R}^d}\Big(\frac{1}{h}(\rho_h(t,x)-\rho_h(t-h,x))\varphi(t,x)-\beta(\rho_h(t,x))&\\
\cdot\Delta\varphi(t,x)-(D(x)\cdot\nabla\varphi(t,x))b^*(\rho_h(t,x))\Big)dx\,dt=0&\\
\forall\varphi\in C_0^\infty([0,\infty)\times\mathbb{R}^d).&
\end{aligned}\tag{2.57}
$$

Taking into account that

$$
\begin{aligned}
\int_h^\infty\int_{\mathbb{R}^d}\rho_h(t-h,x)\varphi(t,x)dx\,dt=&\int_0^\infty\int_{\mathbb{R}^d}\rho_h(t,x)\varphi(t+h,x)dx\,dt\\
&+\int_0^h\int_{\mathbb{R}^d}\rho_0(x)\varphi(t+h,x)dx\,dt,
\end{aligned}
$$

and, letting $h\to 0$ in (2.57), it follows that (1.24) holds, as claimed.

If $\rho_0\in L^1\cap L^\infty$, it follows as in Theorem 2.2 in [5] that $\rho\in L^\infty((0,T)\times\mathbb{R}^d)$, $\forall T>0$. This completes the proof. □

Remark 2.6. Analyzing the proof of Lemma 2.5, one sees that, as far as concerns the existence of a solution to equation (2.6), Hypothesis (1.5) can be dispensed with. Moreover, the condition $b\ge 0$ in Hypothesis (iii), which was used in Lemma 2.5 to prove (2.47), is no longer necessary for existence of a solution to (2.6) if (1.5) is strengthen to $\gamma_1|r|\le|\beta(r)|\le\alpha_1|r|$, $\forall r\in\mathbb{R}$.

Proof of Theorem 2.2.

Assume that Hypotheses (i)–(v) hold and let $\rho_0\in L^1\cap L^2$. We set $g(r)\equiv\int_0^r\beta(s)ds$. Then, multiplying (1.8) by $\beta(\rho_h^{j+1})\in H^1(\mathbb{R}^d)$ and integrating over $\mathbb{R}^d$, we get

$$
\begin{aligned}
&\int_{\mathbb{R}^d}g(\rho_h^{j+1}(x))dx+h\int_{\mathbb{R}^d}|\nabla\beta(\rho_h^{j+1}(x))|^2dx\\
&\quad\le h\int_{\mathbb{R}^d}b^*(\rho_h^{j+1}(x))D(x)\cdot\nabla\beta(\rho_h^{j+1}(x))dx+\int_{\mathbb{R}^d}g(\rho_h^j(x))dx,\\
&\qquad\qquad j=0,1,...,N.
\end{aligned}\tag{2.58}
$$

Next, by (2.58) we get

$$
\begin{aligned}
&\int_{\mathbb{R}^d}g(\rho_h(t,x))dx+\int_0^t\int_{\mathbb{R}^d}|\nabla\beta(\rho_h(s,x))|^2dx\,ds\\
&\quad\le\int_{\mathbb{R}^d}g(\rho_0(x))dx+C\int_0^t\int_{\mathbb{R}^d}\rho_h^2(s,x)ds\,dx,\ \forall\,t\in(0,T),
\end{aligned}
$$

and, since by assumption (v) we have $\alpha_0 r^2 \le g(r) \le \alpha_1 r^2$, $\forall r \in \mathbb{R}$, we get

$$|\rho_h(t)|_2^2 + |\beta(\rho_h)|^2_{L^2(0,T;H^1)} \le C_T|\rho_0|_2^2, \quad \forall\, h,\ t \in [0,T]. \tag{2.59}$$

For $h \to 0$, $\rho_h \to \rho$ in $L^\infty(0,T;L^1)$ and so, along a subsequence $\rho_h \to \rho$, a.e. in $(0,T)\times\mathbb{R}^d$, we have

$$\beta(\rho_h) \to \beta(\rho), \text{ a.e. on } (0,T)\times\mathbb{R}^d. \tag{2.60}$$

Then, by (1.5) it follows that $\{\beta(\rho_h)\}$ is Lebesgue equi-integrable on $(0,T)\times\mathbb{R}^d$. Hence, by the generalized Lebesgue convergence theorem (see, e.g., [10, Theorem 21.4]) it follows that $\beta(\rho_h) \to \beta(\rho)$ in $L^1((0,T)\times\mathbb{R}^d)$. Then, we have along a subsequence $\{h\} \to 0$

$$\begin{aligned}
\rho_h \;&\to\; \rho \qquad\quad \text{strongly in } L^1((0,T)\times\mathbb{R}^d), && (2.61)\\
&\qquad\qquad\quad\ \ \text{weakly}^* \text{ in } L^\infty(0,T;L^2), && (2.62)\\
\beta(\rho_h) \;&\to\; \beta(\rho) \quad\ \text{strongly in } L^1_{\rm loc}((0,T)\times\mathbb{R}^d), && (2.63)\\
&\qquad\qquad\quad\ \ \text{weakly}^* \text{ in } L^\infty(0,T;L^2),\\
\nabla\beta(\rho_h) \;&\to\; \nabla\beta(\rho) \quad \text{weakly in } L^2(0,T;(L^2(\mathbb{R}^d))^d). && (2.64)
\end{aligned}$$

Moreover, letting $h \to 0$ in (2.59), we get

$$|\rho(t)|_2^2 + \|\beta(\rho)\|^2_{L^2(0,T;H^1)} \le C_T|\rho_0|_2^2.$$

This yields

$$\Delta\beta(\rho), \operatorname{div}(Db^*(\rho)) \in L^2(0,T;H^{-1}(\mathbb{R}^d)),\ \forall\, T > 0. \tag{2.65}$$

By (1.10), we get

$$\begin{aligned}
&\int_h^T \left(\psi(t), \left(\frac{\rho_h(t)-\rho_h(t-h)}{h}\right)\right)_2 dt + \int_h^T {}_{H^1}\langle \psi(t), A_0(\rho_h(t))\rangle_{H^{-1}}\, dt\\
&\quad - \int_0^h (\psi(t),\rho_h(t))_2 dt + \int_{T-h}^T (\psi(t),\rho_h(t))_2 dt = 0,\ \forall \psi \in C_0^\infty((0,T)\times\mathbb{R}^d).
\end{aligned}$$

Equivalently,

$$\begin{aligned}
-\int_0^{T-h} &\left(\frac{\psi(t+h)-\psi(t)}{h}, \rho_h(t)\right)_2 dt + \int_h^T {}_{H^1}\langle \psi(t) A_0(\rho_h(t))\rangle_{H^{-1}}\, dt\\
&\qquad\qquad - \int_0^h (\psi(t),\rho_h(t))_2 + \int_{T-h}^T (\psi(t),\rho_h(t)) dt = 0.
\end{aligned}$$

Taking into account (2.61), (2.64), we get for $h \to 0$

$$-\int_0^T (\rho(t), \psi'(t))_2 dt + \int_0^T {}_{H^1}\langle \psi(t), A_0(\rho(t))\rangle_{H^{-1}}\, dt = 0,\ \forall\, \psi \in C_0^\infty((0,T)\times\mathbb{R}^d)$$

and, by (2.65), $A_0(\rho) \in L^2(0,T;H^{-1}(\mathbb{R}^d))$. This means that

$$\frac{d\rho}{dt} = \Delta\beta(\rho) - \mathrm{div}(Db^*(\rho)) = -A_0(\rho) \ \text{ in } L^2(0,T;H^{-1}(\mathbb{R}^d)).$$

Hence

$$\frac{d\rho}{dt} \in L^2(0,T;H^{-1}(\mathbb{R}^d)),$$

and, therefore, $\rho : [0,T] \to H^{-1}(\mathbb{R}^d)$ is absolutely continuous and satisfies (2.1)–(2.3). This completes the proof. $\square$

3 The uniqueness of distributional solutions to NFPE

In this section, we shall prove the uniqueness of distributional solutions to (1.1) under the following Hypotheses:

(j) $\beta \in C^1(\mathbb{R}),\ \beta'(r) \geq 0,\ \forall\, r \in \mathbb{R},\ \beta(0) = 0.$

(jj) $D \in L^\infty(\mathbb{R}^d;\mathbb{R}^d)$.

(jjj) $b \in C^1(\mathbb{R})$.

(jv) For each compact $K \subset \mathbb{R}$ there exists $\alpha_K \in (0,\infty)$ such that $|b'(r)r + b(r)| \leq \alpha_K|\beta'(r)|,\ \forall\, r \in \mathbb{R}$.

We note that Hypotheses (j)–(jjj) are weaker than (i)–(iii), while (jv) is equivalent to (iv).

Remark 3.1. If $\beta'(r) > 0,\ \forall\, r \in \mathbb{R}$, then (jv) is automatically fulfilled due to (jjj).

Theorem 3.2. *Let $d \geq 1$, $T > 0$, and let $y_1, y_2 \in L^\infty((0,T)\times\mathbb{R}^d)$ be two distributional solutions to* (1.1) *on $(0,T)\times\mathbb{R}^d$ (in the sense of* (1.24)*) such that $y_1 - y_2 \in L^1((0,T)\times\mathbb{R}^d) \cap L^\infty(0,T;L^2)$ and*

$$\lim_{t\to 0}\ \operatorname*{ess\,sup}_{s\in(0,t)} |(y_1(s) - y_2(s), \varphi)_2| = 0,\ \ \forall \varphi \in C_0^\infty(\mathbb{R}^d). \tag{3.1}$$

Then $y_1 \equiv y_2$.

Proof. Replacing, if necessary, the functions β and b by

$$\beta_N(r) = \begin{cases} \beta(r) & \text{if } |r| \le N, \\ \beta'(N)(r-N) + \beta(N) & \text{if } r > N, \\ \beta'(-N)(r+N) + \beta(-N) & \text{if } r < -N, \end{cases}$$

and

$$b_N(r) = \begin{cases} b(r) & \text{if } |r| \le N, \\ b'(N)(r-N) + b(N) & \text{if } r > N, \\ b'(-N)(r+N) + b(-N) & \text{if } r < -N, \end{cases}$$

where $N \ge \max\{|y_1|_\infty, |y_2|_\infty\}$, we may assume that

$$\beta', b' \in C_b(\mathbb{R}), \tag{3.2}$$

and, therefore, by (j) we have

$$(\beta(r) - \beta(\bar{r}))(r - \bar{r}) \ge \alpha_4 |\beta(r) - \beta(\bar{r})|^2, \quad \forall\, r, \bar{r} \in \mathbb{R}, \tag{3.3}$$

where $\alpha_4 > 0$. We set

$$\begin{gathered} \Phi_\varepsilon(y) = (\varepsilon I - \Delta)^{-1} y, \ \forall\, y \in L^2, \\ z = y_1 - y_2, \ w = \beta(y_1) - \beta(y_2), \ b^*(y_i) \equiv b(y_i) y_i, \ i = 1, 2. \end{gathered} \tag{3.4}$$

It is well known that $\Phi_\varepsilon : L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d)$, $\forall p \in [1, \infty]$ and

$$\varepsilon |\Phi_\varepsilon(y)|_p \le |y|_p, \quad \forall y \in L^p, \ \varepsilon > 0. \tag{3.5}$$

Moreover, $\Phi_\varepsilon(y) \in C_b(\mathbb{R}^d)$ if $y \in L^1 \cap L^\infty$.

By (1.24), we have

$$z_t - \Delta w + \operatorname{div} D(b^*(y_1) - b^*(y_2)) = 0 \text{ in } \mathcal{D}'((0,T) \times \mathbb{R}^d).$$

We set

$$z_\varepsilon = z * \theta_\varepsilon, \ w_\varepsilon = w * \theta_\varepsilon, \ \zeta_\varepsilon = (D(b^*(y_1) - b^*(y_2))) * \theta_\varepsilon,$$

where $\theta \in C_0^\infty(\mathbb{R}^d)$, $\theta_\varepsilon(x) \equiv \varepsilon^{-d} \theta\left(\frac{x}{\varepsilon}\right)$ is a standard mollifier. We note that $z_\varepsilon, w_\varepsilon, \zeta_\varepsilon, \Delta w_\varepsilon, \operatorname{div} \zeta_\varepsilon \in L^2(0,T; L^2)$ and we have

$$(z_\varepsilon)_t - \Delta w_\varepsilon + \operatorname{div} \zeta_\varepsilon = 0 \text{ in } \mathcal{D}'(0,T; L^2). \tag{3.6}$$

This yields $\Phi_\varepsilon(z_\varepsilon), \Phi_\varepsilon(w_\varepsilon), \operatorname{div}\Phi_\varepsilon(\zeta_\varepsilon) \in L^2(0,T;L^2)$ and

$$(\Phi_\varepsilon(z_\varepsilon))_t = \Delta\Phi_\varepsilon(w_\varepsilon) - \operatorname{div}\Phi_\varepsilon(\zeta_\varepsilon) = \varepsilon\Phi_\varepsilon(w_\varepsilon) - w_\varepsilon - \operatorname{div}\Phi_\varepsilon(\zeta_\varepsilon) \\ \text{in } \mathcal{D}'(0,T;L^2). \tag{3.7}$$

By (3.6), (3.7) it follows that $(z_\varepsilon)_t = \frac{d}{dt}\, z_\varepsilon$, $(\Phi_\varepsilon(z))_t = \frac{d}{dt}\,\Phi_\varepsilon(z_\varepsilon) \in L^2(0,T;L^2)$, where $\frac{d}{dt}$ is taken in the sense of L^2-valued vectorial distributions on $(0,T)$. This implies that $z_\varepsilon, \Phi_\varepsilon(z_\varepsilon) \in H^1(0,T;L^2)$ and both $[0,T] \ni t \mapsto z_\varepsilon(t) \in L^2$ and $[0,T] \ni t \to \Phi_\varepsilon(z_\varepsilon(t)) \in L^2$ is absolutely continuous. (See, e.g., [1], p. 23.) As a matter of fact, it follows by (3.5) and (3.7) that

$$\Phi_\varepsilon(z_\varepsilon), \Phi_\varepsilon(w_\varepsilon) \in L^2((0,T);C_b(\mathbb{R}^d)\cap L^2). \tag{3.8}$$

We set $h_\varepsilon(t) = (\Phi_\varepsilon(z_\varepsilon(t)), z_\varepsilon(t))_2$ and get, therefore,

$$\begin{aligned} h'_\varepsilon(t) &= 2(z_\varepsilon(t), (\Phi_\varepsilon(z_\varepsilon(t)))_t)_2 \\ &= 2(\varepsilon\Phi_\varepsilon(w_\varepsilon(t)) - w_\varepsilon(t) - \operatorname{div}\Phi_\varepsilon(\zeta_\varepsilon(t)), z_\varepsilon(t))_2 \\ &= 2\varepsilon(\Phi_\varepsilon(z_\varepsilon(t)), w_\varepsilon(t))_2 + 2(\nabla\Phi_\varepsilon(z_\varepsilon(t)), \zeta_\varepsilon(t))_2 \\ &\quad -2(z_\varepsilon(t), w_\varepsilon(t))_2, \ \text{a.e. } t \in (0,T), \end{aligned} \tag{3.9}$$

where, $(\cdot,\cdot)_2$ is the scalar product in L^2. By (3.7)–(3.9) it follows that $t \to h_\varepsilon(t)$ is absolutely continuous on $[0,T]$.

Since, by (3.3), (3.4),

$$(z_\varepsilon(t), w_\varepsilon(t))_2 \ge \alpha_4 |\,|w(t)| * \theta_\varepsilon|_2^2 + \gamma_\varepsilon(t), \tag{3.10}$$

where

$$\gamma_\varepsilon(t) := (z_\varepsilon(t), w_\varepsilon(t))_2 - (z(t), w(t))_2, \tag{3.11}$$

we get, therefore, by (3.3), (3.9) and (jv) that, for $\alpha_3 := \alpha_{[-N,N]}$,

$$\begin{aligned} 0 \le h_\varepsilon(t) &\le h_\varepsilon(0+) + 2\varepsilon\int_0^t (\Phi_\varepsilon(z_\varepsilon(s)), w_\varepsilon(s))_2 ds - 2\alpha_4 \int_0^t |\,|w(s)| * \theta_\varepsilon|_2^2 ds \\ &+2\alpha_3|D|_\infty \int_0^t |\nabla\Phi_\varepsilon(z_\varepsilon(s))|_2 ||w(s)| * \theta_\varepsilon|_2 ds + 2\int_0^t |\gamma_\varepsilon(s)| ds,\ \forall t \in [0,T]. \end{aligned} \tag{3.12}$$

Moreover, since $z \in L^\infty((0,T)\times\mathbb{R}^d)$ and by (3.5) we obtain

$$\varepsilon|\Phi_\varepsilon(z_\varepsilon(t))|_\infty \le |z_\varepsilon(t)|_\infty \le |z(t)|_\infty, \quad \text{a.e. } t \in (0,T). \tag{3.13}$$

Taking into account that $t \to \Phi_\varepsilon(z_\varepsilon(t))$ is $L^2(\mathbb{R}^d)$ continuous on $[0,T]$, there exists $f \in L^2$ such that

$$\lim_{t\to 0} \Phi_\varepsilon(z_\varepsilon(t)) = f \ \text{ in } L^2.$$

Furthermore, for every $\varphi \in C_0^\infty(\mathbb{R}^d)$, $s \in (0,T)$,

$$|h_\varepsilon(s) \le |\Phi_\varepsilon(z_\varepsilon(s)) - f|_2 |z_\varepsilon(s)|_2 + |f-\varphi|_2 |z_\varepsilon(s)|_2 + |(\varphi * \theta_\varepsilon, z(s))_2|.$$

Hence, by (3.1),

$$\begin{aligned}|h_\varepsilon(0+)| &= \lim_{t\downarrow 0} |h_\varepsilon(t)| = \lim_{t\to 0} \operatorname*{ess\,sup}_{s\in(0,t)} |h_\varepsilon(s)| \\ &\le \left(\lim_{t\to 0} |\Phi_\varepsilon(z_\varepsilon(t)) - f|_2 + |f-\varphi|_2\right) |z_\varepsilon|_{L^\infty(0,T;L^2)} \\ &+ \lim_{t\to 0} \operatorname*{ess\,sup}_{s\in(0,t)} |(\varphi * \theta_\varepsilon, z(s))_2| = |f-\varphi|_2 |z_\varepsilon|_{L^\infty(0,T;L^2)}.\end{aligned}$$

Since $C_0^\infty(\mathbb{R}^d)$ is dense in $L^2(\mathbb{R}^d)$, we find

$$h_\varepsilon(0+) = 0. \tag{3.14}$$

On the other hand, taking into account that, for a.e. $t \in (0,T)$,

$$\varepsilon\Phi_\varepsilon(z_\varepsilon(t)) - \Delta\Phi_\varepsilon(z_\varepsilon(t)) = z_\varepsilon(t), \tag{3.15}$$

we get that

$$\varepsilon|\Phi_\varepsilon(z_\varepsilon(t))|_2^2 + |\nabla\Phi_\varepsilon(z_\varepsilon(t))|_2^2 = (z_\varepsilon(t), \Phi_\varepsilon(z_\varepsilon(t)))_2, \text{ for a.e. } t \in (0,T). \tag{3.16}$$

We also have by (3.13)

$$\begin{aligned}\varepsilon|(\Phi_\varepsilon(z_\varepsilon(t), w_\varepsilon(t)))_2| \le \varepsilon|\Phi_\varepsilon(z_\varepsilon(t))|_\infty |w_\varepsilon(t)|_1 \le |z(t)|_\infty |w(t)|_1, \\ \text{for a.e. } t \in (0,T),\end{aligned} \tag{3.17}$$

while

$$\lim_{\varepsilon\to 0} \varepsilon(\Phi_\varepsilon(z_\varepsilon(s)), w_\varepsilon(s))_2 = 0, \quad \text{a.e. } s \in (0,T), \tag{3.18}$$

To prove (3.18), one proceeds as in the proof of [14, Lemma 1]. Namely, we may write, for a.e. $t \in (0,T)$,

$$\varepsilon\Phi_\varepsilon(z_\varepsilon(t))(x) = \varepsilon^{\frac{d}{2}} \int_{\mathbb{R}^d} K(\sqrt{\varepsilon}\,(x-\xi)) z_\varepsilon(t)(\xi) d\xi, \ x \in \mathbb{R}^d,$$

where K is the probability kernel associated with the operator $(I-\Delta)^{-1}$, that is (see [25]),

$$\Phi_1(y)(x) = \int_{\mathbb{R}^d} K(x-\xi) y(\xi) d\xi, \ \forall x \in \mathbb{R}^d.$$

This yields for a.e. $x \in \mathbb{R}^d$

$$|\varepsilon\Phi_\varepsilon(z_\varepsilon(t))(x)| \le C_r \varepsilon^{\frac{d}{2}} |z_\varepsilon(t)|_1 + \varepsilon^{\frac{d}{2}} |z_\varepsilon(t)|_\infty \int_{[\sqrt{\varepsilon}|x-\xi| \le r]} K(\sqrt{\varepsilon}(x-\xi)) d\xi, \ \forall r > 0.$$

Since $C_r = \sup\{K(x);\ |x| \geq r\} < \infty$, $\forall r > 0$, we get

$$\limsup_{\varepsilon \downarrow 0} (\varepsilon(|\Phi_\varepsilon(z_\varepsilon(t))|_\infty) \leq |z(t)|_\infty \int_{[|\xi| \leq r]} K(\xi) d\xi,\ \forall r > 0, \quad \text{for a.e. } t \in (0,T). \tag{3.19}$$

Since $K \in L^1(\mathbb{R}^d)$, letting $r \to 0$ in (3.19) it follows that

$$\limsup_{\varepsilon \to 0} (\varepsilon |\Phi_\varepsilon(z_\varepsilon(t))|_\infty) = 0,\ \text{ a.e. } t \in (0,T), \tag{3.20}$$

which implies (3.18), as claimed.

By (3.17)–(3.18) and the dominated convergence theorem, it follows that

$$\lim_{\varepsilon \to 0} \varepsilon \int_0^t (\Phi_\varepsilon(z_\varepsilon(s)), w_\varepsilon(s))_2 ds = 0,\ t \in [0,T]. \tag{3.21}$$

On the other hand, since $\lim\limits_{\varepsilon \to 0} \gamma_\varepsilon = 0$, a.e. on $[0,T]$ and $|\gamma(t)| \leq 2|z(t)|_\infty |w(t)|_1$ for a.e. $t \in (0,T)$, we have by (3.12)–(3.21) that

$$|\nabla \Phi_\varepsilon(z_\varepsilon(t))|_2^2 \leq \eta_\varepsilon(t) + \alpha_5 |D|_\infty^2 \int_0^t |\nabla \Phi_\varepsilon(z_\varepsilon(s))|_2^2 ds, \quad \text{for a.e. } t \in (0,T), \tag{3.22}$$

where $\alpha_5 > 0$ and $\lim\limits_{\varepsilon \to 0} \eta_\varepsilon(t) = 0$ for all $t \in [0,T]$.

In particular, by (3.22), it follows that

$$|\nabla \Phi_\varepsilon(z_\varepsilon(t))|_2^2 \leq \eta_\varepsilon(t) \exp(\alpha_6 |D|_\infty^2 t),\ \text{for a.e. } t \in [0,T]. \tag{3.23}$$

This implies that $\nabla \Phi_\varepsilon(z_\varepsilon(t)) \to 0$ in $\mathcal{D}'(\mathbb{R}^d)$ as $\varepsilon \to 0$ for a.e. $t \in (0,T)$, hence by (3.20), so does the left hand side of (3.15). Thus, $0 = \lim\limits_{\varepsilon \to 0} z_\varepsilon(t) = z(t)$ in $\mathcal{D}'(\mathbb{R}^d)$ for a.e. $t \in (0,T)$, which implies $y_1 \equiv y_2$. $\square$

Remark 3.3. Let $\rho \in L^1((0,T) \times \mathbb{R}^d)$ such that $\beta(\rho) \in L^1((0,T) \times \mathbb{R}^d)$ and ρ is a solution to (1.24). Then, it is elementary to check that

$$\int_{\mathbb{R}^d} \rho(t,x) dx = \int_{\mathbb{R}^d} \rho_0(dx)\ \text{ for } dt\text{-a.e. } t \in (0,T).$$

Hence, if ρ_0 is nonnegative and $\rho \geq 0$, a.e. on $(0,T) \times \mathbb{R}^d$, it follows by Lemma 2.3 in [24] that there exists a $dt \otimes dt$-version $\widetilde{\rho}$ of ρ such that $[0,T] \ni t \mapsto \widetilde{\rho}(t,x)dx$ is narrowly continuous and $\widetilde{\rho}(0,x)dx = \rho_0(dx)$.

Remark 3.3 then implies the following consequence of Theorem 3.2.

Corollary 3.4. *Let $\rho_0 \in \mathcal{M}(\mathbb{R}^d)$, ρ_0 nonnegative, and $y_1, y_2 \in L^\infty((0,T)\times\mathbb{R}^d)$ $\cap L^1((0,T)\times\mathbb{R}^d)$ be two nonnegative solutions to* (1.24)*. Then, $y_1 \equiv y_2$.*

Proof. It follows by Remark 3.3 that $y_1, y_2 \in L^\infty(0,T;L^1 \cap L^\infty)$, hence $y_1, y_2 \in L^\infty(0,T;L^p)$ for all $p \in [1,\infty]$, in particular $y_1 - y_2 \in L^\infty(0,T;L^2)$. Furthermore, let $\widetilde{y}_1, \widetilde{y}_2$ be the $dt \otimes dx$-versions from Remark 3.3. Then, for every $\varphi \in C_0^\infty(\mathbb{R}^d)$,

$$\lim_{t\to 0} \operatorname*{ess\,sup}_{s\in(0,t)} |(y_1(s) - y_2(s), \varphi)_2| = \lim_{t\to 0} \left| \int_{\mathbb{R}} (\widetilde{y}_1(s,x) - \widetilde{y}_2(s,x))\varphi(x)dx \right| = 0.$$

So, (3.1) holds and Theorem 3.2 implies the assertion. □

Remark 3.5. Theorem 5.2 in [5] asserts that, under the assumptions (k), (kk), (kkk) and (6.3) on β, b and D, specified in [5, Section 4], there exists a distributional solution to (1.24) above for every $\rho_0 \in \mathcal{M}(\mathbb{R}^d)$. If (in the notation of Remark 3.3) $S(t)\rho_0 := \widetilde{\rho}(t,x)dx$, $t \geq 0$, denotes the narrowly continuous solution to (1.24), it was left as an open question in [5] whether the semigroup property

$$S(t+s)\rho_0 = S(t)S(s)\rho_0;\ t,s > 0, \tag{3.24}$$

holds (see Remark 5.3 in [5]). If, in addition to the above mentioned assumptions from [5] on β, b and D, we assume that our condition (jv) holds, then Corollary 3.4 implies that (3.24) holds. Indeed, for $s > 0$ fixed, both functions in $t \geq 0$ from the left and right hand sides of (3.24) are in $L^\infty((0,T)\times\mathbb{R}^d) \cap L^\infty(0,T;L^1)$ by [5, Theorem 5.2, (5.6), (5.7)] and [5, Theorem 2.2, (2.6)], respectively, and both solve (1.24). Hence, Corollary 3.4 implies that they are equal.

4 Weak uniqeness of the corresponding McKean–Vlasov SDE

Our next aim is to prove weak uniqueness of the McKean–Vlasov SDE (1.4) corresponding to the Fokker–Planck equation (1.1) (resp. (1.24)). Here, we allow the law of X_0 to be a general probability measure ρ_0, not necessarily absolutely continuous with respect to Lebesgue measure. By Corollary 3.4, it is immediate to prove that the time-marginal law densities of all weak solutions to (1.4) coincide, provided they are bounded. To prove that also their laws on path space coincide, requires to also prove the so-called "linearized uniqueness". This was already noted in [6], [7].

Theorem 4.1. (Linearized Uniqueness) *Assume that Hypotheses* (j)–(jv) *hold. Let $T>0$ and $u\in L^\infty((0,T)\times\mathbb{R}^d)$. Let $y_1,y_2\in L^\infty((0,T)\times\mathbb{R}^d)$ such that $y_1-y_2\in L^1((0,T)\times\mathbb{R}^d)\cap L^\infty(0,T;L^2)$ and y_1,y_2 are solutions to the following linearized version of* (1.24)

$$\begin{aligned}\int_0^\infty\int_{\mathbb{R}^d}\left(\varphi_t+\frac{\beta(u)}{u}\Delta\varphi+b(u)D\cdot\nabla\varphi\right)\rho\,dx\,dt+\int_{\mathbb{R}^d}\varphi(0,x)\rho_0(dx)=0,\\ \forall\varphi\in C_0^\infty([0,T)\times\mathbb{R}^d),\end{aligned}\tag{4.1}$$

for some $\rho_0\in\mathcal{M}(\mathbb{R}^d)$, where $\frac{\beta(0)}{0}:=\beta'(0)$, such that (3.1) *holds. Then, $y_1\equiv y_2$.*

Proof. First, we note that by (j)–(jv) we have:

$$\frac{\beta(u)}{u},b(u)\in L^\infty((0,T)\times\mathbb{R}^d),\tag{4.2}$$

$$|Db(u)|\le\alpha_7\,\frac{\beta(u)}{u},\ \text{ a.e. on }(0,T)\times\mathbb{R}^d,\tag{4.3}$$

where $\alpha_7:=|D|_\infty\alpha_{[-|u|_\infty,|u|_\infty]}$. Now, we set

$$z:=y_1-y_2,\quad w:=\frac{\beta(u)}{u}\,(y_1-y_2).\tag{4.4}$$

Then, we have, since $\frac{\beta(u)}{u}\ge 0$,

$$wz=\frac{\beta(u)}{u}\,|y_1-y_2|^2\ge\left(\left|\frac{\beta(u)}{u}\right|_\infty+1\right)^{-1}|w|^2,\ \text{ a.e. on }(0,T)\times\mathbb{R}^d,\tag{4.5}$$

and

$$|Db(u)z|\le\alpha_7|w|.\tag{4.6}$$

We set

$$z_\varepsilon=z*\theta_\varepsilon,\ w_\varepsilon=w*\theta_\varepsilon,\ \zeta_\varepsilon=(Db(u)(y_1-y_2))*\theta_\varepsilon,\tag{4.7}$$

where θ_ε is as in the proof of Theorem 3.2. Now, by (4.2)–(4.6), we can repeat the proof of the latter line by line for z_ε, w_ε and ζ_ε in (4.7) to obtain $y_1\equiv y_2$. □

Corollary 4.2. *Let $\rho_0 \in \mathcal{M}(\mathbb{R}^d)$, $\rho_0 \geq 0$, and $y_1, y_2 \in (L^\infty \cap L^1)((0,T)\times\mathbb{R}^d)$ be two nonnegative solutions to* (4.1)*. Then, $y_1 \equiv y_2$.*

Proof. The assertion follows from Theorem 4.1 by analogous arguments as in the proof of Corollary 3.4. □

Theorem 4.3. *Assume that Hypotheses* (j)–(jv) *hold and let $T > 0$. Let $X(t)$, $t \geq 0$, and $\widetilde{X}(t)$, $t \geq 0$, be two solutions to* (1.4) *such that, for*

$$u(t,\cdot) := \frac{d\mathcal{L}_{X(t)}}{dx}, \quad \widetilde{u}(t,\cdot) := \frac{d\mathcal{L}_{\widetilde{X}(t)}}{dx},$$

we have

$$u, \widetilde{u} \in L^\infty((0,T)\times\mathbb{R}^d). \tag{4.8}$$

Then, X and $\overline{X}$ have the same laws, i.e., $\mathbb{P}\circ X^{-1} = \mathbb{P}\circ \widetilde{X}^{\,-1}$.

Proof. We first note that, by narrow continuity,

$$u(0,x)dx = \rho_0(dx) = (\mathbb{P}\circ X(0)^{-1})(dx) = (\mathbb{P}\circ \widetilde{X}(0)^{-1})(dx).$$

Furthermore, again by narrow continuity, (4.8) implies that $u(t,\cdot), \widetilde{u}(t,\cdot) \in L^\infty(\mathbb{R}^d)$, for all $t \in (0,T]$. Furthermore, by Itô's formula, both u and $\widetilde{u}$ satisfy the weak (nonlinear) Fokker–Planck equation (1.24). Hence, since $u, \widetilde{u} \in L^\infty(0,T;L^1) \subset L^1((0,T)\times\mathbb{R}^d)$, Corollary 3.4 implies $u \equiv \widetilde{u}$. Furthermore, again by Itô's formula, $\mathbb{P}\circ X^{-1}$ and $\mathbb{P}\circ\widetilde{X}^{-1}$ satisfy the martingale problem with the initial condition ρ_0 for the linear Kolmogorov operator

$$L_u := \frac{\beta(u)}{u}\Delta + b(u)D\cdot\nabla.$$

Hence, by Corollary 4.2, the assertion follows from Lemma 2.12 in [27].

Here, for $s \in [0,T]$, the set $\mathcal{R}_{[s,T]}$, which appears in that lemma, is chosen to be the set of all narrowly continuous, probability measure-valued solutions of (4.1) having, for each $t \in [s,T]$, $t > 0$, a density $u(t,\cdot) \in L^\infty(\mathbb{R}^d)$ such that $v \in L^\infty((0,R)\times\mathbb{R}^d)$. □

Remark 4.4. For existence of weak solutions to (1.4) with time marginals in $L^\infty((0,T)\times\mathbb{R}^d)$ given by a solution to (1.24) for initial conditions in $L^1\cap L^\infty$, we refer to [5, Theorem 6.1(a)] in the much improved recent arXiv version. In particular, the conditions stated there are weaker than conditions (i)–(iv) from Section 1 of the present paper. But, as seen, under the latter strong conditions we have even mild solutions ρ to (1.1) (and hence solutions to (1.24)) with initial conditions ρ_0 merely in L^1, as follows from Theorem 2.1, and we also have $\beta(\rho) \in L^1((0,T) \times \mathbb{R}^d)$. Hence, by the general results in [4, Section 2], we get a weak solution to (1.4) with time-marginals $\rho(t,x)dx$, $t \in [0,T]$. Furthermore, additionally assuming that (v) holds, by Theorem 2.2 we even have that this mild solution ρ is a strong solution to (1.1) in H^{-1} in the sense of (2.3) if $\rho_0 \in L^1 \cap L^2$.

5 Weak differentiability of the nonlinear Fokker–Planck flow

Though the continuous semigroup $S(t) : L^1 \to L^1$, $t \geq 0$, defined by Theorem 2.1 is not differentiable on $(0,\infty)$, it is however differentiable in the distribution space $H^{-1} = H^{-1}(\mathbb{R}^d)$. Namely, consider in the space H^{-1} the operator $A_1 : D(A_1) \subset H^{-1} \to H^{-1}$,

$$\begin{aligned} A_1(y) &= -\Delta\beta(y) + \operatorname{div}(Db(y)y), \quad \forall\, y \in D(A_1), \\ D(A_1) &= \{y \in L^2; \beta(y) \in H^1(\mathbb{R}^d)\}. \end{aligned}$$

(We note that $D(A_1)$ is dense in H^{-1}.)

We shall assume here that besides (i)–(iv) the following hypothesis holds.

(vi) $\beta \in \operatorname{Lip}(\mathbb{R})$.

Lemma 5.1. *Let $d \geq 1$. Then, under Hypotheses* (i)–(v)*, the operator A_1 is quasi-m-accretive in H^{-1}, that is, $\omega I + A_1$ is m-accretive in H^{-1} for some $\omega \geq 0$.*

Proof. One must prove that

$$R(I + \lambda A_1) = H^{-1}, \quad \forall\, \lambda \in (0, \omega^{-1}), \tag{5.1}$$

$$\|(I+\lambda A_1)^{-1}u - (I+\lambda A_1)^{-1}v\|_{H^{-1}} \leq (1-\lambda\omega)^{-1}\|u-v\|_{H^{-1}},\ \forall\, u,v \in H^{-1}. \tag{5.2}$$

To prove (5.1), consider for $f \in H^{-1}$ the equation

$$y - \lambda\Delta\beta(y) + \lambda \operatorname{div}(Db(y)y) = f \text{ in } \mathcal{D}'(\mathbb{R}^d), \tag{5.3}$$

and approximate it by (2.44). Arguing as in the proof of Lemma 2.4, we get for the solution y_ε to (2.44) the estimate.

Namely,

$$|y_\varepsilon|_2^2+\lambda|\nabla\beta(y_\varepsilon)|_2^2|+\varepsilon|\nabla y_\varepsilon|_2^2\le C\|f\|^2_{H^{-1}},\quad \forall\varepsilon>0,\ 0<\lambda<\lambda_0=\omega^{-1}.$$

Then, letting $\varepsilon\to 0$, we get as above that there is the limit

$$y=\lim_{\varepsilon\to 0}y_\varepsilon\ \text{ strongly in } L^2_{\rm loc} \text{ and weakly in } H^{-1},$$

and

$$y+\lambda A_1(y)=f,$$
$$|y|_2^2+|\nabla\beta(y)|_2^2\le C\|f\|^2_{H^{-1}}.$$

Moreover, subtracting equations (5.3) for two solutions $y_1,y_2\in D(A_1)$ corresponding to $f_1,f_2\in H^{-1}$, multiplying by $(I-\Delta)^{-1}(y_1-y_2)$ and integrating on $\mathbb{R}^d$, we get

$$\begin{aligned}&\|y_1-y_2\|^2_{H^{-1}}+\lambda(\beta(y_1)-\beta(y_2),y_1-y_2)_2\\&\qquad\le\|f_1-f_2\|_{H^{-1}}\|y_1-y_2\|_{H^{-1}}+\lambda\|\beta\|_{\rm Lip}|y_1-y_2|_2\|y_1-y_2\|_{H^{-1}}\\&\qquad+\lambda|D|_\infty|b^*(y_1)-b^*(y_2)|_2\|y_1-y_2\|_{H^{-1}}.\end{aligned}$$

This yields

$$\|y_1-y_2\|_{H^{-1}}\le(1-\lambda\omega)^{-1}\|f_1-f_2\|_{H^{-1}},\ \forall\,\lambda\in(0,\omega^{-1}),$$

and so (5.2) follows.

By Lemma 5.1, we infer that the operator $-A_1$ generates a continuous quasi-contractive semigroup in H^{-1}. More precisely, we have (see [1], p. 143)

Theorem 5.2. *Assume that Hypotheses* (i)–(iv) *hold. Then, for each $y_0\in D(A_1)$, there is a unique absolutely continuous function $y:[0,\infty)\to H^{-1}(\mathbb{R}^d)$ t-differentiable from the right, such that*

$$\begin{aligned}&\frac{d}{dt}\,y(t)+A_1(y(t))=0,\quad a.e.\ t>0,\\&y(0)=y_0,\end{aligned}\tag{5.4}$$

and

$$\frac{d^+}{dt}\,y(t)+A_1(y(t))=0,\quad\forall\,t\ge 0.\tag{5.5}$$

Moreover, $S_1(t)y_0\equiv y(t)$ satisfies

$$S_1(t+s)y_0 = S_1(t)S_1(s)y_0,\ \forall\, t,s \ge 0,\ y_0 \in D(A_1), \tag{5.6}$$

$$\|S_1(t)y_0 - S_1(t)\bar{y}_0\|_{H^{-1}} \le \exp(\omega t)\|y_0 - \bar{y}_0\|_{H^{-1}},\ \forall t \ge 0,\ y_0, \bar{y}_0 \in D(A_1). \tag{5.7}$$

$$S_1(t)y_0 = \lim_{n\to\infty} \left(I + \frac{t}{n}\, A_1\right)^{-n} y_0 \ \ in\ H^{-1},\ \forall\, t \ge 0. \tag{5.8}$$

The semigroup $\{S_1(t)\}_{t\ge0}$ extends by density on the closure $\overline{D(A_1)} = H^{-1}$ of $D(A_1)$ in H^{-1} and by (5.5) it follows that $-A_1$ is the infinitesimal generator of $S_1(t)$ in H^{-1}. Since $S(t)$ and $S_1(t)$ are given by the same exponential formula (1.18) (respectively (5.8)) on L^1 (respectively H^{-1}), we infer that

$$S(t)\rho_0 = S_1(t)\rho_0,\ \forall\, t \ge 0,\ \rho_0 \in L^1 \cap H^{-1}.$$

This implies that the *Fokker–Planck semigroup* $S(t)$ *given by Theorem* 2.1 *is* H^{-1} *differentiable from the right in* t *on* $L^1 \cap H^{-1}$.

6 Fokker–Planck equations with x-dependent coefficients

Theorems 2.1, 2.2, as well as Theorem 3.2, extend *mutatis-mutandis* to nonlinear Fokker–Planck equations of the form

$$\begin{aligned} &\rho_t(t,x) - \Delta\beta(x,\rho(t,x)) + \operatorname{div}(D(x)b(\rho(t,x))\rho(t,x)) = 0, \\ &\hspace{16em} t \ge 0,\ x \in \mathbb{R}^d, \\ &\rho(0,x) = \rho_0(x),\ x \in \mathbb{R}^d, \end{aligned} \tag{6.1}$$

where $\beta : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R},\ b : \mathbb{R} \to \mathbb{R},\ D : \mathbb{R}^d \to \mathbb{R}$ satisfy (ii), (iii) and

(i)′ $\beta \in C^2(\mathbb{R}^d \times \mathbb{R}),\ \beta_r(x,r) > 0,\ \forall r \ne 0,\ x \in \mathbb{R}^d,\ \beta(x,0) \equiv 0$ and

$$h_N(x) = \sup\{|\Delta_x\beta(x,r)|;\ |r| \le N\} \in L^1(\mathbb{R}^d),\ \forall N > 0, \tag{6.2}$$

$$|\beta(x,r)| \le \alpha_2|r|,\ \forall x \in \mathbb{R}^d,\ r \in \mathbb{R}. \tag{6.3}$$

(iv)′ $$|b^*(r) - b^*(\bar{r})| \le \alpha_2|\beta(x,r) - \beta(x,\bar{r})|,\ \forall x \in \mathbb{R}^d,\ r,\bar{r} \in \mathbb{R}. \tag{6.4}$$

The definition of the mild solution ρ to (6.1) is that given in (1.6)–(1.9). We have

Theorem 6.1. *Under Hypotheses* (i)$'$, (ii), (iii), (iv)$'$, *for each* $\rho_0 \in L^1$ *there is a unique mild solution* ρ *to* (6.1) *which has all properties mentioned in Theorem* 2.1.

The proof is exactly the same as in the previous case and relies on Proposition 2.3, where $A_0 : L^1 \to L^1$ is the operator

$$\begin{aligned} A_0(y) &= -\Delta\beta(x,y) + \operatorname{div}(D(x), b(y)y), \\ D(A_0) &= \{y \in L^1; -\Delta\beta(x,y) + \operatorname{div}(D(x)b(y)y) \in L^1\}. \end{aligned}$$

The extension of Theorem 2.2 to (6.1) is also immediate, and so we omit the details. As regards the uniqueness of a distributional solution to equation (6.1), instead of (j) the following hypothesis will be assumed:

(j)$'$ $\beta, \beta_r \in C(\mathbb{R}^d \times \mathbb{R}), \beta_r \in L^\infty(\mathbb{R}^d \times (-N,N))$, $\forall N > 0$, $\beta_r(x,r) \geq 0$, $\forall x \in \mathbb{R}^d$, $r \in \mathbb{R}$, $\beta(x,0) \equiv 0$.

Then, we obtain results analogous to Theorem 3.2 and Corollary 3.4.

The proofs are identical because also in this case one may assume that (see (3.2))

$$\beta \in C_b(\mathbb{R}^d \times \mathbb{R});\ b \in C_b(\mathbb{R}),\ \beta_r \in L^\infty(\mathbb{R}^d \times \mathbb{R}).$$

We also note that Theorem 5.2 remains valid in this case if, besides Hypotheses (i)$'$, (ii), (iii), (iv)$'$, one assume that

$$\beta_r \in L^\infty(\mathbb{R}^d \times \mathbb{R}).$$

Then, it follows as in Lemma 5.1 that (5.1)–(5.2) hold, and so, the operator A_1 generates in $H^{-1}(\mathbb{R}^d)$ a semigroup $S_1(t)$ satisfying (5.4)–(5.8).

As in Section 4, these results imply weak uniqueness results for probabilistically weak solutions to the McKean–Vlasov equation

$$\begin{aligned} &dX(t) = D(X(t))b(\rho(t,X(t)))dt + \sqrt{\frac{2\beta(X(t),\rho(t,X(t)))}{\rho(t,X(t))}}\, dW(t), \\ &X(0) = X_0. \end{aligned}$$

Details on this will be given a forthcoming survey paper which is in preparation.

Acknowledgement. The authors are indebted to Haim Brezis for useful suggestions regarding the proof of estimates (2.9).
This work was supported by the DFG through CRC 1283 and by UEFISCDI (Romanian) grant through PN-III-ID-PCE 2021-3.

References

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

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

[3] Barbu, V., Röckner, M., Probabilistic representation for solutions to nonlinear FP equations, *SIAM J. Math. Anal.*, 50 (2018), 2588-2607.

[4] Barbu, V., Röckner, M., From Fokker–Planck equations to solutions of distribution dependent, *SDE, Annals of Probability*, 48 (2020), 1902-1920.

[5] Barbu, V., Röckner, M., Solutions for nonlinear Fokker–Planck equations with measures as initial data and McKean-Vlasov equations, *J. Functional Anal.*, 280 (7) (2021), 1-35.

[6] Barbu, V., Röckner, M., Uniqueness for nonlinear Fokker-Planck equations and weak uniqueness for McKean-Vlasov SDEs, *Stoch. PDE Anal. Comput.*, 9 (2021), 702-713.

[7] Barbu, V., Röckner, M., Stochastic PDEs, Correction to: Uniqueness for nonlinear Fokker–Planck equations and weak uniqueness for McKean–Vlasov SDEs, *Stoch. PDE: Anal. Comp.* https://doi.org/10.1007/S40072-020-0018-8.

[8] Barbu, V., Röckner, M., *The evolution to equilibrium of solutions to nonlinear Fokker-Planck equations*, arXiv:1904.08291, *Indiana University Journal* (to appear, 2022).

[9] Barbu, V., Röckner, M., Nonlinear Fokker–Planck equations with time-dependent coefficients, (submitted).

[10] Bauer, H., *Measure and Integration Theory*, de Gruyter, 2001.

[11] Belaribi, N., Russo, F., Uniqueness for Fokker–Planck equations with measurable coefficients and applications to the fast diffusion equations, *Electron. J. Probab.*, 17 (2012), 1-28.

[12] Benilan, Ph., Brezis, H., Crandall, M.G., A semilinear elliptic equation in $L^1(\mathbb{R}^N)$, *Ann. Scuola Norm. Sup. Pisa*, Serie IV, vol. II (1975), 523-555.

[13] Brezis, H., *Functional Analysis, Sobolev Spaces and Partial differential equations*, Springer, 2011.

[14] 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.

[15] Brezis, H., Strauss, W., Semilinear elliptic equations in L^1, *J. Math. Soc. Japan*, 25 (1973), 565-590.

[16] Calderon, A.P., Zygmund, A., On singular integrals, *American J. Math.*, 78 (1956), 289-309.

[17] Carillo, J.A., Gualdani, M.P., Jürgel, A., Convergence of an entropic semi-discretization for nonlinear Fokker–Planck equations in $\mathbb{R}^d$, *Publicaciones Mathematique,* 25 (2008), 413-433.

[18] Crandall, M.G., The semigroup approach to first order quasilinear equations in several space variables, *Israel J. Math.*, 12 (1972), 108-132.

[19] Diaz, J.L., Galiano, G., Jürgel, A., On a quasilinear degenerate system arising in semiconductor theory. Existence and uniqueness of solutions, *Nonlinear Anal. Real World Appl.*, 2 (2001), 305-331.

[20] Frank, T.D., *Nonlinear Fokker–Planck Equations*, Springer, Berlin. Heidelberg. New York, 2005.

[21] Lasry, J.M., Lions, P.L., Mean field games, *Japan J. Math.*, 2 (2007), 229-260.

[22] Porretta, A., Weak solutions to Fokker–Planck equations and mean field games, *Archive Rat. Mech. Anal.*, 216 (2015), 1-62.

[23] Pierre, M., Uniqueness of the solutions of $u_t - \Delta\varphi(u) = 0$ with initial data measure, *Nonlinear Anal. Theory Methods Appl.*, 6 (2) (1982), 175-187.

[24] Rehmeier, M., Flow selections for (nonlinear) Fokker–Planck–Kolmogorov equations, arXiv:2201.04539v1.

[25] Stein, E., Weiss, G., *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, Princeton, N.J., 1971.

[26] Schwammle, V., Nobre, F.D., Curado, F.M.F., Consequences of the H-theorem from nonlinear Fokker–Planck equations, *Phys. Review*, E 76 (2007), 041123.

[27] Trevisan, D., Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients, *Electron J. Probab.,* 21 (2016), 22-41.