Nonlocal, nonlinear Fokker–Planck equations and nonlinear martingale problems

Viorel Barbu* José Luís da Silva† Michael Röckner‡§

Abstract

This work is concerned with the existence of mild solutions and the uniqueness of distributional solutions to nonlinear Fokker–Planck equations with nonlocal operators $\Psi(-\Delta)$, where Ψ is a Bernstein function. As applications, the existence and uniqueness of solutions to the corresponding nonlinear martingale problems are proved. Furthermore, it is shown that these solutions form a nonlinear Markov process in the sense of McKean such that their one-dimensional time marginal law densities are the solutions to the nonlocal nonlinear Fokker–Planck equation. Hence, McKean's program envisioned in his PNAS paper from 1966 is realized for these nonlocal PDEs.

MSC: 60H15, 47H05, 47J05.

Keywords: Fokker–Planck equation, Bernstein function, distributional solution, mild solution, nonlinear martingale problem.

1 Introduction

We are concerned here with the following nonlinear, nonlocal Fokker–Planck equation of Nemytskii-type

$$\begin{aligned} &u_t + \Psi(-\Delta)(\beta(u)) + \operatorname{div}(Db(u)u) = 0, \quad \text{on } (0,\infty)\times\mathbb{R}^d, \\ &u(0,x) = u_0(x),\ x\in\mathbb{R}^d, \end{aligned} \tag{1.1}$$

*Al.I. Cuza University and Octav Mayer Institute of Mathematics of Romanian Academy, Iaşi, Romania. Email: vbarbu41@gmail.com

†CIMA – Faculty of Exact Sciences and Engineering, University of Madeira, 9020-105 Funchal, Portugal. Email: joses@staff.uma.pt

‡Fakultät für Mathematik, Universität Bielefeld, D-33501 Bielefeld, Germany. Email: roeckner@math.uni-bielefeld.de

§Academy of Mathematics and System Sciences, CAS, Beijing, China.

where the functions $\beta : \mathbb{R} \to \mathbb{R}$, $D : \mathbb{R}^d \to \mathbb{R}^d$, $d \geq 2$, $b : \mathbb{R} \to \mathbb{R}$ are to be made precise below, and Ψ is a Bernstein function.

The operator $\Psi(-\Delta)$ is defined as follows. Let $S' := S'(\mathbb{R}^d)$ be the dual of the Schwartz test function space $S := S(\mathbb{R}^d)$. Define

$$D_\Psi := \{u \in S' : \mathcal{F}(u) \in L^1_{\rm loc}(\mathbb{R}^d), \Psi(|\xi|^2)\mathcal{F}(u) \in S'\}(\supset L^1(\mathbb{R}^d)) \tag{1.2}$$

and $\Psi(-\Delta) : D_\Psi \to S'$ by

$$\mathcal{F}(\Psi(-\Delta)u)(\xi) := \Psi(|\xi|^2)\mathcal{F}(u)(\xi),\ \xi \in \mathbb{R}^d,\ u \in D_\Psi, \tag{1.3}$$

where $\mathcal{F}$ stands for the Fourier transform on $\mathbb{R}^d$, i.e.,

$$\mathcal{F}(u)(\xi) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{ix\cdot\xi} u(x)dx,\ \xi \in \mathbb{R}^d,\ u \in L^1(\mathbb{R}^d).$$

($\mathcal{F}$ extends from S' to itself.)

Furthermore, $\Psi : [0,\infty) \to [0,\infty)$ is a a Bernstein function, i.e. an infinitely differentiable completely monotone function, that is,

$$(-1)^k \Psi^{(k)}(r) \geq 0,\ \forall r \geq 0,\ k = 1, 2, \dots .$$

A Bernstein function Ψ admits the unique representation (see [22], p. 21)

$$\Psi(r) = a_1 + a_2 r + \int_0^\infty (1 - e^{-rt})\mu(dt),\ \forall r \geq 0, \tag{1.4}$$

where $a_1, a_2 \geq 0$ and μ is a positive measure on $(0,\infty)$ such that

$$m := \int_0^\infty (1 \wedge t)\mu(dt) < \infty, \tag{1.5}$$

which implies

$$\Psi(r) \leq m(1 + r),\ r \geq 0. \tag{1.6}$$

Given a Bernstein function Ψ, there is a unique convolution semigroup of sub-probability (if $a_1 = 0$, probability) measures $(\eta_t^\Psi)_{t\geq 0}$ on $(0,\infty)$ such that

$$\mathcal{L}(\eta_t^\Psi)(\lambda) = e^{-t\Psi(\lambda)},\ \forall \lambda \geq 0, \tag{1.7}$$

where $\mathcal{L}(\eta_t^\Psi)$ is the Laplace transform of η_t^Ψ (see [22], p. 49).

A standard example is $\Psi(r) \equiv r^s$, $s \in (0,1)$, which corresponds to $a_1, a_2 = 0$ and

$$\mu(dt) = \frac{s}{\Gamma(1-s)}\ t^{-s-1}dt. \tag{1.8}$$

The hypotheses below will be in effect in Sections 2 and 4.1.

(i) $\beta\in C^1(\mathbb{R})\cap\mathrm{Lip}(\mathbb{R}),\ \beta(0)=0,\ \beta'(r)>0,\ \forall\, r\ne 0.$

(ii) $D\in(C^1\cap C_b)(\mathbb{R}^d;\mathbb{R}^d),\ \mathrm{div}\,D\in L^2_{\rm loc}(\mathbb{R}^d).$

(iii) $b\in C_b(\mathbb{R}).$

(iv) $(\mathrm{div}\,D)^-\in L^\infty,\ b(r)\ge 0,\ \forall r\in\mathbb{R}.$

(v) $\Psi:[0,\infty)\to[0,\infty)$ is a Bernstein function of the form (1.4) with $a_1,a_2=0$ which satisfies for some $s\in\left(\frac12,1\right)$ and $C\in(0,\infty)$,

$$\Psi(r)\ge Cr^s,\ \forall r\ge 0. \tag{1.9}$$

Here, we shall study the existence of a *mild solution* to equation (1.1) (see Definition 1.1 below) and also the uniqueness of distributional solutions. As regards the existence, we shall follow the semigroup methods used in [2]–[4] and in the special case $\Psi(r)\equiv r^s$, $s\in\left(\frac12,1\right)$ in [6]. Namely, we shall represent (1.1) as an abstract differential equation in $L^1(\mathbb{R}^d)$ of the form

$$\begin{aligned}&\frac{du}{dt}+A(u)=0,\quad t\ge 0,\\ &u(0)=u_0,\end{aligned} \tag{1.10}$$

where $A:D(A)\subset L^1(\mathbb{R}^d)\to L^1(\mathbb{R}^d)$ is an m-accretive realization in $L^1(\mathbb{R}^d)$ of the operator

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

to be made precise later on. (Here, div and Δ are taken in the sense of Schwartz distributions on $\mathbb{R}^d$.)

Definition 1.1. A function $u\in C([0,\infty);L^1(\mathbb{R}^d))$ is said to be a *mild solution* to (1.1) if, for each $0<T<\infty$,

$$u(t)=\lim_{h\to 0}u_h(t)\ \text{in}\ L^1(\mathbb{R}^d)\ \text{uniformly in}\ t\in[0,T], \tag{1.12}$$

where u_h is for each $h>0$ the solution to the equation

$$\begin{aligned}&\frac1h\,(u_h(t)-u_h(t-h)+A_0u_h(t))=0,\quad \forall t\ge 0,\\ &u_h(t)=u_0,\qquad\qquad\qquad\qquad\qquad\ \ \forall t\le 0.\end{aligned} \tag{1.13}$$

We note that, if u is a mild solution to (1.1), then it is also a *Schwartz distributional solution*, that is,

$$\begin{aligned}\int_0^\infty\int_{\mathbb{R}^d}&(u(t,x)\varphi_t(t,x)-\Psi(-\Delta)(\varphi(t,x))\beta(u(t,x))\\&+b(u(t,x))u(t,x)D(x)\cdot\nabla\varphi(t,x))dtdx\\&+\int_{\mathbb{R}^d}\varphi(0,x)u_0(dx)=0,\ \forall\,\varphi\in C_0^\infty([0,\infty)\times\mathbb{R}^d),\end{aligned}\tag{1.14}$$

where u_0 is a measure of finite variation on $\mathbb{R}^d$. The main existence result for equation (1.1) is given by Theorem 2.4 below, which amounts to saying that under Hypotheses (i)–(v) there is a mild solution u represented as $u(t)=S(t)u_0$, $t\ge 0$, where $S(t)$ is a continuous semigroup of nonlinear contractions in L^1. In Section 3, the uniqueness of distributional solutions to (1.1), (1.14) respectively, in the class $(L^1\cap L^\infty)((0,T)\times\mathbb{R}^d)\cap L^\infty(0,T;L^2)$ will be proved under Hypotheses (j)–(jjj) on D,b and β stated in Section 3. We would like to stress at this point that as in [5] (see, also, the pioneering papers [7] and [17]), where such a result was proved for local nonlinear Fokker–Planck equations (i.e. of type (1.1) with $-\Delta$ replacing $\Psi(-\Delta)$), we prove uniqueness in the large class of distributional solutions without any a priori-restrictions such as e.g. weak differentiability of the solutions. Therefore, our uniqueness results are considerably stronger than uniqueness results within the much smaller classes of mild solutions or in the local case of entropy solutions (see, e.g., [8]). In Section 4 we apply our analytic results on (1.1) to prove existence and uniqueness results (see Theorems 4.1 and 4.3, resp.) for the nonlinear martingale problem corresponding to (1.1) in the spirit of [14], where the latter were studied in the case of local Fokker–Planck equations with stronger regularity assumptions on the coefficients, not covering Nemytskii-type equations as above. Furthermore, we prove that the solutions to such a nonlinear martingale problem form a nonlinear Markov process in the sense of McKean [16] (see Theorem 4.4 below). The existence proof (Theorem 4.1) is based on Theorem 2.4 and [21], in which a nonlocal analogue of the superposition principle (see [12], [23]) is proved. The uniqueness proof (Theorem 4.3) is based on Theorems 3.1 and 3.3. As a consequence of this and Corollary 3.8 in [18], we finally obtain Theorem 4.4, which in particular realizes McKean's vision from [16], i.e. to identify the solutions of nonlinear PDEs as the one-dimensional time marginal law densities of a nonlinear Markov process, for nonlocal PDEs as in (1.1).

In the special case $\Psi(r)\equiv r^s$, $s\in\left(\frac{1}{2},1\right)$, equation (1.1) was studied [6] (see, also, [10], [11], [24] for a direct approach to existence theory in the case $\Psi(r) = r^s$ and $D \equiv 0$). Though the strategy in the present paper to prove the existence and uniqueness of solutions to (1.1) follows in great lines that developed in [6]. There is, however, not a large overlap and the specific arguments involved here are quite different. In particular, the probabilistic applications in Section 4 are much more difficult in the case of Bernstein functions.

Notation. $L^p(\mathbb{R}^d) = L^p$, $p \in [1,\infty]$ is the standard space of Lebesgue p-integrable functions on $\mathbb{R}^d$. We denote the corresponding local space by L^p_{loc} and the norm of L^p by $|\cdot|_p$. The inner product in L^2 is denoted by $(\cdot,\cdot)_2$. Let $H^\sigma(\mathbb{R}^d) = H^\sigma$, $0 < \sigma < \infty$, denote the standard Bessel space on $\mathbb{R}^d$ and $H^{-\sigma}$ its dual space. Let $C_b(\mathbb{R})$ denote the space of continuous and bounded functions on $\mathbb{R}$ and $C^1(\mathbb{R})$ the space of continuously differentiable functions on $\mathbb{R}$, and likewise $C^1(\mathbb{R}^d,\mathbb{R}^d)$ the space of continuously differentiable vector fields from $\mathbb{R}^d$ to $\mathbb{R}^d$. For any $T > 0$ and a Banach space $\mathcal{X}$, $C([0,T];\mathcal{X})$ denotes the space of $\mathcal{X}$-valued continuous functions on $[0,T]$ and by $L^p(0,T;\mathcal{X})$ the space of $\mathcal{X}$-valued L^p-Bochner integrable functions on $(0,T)$. $C^k_c(\mathcal{O})$, $\mathcal{O} \subset \mathbb{R}^d$, denotes the space of k-differentiable functions with compact support in $\mathcal{O}$ and $\mathcal{D}'(\mathcal{O})$ the space of Schwartz distributions on $\mathcal{O}$. $C^\infty_0([0,\infty)\times\mathbb{R}^d)$ denotes the space of differentiable functions on $[0,\infty)\times\mathbb{R}^d$ with compact in $[0,\infty)\times\mathbb{R}^d$. $S'(\mathbb{R}^d) = S'$ denotes the space of tempered distributions on $\mathbb{R}^d$ and $\mathcal{F}(y)$ the Fourier transform of $y \in S'(\mathbb{R}^d)$.

2 Existence of a mild solution

We first note that by (1.6) and (1.9) all functions

$$\mathbb{R}^d \ni \xi \to (\Psi(\varepsilon + |\xi|^2))^\alpha,\ \varepsilon > 0,\ \alpha \in \mathbb{R},$$

are multipliers on S, hence on S'. Hence, we may define the maps $\Psi(\varepsilon I - \Delta) : S' \to S'$ by

$$\Psi(\varepsilon I - \Delta)u := \mathcal{F}^{-1}(\Psi(\varepsilon + |\xi|^2)\mathcal{F}u,\ u \in S',$$

which are clearly linear homeomorphisms, and the following Hilbert spaces:

$$H^\Psi := H^\Psi(\mathbb{R}^d) := \left\{u \in S' : \sqrt{\Psi(1+|\xi|^2)}\,\mathcal{F}(u) \in L^2\right\}$$

and

$$\dot{H}^\Psi := \dot{H}^\Psi(\mathbb{R}^d) := \left\{u \in S' : \mathcal{F}(u) \in L^1_{\text{loc}} \text{ and } \sqrt{\Psi(|\xi|^2)}\,\mathcal{F}(u) \in L^2\right\}$$

with respective norms

$$|u|^2_{H^\Psi} := \int_{\mathbb{R}^d} \Psi(1+|\xi|^2)|\mathcal{F}(u)(\xi)|^2 d\xi$$
$$|u|^2_{\dot{H}^\Psi} := \int_{\mathbb{R}^d} \Psi(|\xi|^2)|\mathcal{F}(u)(\xi)|^2 d\xi.$$

We denote the corresponding inner products by $\langle \cdot,\cdot\rangle_{H^\Psi}$ and $\langle \cdot,\cdot\rangle_{\dot{H}^\Psi}$, respectively. By (1.6) and (1.9) we have the continuous embeddings

$$H^1 \subset H^\Psi \subset H^s \subset \dot{H}^s, \quad H^\Psi \subset \dot{H}^\Psi \subset \dot{H}^s, \tag{2.1}$$

where $\dot{H}^s$ denotes the usual homogeneous Sobolev space of order s. We note that $\dot{H}^s$ is only complete, if $s < \frac{d}{2}$, which holds in our case since $s<1$, $d \geq 2$. Clearly, H^Ψ is complete. But since by (1.9)

$$\frac{1}{\Psi(|\xi|^2)} \leq \frac{1}{C}\ \frac{1}{|\xi|^{2s}} \in L^1_{\text{loc}},$$

it is easy to show that also $\dot{H}^\Psi$ is complete.

Furthermore, for $\varepsilon \geq 0$ we define

$$D_{\varepsilon,2} := \{u \in L^2 : \Psi(\varepsilon I - \Delta)u \in L^2\}.$$

Then, it is elementary to check that $(\Psi(\varepsilon I - \Delta), D_{\varepsilon,2})$ is a nonnegative definite self-adjoint operators on L^2 (to which the usual operator calculus on L^2 applies) and that for all $u \in D_{0,2}$

$$|u|^2_{\dot{H}^\Psi} = \int_{\mathbb{R}^d} \left|\sqrt{\Psi(-\Delta)}\,u\right|^2 d\xi.$$

Furthermore, $D_{\varepsilon,2} := D_{1,2},\ \forall \varepsilon > 0$ and

$$|u|^2_{H^{\Psi,\varepsilon}} = \int_{\mathbb{R}^d} \left|\sqrt{\Psi(\varepsilon I - \Delta)}\,u\right|^2 d\xi,$$

where $|\cdot|_{H^{\Psi,\varepsilon}}$ is defined analogously to $|\cdot|_{H^\Psi}$ with $\Psi(\varepsilon + |\xi|^2)$ replacing $\Psi(1+|\xi|^2)$. Then, we have the following

Lemma 2.1.

(i) $\dot{H}^\Psi$ *and* $H^{\Psi,\varepsilon}$, $\varepsilon > 0$, *as well as* H^α, $\alpha \in (0,1]$, *are invariant under composition with Lipschitz continuous functions* $\varphi : \mathbb{R} \to \mathbb{R}$ *with* $\varphi(0) = 0$.

(ii) *For all* $\alpha \in \mathbb{R}$ *and* $\varepsilon > 0$, *we have for the inverse* $\Psi(\varepsilon I - \Delta)^{-1} : S' \to S'$ *of* $\Psi(\varepsilon I - \Delta) : S' \to S'$ *that*

$$\Psi(\varepsilon I - \Delta)^{-1}(H^\alpha) \subset H^{\alpha+2s}$$

and that the operator $\Psi(\varepsilon I - \Delta)^{-1} : H^\alpha \to H^{\alpha+2s}$ *is continuous.*

Proof. (i): This is an immediate consequence of [13, Section 1.5, in particular, Theorem 1.5.3 and Example 1.5.2], since, as it is proved there, $\left(\dot{H}^\Psi, \langle\cdot,\cdot\rangle_{\dot{H}^\Psi}\right)$, $\left(\dot{H}^{\Psi,\varepsilon}, \langle\cdot,\cdot\rangle_{\dot{H}^{\Psi,\varepsilon}}\right)$, $\varepsilon > 0$, and H^α, $\alpha \in (0,1]$ are transient Dirichlet spaces.

(ii): Using the definition of H^α, $\alpha \in \mathbb{R}$, in terms of Fourier transforms, the proof is elementary by (1.9). □

We shall now prove the following key lemma which is similar to Lemma 2.1 in [6]. (See, also, [2]–[4].)

Lemma 2.2. *Let* $\lambda_0 > 0$ *be as defined in* (2.32) *below. Then, under Hypotheses* (i)–(v) *there is a family of operators* $\{J_\lambda : L^1 \to L^1; \lambda > 0)\}$, *which for all* $\lambda \in (0, \lambda_0)$ *satisfies*

$$(I + \lambda A_0)(J_\lambda(f)) = f,\ \forall\, f \in L^1, \tag{2.2}$$

$$|J_\lambda(f_1) - J_\lambda(f_2)|_1 \le |f_1 - f_2|_1,\ \forall f_1, f_2 \in L^1, \tag{2.3}$$

$$J_{\lambda_2}(f) = J_{\lambda_1}\left(\frac{\lambda_1}{\lambda_2} f + \left(1 - \frac{\lambda_1}{\lambda_2}\right) J_{\lambda_2}(f)\right),\ \forall f \in L^1,\ \lambda_1, \lambda_2 > 0, \tag{2.4}$$

$$\int_{\mathbb{R}^d} J_\lambda(f)dx = \int_{\mathbb{R}^d} f\,dx,\ \forall f \in L^1, \tag{2.5}$$

$$J_\lambda(f) \ge 0, \quad \textit{a.e. on } \mathbb{R}^d,\ \textit{if } f \ge 0,\ \textit{a.e. on } \mathbb{R}^d, \tag{2.6}$$

$$|J_\lambda(f)|_\infty \le (1 + ||D| + (\operatorname{div} D)^-|_\infty^{\frac{1}{2}})|f|_\infty,\ \forall\, f \in L^1 \cap L^\infty, \tag{2.7}$$

$$\beta(J_\lambda(f)) \in H^\Psi \cap L^1 \cap L^\infty,\ \forall f \in L^1 \cap L^\infty. \tag{2.8}$$

Proof of Lemma 2.2. We shall first prove the existence of a solution to the equation

$$y+\lambda A_0(y)=f \text{ in } S', \tag{2.9}$$

where $f\in L^1$. We consider the approximating equation

$$y+\lambda\Psi(\varepsilon I-\Delta)(\beta_\varepsilon(y))+\lambda\,\mathrm{div}(D_\varepsilon b_\varepsilon(y)y)=f \text{ in } S', \tag{2.10}$$

where $\varepsilon\in(0,1]$, $\beta_\varepsilon(r):=\beta(r)+\varepsilon r$ and

$$D_\varepsilon:=\eta_\varepsilon D,\ \eta_\varepsilon\in C_0^1(\mathbb{R}^d),\ 0\le\eta_\varepsilon\le 1,\ |\nabla\eta_\varepsilon|\le 1,\ \eta_\varepsilon(x)=1 \text{ if } |x|<\frac{1}{\varepsilon}.$$

We have

$$\begin{gathered}|D_\varepsilon|\in L^2\cap L^\infty,\ |D_\varepsilon|\le|D|,\ \lim_{\varepsilon\to 0}D_\varepsilon(x)=D(x),\ \text{a.e. } x\in\mathbb{R}^d,\\ \mathrm{div}\,D_\varepsilon\in L^2,\ (\mathrm{div}\,D_\varepsilon)^-\le(\mathrm{div}\,D)^-+\mathbb{1}_{\left[|x|>\frac{1}{\varepsilon}\right]}|D|.\end{gathered} \tag{2.11}$$

The function b_ε is defined by

$$b_\varepsilon(r)\equiv\frac{(b*\varphi_\varepsilon)(r)}{1+\varepsilon|r|},\ \forall\, r\in\mathbb{R},$$

where $*$ is the convolution product and $\varphi_\varepsilon(r)=\frac{1}{\varepsilon}\,\varphi\left(\frac{r}{\varepsilon}\right)$ is a standard mollifier. We set $b^*_\varepsilon(r):=b_\varepsilon(r)r$, $r\in\mathbb{R}$, and note that b^*_ε is bounded, Lipschitz and $b^*_\varepsilon(0)=0$.

Let us assume first that $f\in L^2$ and rewrite (2.9) as

$$F_{\varepsilon,\lambda}(y)=f \text{ in } S', \tag{2.12}$$

where $F_{\varepsilon,\lambda}:L^2\to S'$ is defined by

$$F_{\varepsilon,\lambda}(y):=y+\lambda\Psi(\varepsilon I-\Delta)\beta_\varepsilon(y)+\lambda\,\mathrm{div}(D_\varepsilon b^*_\varepsilon(y)),\ \forall y\in L^2.$$

We set

$$G_\varepsilon(y):=\Psi(\varepsilon I-\Delta)(y),\ y\in S'.$$

Now, we shall show that (2.12) has a unique solution $y_\varepsilon\in L^2$. To this end, we rewrite it as

$$G_\varepsilon^{-1}(F_{\varepsilon,\lambda}(y))=G_\varepsilon^{-1}f(\in H^{2s}),$$

that is,

$$G_\varepsilon^{-1}y+\lambda\beta_\varepsilon(y)+\lambda G_\varepsilon^{-1}\,\mathrm{div}(D_\varepsilon b^*_\varepsilon(y))=G_\varepsilon^{-1}f. \tag{2.13}$$

Since $D_\varepsilon b_\varepsilon^*(y) \in L^2$, we have $\mathrm{div}(D_\varepsilon b_\varepsilon^*(y)) \in H^{-1}$, and so, by Lemma 2.1 (ii) and because $s > \frac{1}{2}$, we have that $G_\varepsilon^{-1} F_{\varepsilon,\lambda} : L^2 \to L^2$ is continuous. Now, it is easy to see that (2.13) has a unique solution, $y_\varepsilon \in L^2$ for small enough λ, because, by (2.13) we have, for $y_1, y_2 \in L^2$,

$$\begin{aligned}
&(G_\varepsilon^{-1}(F_{\varepsilon,\lambda}(y_2) - F_{\varepsilon,\lambda}(y_1)), y_2 - y_1)_2 \\
&\quad = (G_\varepsilon^{-1}(y_2 - y_1), y_2 - y_1)_2 + \lambda(\beta_\varepsilon(y_2) - \beta_\varepsilon(y_1), y_2 - y_1)_2 \\
&\quad -\lambda\ {}_{H^{-1}}\langle \mathrm{div}(D_\varepsilon(b_\varepsilon^*(y_2) - b_\varepsilon^*(y_1))), G_\varepsilon^{-1}(y_2 - y_1)\rangle_{H^1} \\
&\quad \geq |y_2 - y_1|_{H^{-s}}^2 + \lambda\varepsilon |y_2 - y_1|_2^2 \\
&\qquad -\lambda c_1 |D_\varepsilon(b_\varepsilon^*(y_2) - b_\varepsilon^*(y_1))|_2 |\nabla G_\varepsilon^{-1}(y_2 - y_1)|_2 \\
&\quad \geq |y_2 - y_1|_{H^{-s}}^2 + \lambda\varepsilon |y_2 - y_1|_2^2 - \lambda c_\varepsilon |D|_\infty \mathrm{Lip}(b_\varepsilon^*) |y_2 - y_1|_2 |y_2 - y_1|_{H^{1-2s}},
\end{aligned} \tag{2.14}$$

where $c_\varepsilon \in (0, \infty)$ is independent of λ, y_1, y_2. Since $-s < 1 - 2s < 0$, by interpolation we have for $\theta := \frac{2s-1}{s}$ that

$$|y_2 - y_1|_{H^{1-2s}} \leq |y_2 - y_1|_2^{1-\theta} |y_2 - y_1|_{H^{-s}}^{\theta},$$

and so, by Young's inequality we find that the left hand side of (2.14) dominates

$$\lambda(\varepsilon - \lambda c_\varepsilon)|y_2 - y_1|_2^2 + \frac{1}{2}|y_2 - y_1|_{H^{-s}}^2$$

for some $c_\varepsilon \in (0, \infty)$ independent of λ, y_1 and y_2. Hence, for some $\lambda_\varepsilon \in (0, \infty)$, we conclude that $G_\varepsilon^{-1} F_{\varepsilon,\lambda}$ is strictly monotone on L^2 for $\lambda \in (0, \lambda_\varepsilon)$.

By Lemma 2.1 (ii), it follows from (2.13) that $\beta_\varepsilon(y_\varepsilon) \in H^{2s-1}$, hence by Lemma 2.1 (i), because $s < 1$ and the inverse of β_ε is Lipschitz and zero at zero, the solution y_ε also belongs to H^{2s-1}, hence $b_\varepsilon^*(y_\varepsilon) \in H^{2s-1}$. Since $s \in \left(\frac{1}{2}, 1\right)$ and $D \in C^1(\mathbb{R}^d; \mathbb{R}^d)$, by simple bootstrapping (2.13) implies that $y_\varepsilon \in H^1$, and therefore, by (2.13) we have

$$\beta_\varepsilon(y_\varepsilon) \in G_\varepsilon^{-1}(L^2) \ \text{ and so } \ G_\varepsilon \beta_\varepsilon(y_\varepsilon) \in L^2. \tag{2.15}$$

Furthermore, for $f \in L^2$ and $\lambda \in (0, \lambda_\varepsilon)$, y_ε is the unique solution of (2.12).

Assume now that $\lambda \in (0, \lambda_\varepsilon)$ and that $f \geq 0$, a.e. on $\mathbb{R}^d$. Then, we have

$$y_\varepsilon \geq 0, \quad \text{a.e. on } \mathbb{R}^d. \tag{2.16}$$

To prove this, consider the function

$$\eta_\delta(r) = \begin{cases} -1 & \text{for} \quad r \leq -\delta, \\ \dfrac{r}{\delta} & \qquad\ \ r \in (-\delta, 0), \\ 0 & \text{for} \quad r \geq 0, \end{cases} \tag{2.17}$$

where $\delta > 0$, and multiply (2.10), where $y = y_\varepsilon$, by $\eta_\delta(\beta_\varepsilon(y_\varepsilon))$ ($\in H^1$) and integrate over $\mathbb{R}^d$. By (2.15) we get

$$\begin{aligned}\int_{\mathbb{R}^d} y_\varepsilon \eta_\delta(\beta_\varepsilon(y_\varepsilon))dx + \lambda \int_{\mathbb{R}^d} G_\varepsilon(\beta_\varepsilon(y_\varepsilon))\eta_\delta(\beta_\varepsilon(y_\varepsilon))dx \\ = \int_{\mathbb{R}^d} f\eta_\delta(\beta_\varepsilon(y_\varepsilon))dx + \lambda \int_{\mathbb{R}^d} D_\varepsilon b_\varepsilon^*(y_\varepsilon)\eta'_\delta(\beta_\varepsilon(y_\varepsilon)) \cdot \nabla\beta_\varepsilon(y_\varepsilon)dx.\end{aligned} \tag{2.18}$$

On the other hand, we have

$$\begin{aligned}\int_{\mathbb{R}^d} G_\varepsilon(u)\varphi(u)dx = \langle u, \varphi(u)\rangle_{H^{\Psi,\varepsilon}} \geq \frac{1}{2}(\mathrm{Lip}(\varphi))^{-1} \langle \varphi(u), \varphi(u)\rangle_{H^{\Psi,\varepsilon}} \geq 0, \\ \forall u \in G_\varepsilon^{-1}(L^2)(\subset H^\Psi),\end{aligned} \tag{2.19}$$

for all non constant, nondecreasing Lipschitz functions $\varphi : \mathbb{R} \to \mathbb{R}$. This is a well-known inequality in the theory of Dirichlet forms (see, e.g., [20, Examples 6.4 and 6.5]). By (2.15), this yields

$$\int_{\mathbb{R}^d} G_\varepsilon(\beta_\varepsilon(y_\varepsilon))\eta_\delta(\beta_\varepsilon(y_\varepsilon))dx \geq 0. \tag{2.20}$$

Taking into account that $|\beta_\varepsilon(y_\varepsilon)| \geq \varepsilon|y_\varepsilon|$, we have

$$\begin{aligned}&\left|\int_{\mathbb{R}^d} D_\varepsilon b_\varepsilon(y_\varepsilon)y_\varepsilon\eta'_\delta(\beta_\varepsilon(y_\varepsilon))\nabla\beta_\varepsilon(y_\varepsilon)dx\right| \\ &\qquad \leq \frac{1}{\delta}\,|b|_\infty \int_{\widetilde{E}_\varepsilon^\delta} |y_\varepsilon|\,|\nabla\beta_\varepsilon(y_\varepsilon)|\,|D_\varepsilon|dx \\ &\qquad \leq \frac{1}{\varepsilon}\,|b|_\infty\|D_\varepsilon\|_{L^2}\left(\int_{\widetilde{E}_\varepsilon^\delta} |\nabla\beta_\varepsilon(y_\varepsilon)|^2dx\right)^{\frac{1}{2}} \to 0 \text{ as } \delta \to 0.\end{aligned} \tag{2.21}$$

Here $\widetilde{E}_\varepsilon^\delta = \{-\delta < \beta_\varepsilon(y_\varepsilon) \leq 0\}$ and we used that $\nabla\beta_\varepsilon(y_\varepsilon) = 0$, a.e. on $\{\beta_\varepsilon(y_\varepsilon) = 0\}$.

Since $\mathrm{sign}\,\beta_\varepsilon(r) \equiv \mathrm{sign}\,r$, by (2.18)–(2.21) we get, for $\delta \to 0$, that $y_\varepsilon^- = 0$, a.e. on $\mathbb{R}^d$ and so (2.16) holds.

If $\lambda \in (0, \lambda_\varepsilon)$ and $y_\varepsilon = y_\varepsilon(\lambda, f)$ is the solution to (2.10) in L^2, we have for $f_1, f_2 \in L^1 \cap L^2$

$$\begin{aligned}y_\varepsilon(\lambda, f_1) - y_\varepsilon(\lambda, f_2) + \lambda G_\varepsilon(\beta_\varepsilon(y_\varepsilon(\lambda, f_1)) - \beta_\varepsilon(y_\varepsilon(\lambda, f_2))) \\ +\lambda \text{ div } D_\varepsilon(b_\varepsilon^*(y_\varepsilon(\lambda, f_1)) - b_\varepsilon^*(y_\varepsilon(\lambda, f_2))) = f_1 - f_2.\end{aligned} \tag{2.22}$$

Now, for $\delta > 0$ consider the function

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

If multiply (2.22) by $\mathcal{X}_\delta(\beta_\varepsilon(y_\varepsilon(\lambda, f_1)) - \beta_\varepsilon(y_\varepsilon(\lambda, f_2)))$, we get

$$\begin{aligned}
&\int_{\mathbb{R}^d} (y_\varepsilon(\lambda, f_1) - y_\varepsilon(\lambda, f_2))\mathcal{X}_\delta(\beta_\varepsilon(y_\varepsilon(\lambda, f_1)) - \beta_\varepsilon(y_\varepsilon(\lambda, f_2)))dx \\
&\quad \leq \lambda \, \frac{1}{\delta} \int_{E_\varepsilon^\delta} |b_\varepsilon^*(y_\varepsilon(\lambda, f_1)) - b_\varepsilon^*(y_\varepsilon(\lambda, f_2))| \, |D_\varepsilon| |\nabla(\beta_\varepsilon(y_\varepsilon(\lambda, f_1)) - \beta_\varepsilon(y_\varepsilon(\lambda, f_2)))| dx \\
&\quad + |f_1 - f_2|_1,
\end{aligned}$$

because, by virtue of (2.19),

$$\int_{\mathbb{R}^d} G_\varepsilon(\beta_\varepsilon(y_\varepsilon, f_1) - \beta_\varepsilon(y_\varepsilon, f_2))\mathcal{X}_\delta(\beta_\varepsilon(y_\varepsilon, f_1) - \beta_\varepsilon(y_\varepsilon, f_2))dx \geq 0.$$

Set $E_\varepsilon^\delta = \{|\beta_\varepsilon(y_\varepsilon(\lambda, f_1)) - \beta_\varepsilon(y_\varepsilon(\lambda, f_2))| \leq \delta\}$.

Since $|\beta_\varepsilon(r_1) - \beta_\varepsilon(r_2)| \geq \varepsilon|r_1 - r_2|$, $D_\varepsilon \in L^2(\mathbb{R}^d; \mathbb{R}^d)$, $b_\varepsilon^* \in \text{Lip}(\mathbb{R})$, $y_\varepsilon(\lambda, f_i) \in H^1$, $i = 1, 2$, and $\nabla(\beta_\varepsilon(y_\varepsilon(\lambda, f_1)) - \beta_\varepsilon(y_\varepsilon(\lambda, f_2))) = 0$, a.e. on $\{\beta_\varepsilon(y_\varepsilon(\lambda, f_1)) - \beta_\varepsilon(y_\varepsilon(\lambda, f_2)) = 0\}$, we get that

$$\lim_{\delta \to 0} \frac{1}{\delta} \int_{E_\varepsilon^\delta} |b_\varepsilon^*(y_\varepsilon(\lambda, f_1)) - b_\varepsilon^*(y_\varepsilon(\lambda, f_2))| |D_\varepsilon| |\nabla(\beta_\varepsilon(y_\varepsilon(\lambda, f_1)) - \beta_\varepsilon(y_\varepsilon(\lambda, f_2)))| dx = 0.$$

This yields

$$|y_\varepsilon(\lambda, f_1) - y_\varepsilon(\lambda, f_2)|_1 \leq |f_1 - f_2|_1, \ \forall \lambda \in (0, \lambda_\varepsilon). \tag{2.23}$$

Hence,

$$|y_\varepsilon(\lambda, f)|_1 \leq |f|_1, \ \forall \, f \in L^1 \cap L^2, \ \lambda \in (0, \lambda_\varepsilon). \tag{2.24}$$

Now, let us remove the restriction on $\lambda \in (0, \lambda_\varepsilon)$. To this purpose, define the operator $A_\varepsilon : D_0(A_\varepsilon) \to L^1$ by

$$\begin{aligned}
A_\varepsilon(y) &:= G_\varepsilon(\beta_\varepsilon(y)) + \text{div}(D_\varepsilon b_\varepsilon^*(y)), \\
D_0(A_\varepsilon) &:= \{y \in L^1 : G_\varepsilon(\beta_\varepsilon(y)) + \text{div}(D_\varepsilon b_\varepsilon^*(y)) \in L^1\},
\end{aligned}$$

and

$$J_\lambda^\varepsilon(f) := y_\varepsilon(\lambda, f),\ f \in L^1 \cap L^2,\ \lambda \in (0, \lambda_\varepsilon).$$

Then, by (2.15) and since $y_\varepsilon(\lambda, f) \in H^1$,

$$\begin{aligned} J_\lambda^\varepsilon(f) \in D_0(A_\varepsilon) \cap H^1 \text{ and } \beta_\varepsilon(J_\lambda^\varepsilon(f)) \in G_\varepsilon^{-1}(L^2),\\ \forall f \in L^1 \cap L^2,\ \lambda \in (0, \lambda_\varepsilon). \end{aligned} \tag{2.25}$$

Furthermore, by (2.23), J_λ^ε extends by continuity to an operator $J_\lambda^\varepsilon : L^1 \to L^1$. We note that the operator $(A_\varepsilon, D_0(A_\varepsilon))$ is closed as an operator on L^1. Hence, by (2.23), $J_\lambda^\varepsilon(L^1) \subset D_0(A_\varepsilon)$ and so, $J_\lambda^\varepsilon(f)$ solves (2.10) for all $f \in L^1$.

Now, define $D(A_\varepsilon) := J_\lambda^\varepsilon(L^1)$ and restrict A_ε to $D(A_\varepsilon)$. It is easy to see that $D(A_\varepsilon)$ is independent of $\lambda \in (0, \lambda_\varepsilon)$. Furthermore, $J_\lambda^\varepsilon(f)$ is the unique solution in $D(A_\varepsilon)$ of (2.10) for all $f \in L^1$, $\lambda \in (0, \lambda_\varepsilon)$.

Now let $0 < \lambda_1 < \lambda_\varepsilon$. Then, for $\lambda \geq \lambda_\varepsilon$, the equation

$$y + \lambda A_\varepsilon(y) = f\ (\in L^1),\ y \in D(A_\varepsilon), \tag{2.26}$$

can be rewritten as

$$y + \lambda_1 A_\varepsilon(y) = \left(1 - \frac{\lambda_1}{\lambda}\right) y + \frac{\lambda_1}{\lambda} f,$$

equivalently,

$$y = J_{\lambda_1}^\varepsilon\left(\left(1 - \frac{\lambda_1}{\lambda}\right) y + \frac{\lambda_1}{\lambda} f\right). \tag{2.27}$$

Taking into account that, by (2.23), $|J_{\lambda_1}^\varepsilon(f_1) - J_{\lambda_1}^\varepsilon(f_2)|_1 \leq |f_1 - f_2|_1$, it follows that (2.27) has a unique solution $y_\varepsilon \in D(A_\varepsilon)$. Let $J_\lambda^\varepsilon(f) := y_\varepsilon$, $\lambda \in [\lambda_\varepsilon, \infty)$, $f \in L^1$, denote this solution to (2.26). Then, $J_\lambda^\varepsilon(f)$ is the unique solution in $D(A_\varepsilon)$ of (2.10) for all $\lambda > 0$, $f \in L^1$. By (2.27) we see that (2.23), (2.24) extend to all $\lambda > 0$, $f \in L^1$.

Let us prove that, for $f \in L^1 \cap L^2$,

$$J_\lambda^\varepsilon(f) \in H^1 \text{ and } \beta_\varepsilon(J_\lambda^\varepsilon(f)) \in G_\varepsilon^{-1}(L^2) \text{ for all } \lambda > 0. \tag{2.28}$$

Here is the argument. Fix $\lambda_1 \in [\lambda_\varepsilon/2, \lambda_\varepsilon)$ and set $\lambda{:=}2\lambda_1$. Define $T : L^1 \to L^1$ by

$$T(y) {:=} J_{\lambda_1}^\varepsilon\left(\frac{1}{2} y + \frac{1}{2} f\right),\ y \in L^1.$$

Then, as just proved, for any $f_0 \in L^1 \cap L^2$ we have $\lim_{n\to\infty} T^n(f_0) = J^\varepsilon_\lambda(f)$ in L^1. It suffices to prove

$$J^\varepsilon_\lambda(f) \in L^2, \tag{2.29}$$

because then $J^\varepsilon_\lambda(f) = J^\varepsilon_{\lambda_1}(g)$ with $g := \frac{1}{2}\, J^\varepsilon_\lambda(f) + \frac{1}{2}\, f \in L^1 \cap L^2$, and so (2.28) follows by (2.25).

To prove (2.29), we note that by (2.25) we have, for $n \in \mathbb{N}$,

$$(I + \lambda_1 A_\varepsilon) T^n(f_0) = \frac{1}{2}\, T^{n-1}(f_0) + \frac{1}{2}\, f$$

with $T^n(f_0) \in H^1$ and $\beta_\varepsilon(T^n(f_0)) \in G^{-1}_\varepsilon(L^2)$. Hence, multiplying this equation, by $T^n(f_0)$ and integrating over $\mathbb{R}^d$ we find

$$\begin{aligned} &|T^n f_0|_2^2 + \lambda_1 (G_\varepsilon(\beta_\varepsilon(T^n(f_0))), \beta_\varepsilon^{-1}(\beta_\varepsilon(T^n(f_0))))_2 \\ &= \lambda_1 \int_{\mathbb{R}^d} (D_\varepsilon b^*_\varepsilon(T^n(f_0))) \cdot \nabla(T^n(f_0)) d\xi + \left(\frac{1}{2} T^{n-1}(f_0) + \frac{1}{2} f, T^n(f_0)\right)_2. \end{aligned} \tag{2.30}$$

We set

$$g_\varepsilon(r) := \int_0^r b^*_\varepsilon(\tau) d\tau,\ \ r \in \mathbb{R}. \tag{2.31}$$

By Hypothesis (iii) we have

$$0 \le g_\varepsilon(r) \le |b^*_\varepsilon|_\infty r,\ \ r \in \mathbb{R},$$

and so the right hand side of (2.30) is equal to

$$-\lambda_1 (\mathrm{div}\ D_\varepsilon, g_\varepsilon(T^n(f_0)))_2 + \left(\frac{1}{2}\, (T^{n-1}(f_0) + f), T^n(f_0)\right)_2.$$

By (2.20) we thus obtain

$$\begin{aligned} |T^n(f_0)|_2^2 \le\ & \lambda_1 |b^*_\varepsilon|_\infty |(\mathrm{div}\, D_\varepsilon)^-|_2 |T^n(f_0)|_2 + \frac{1}{2} |T^n(f_0)|_2^2 \\ & + \frac{1}{4} (|T^{n-1}(f_0)|_2^2 + |f|_2^2), \end{aligned}$$

therefore, using Young's inequality we obtain

$$\begin{aligned} |T^n(f_0)|_2^2 &\le C_\varepsilon + \frac{2}{3}\, |T^{n-1}(f_0)|_2^2, \\ C_\varepsilon &:= \frac{16}{3}\, \lambda_1^2 |b^*_\varepsilon|^2_\infty |(\mathrm{div}\, D_\varepsilon)^-|_2^2 + \frac{2}{3}\, |f|_2^2. \end{aligned}$$

Finally, we get

$$|T^n(f_0)|_2^2 \le C_\varepsilon \sum_{k=0}^{n} \left(\frac{2}{3}\right)^k + \left(\frac{2}{3}\right)^n |f_0|_2^2,\ n \in \mathbb{N}.$$

Hence, we get

$$|J_\lambda^\varepsilon(f)|_2^2 \le \liminf_{n\to\infty} |T^n(f_0)|_2^2 \le 3C_\varepsilon < \infty,$$

so, (2.29) holds for $\lambda = 2\lambda_1$, and finally we get (2.29) for all $\lambda > 0$. $\square$

We set

$$\lambda_0 := \left(\left|(\mathrm{div}\ D)^- + |D|\right|_\infty^{\frac{1}{2}} |b|_\infty\right)^{-1}, \tag{2.32}$$

where $\frac{1}{0} := \infty$. Then, for $f \in L^1 \cap L^\infty$ and $y_\varepsilon := J_\lambda^\varepsilon(f)$, $\lambda > 0$, we have

$$|y_\varepsilon|_\infty \le (1 + \left||D| + (\mathrm{div}\, D)^-\right|_\infty^{\frac{1}{2}})|f|_\infty,\ \forall \lambda \in (0, \lambda_0). \tag{2.33}$$

Here is the argument. If $M_\varepsilon = |(\mathrm{div}\, D_\varepsilon)^-|_\infty^{\frac{1}{2}} |f|_\infty$, then we get by (2.10) that

$$\begin{aligned} &(y_\varepsilon - |f|_\infty - M_\varepsilon) + \lambda G_\varepsilon(\beta_\varepsilon(y_\varepsilon) - \beta_\varepsilon(|f|_\infty + M_\varepsilon)) \\ &\quad + \lambda\Psi(\varepsilon)(\beta_\varepsilon(|f|_\infty + M_\varepsilon)) + \lambda\ \mathrm{div}(D_\varepsilon(b_\varepsilon^*(y_\varepsilon) - b_\varepsilon^*(|f|_\infty + M_\varepsilon))) \le 0, \end{aligned} \tag{2.34}$$

because $G_\varepsilon 1 = \Psi(\varepsilon)$ and $\mathcal{F}(1) = (2\pi)^{\frac{d}{2}}\,\delta_0$. Multiplying with $\mathcal{X}_\delta((\beta_\varepsilon(y_\varepsilon) - \beta_\varepsilon(|f|_\infty + M_\varepsilon)^+)$ in (2.34), letting $\delta \to 0$ and using (2.19), we get by (2.11)

$$y_\varepsilon \le (1 + ||D| + (\mathrm{div}\, D)^-|_\infty^{\frac{1}{2}})|f|_\infty,\ \text{a.e. in } \mathbb{R}^d,$$

and, similarly, for $-y_\varepsilon$ which yields (2.33) for $\lambda \in (0, \lambda_0)$. This yields

$$|J_\lambda^\varepsilon(f)|_1 + |J_\lambda^\varepsilon(f)|_\infty \le c_1,\ \forall \varepsilon > 0,\ \lambda \in (0, \lambda_0), \tag{2.35}$$

where $c_1 = c_1(|f|_1, |f|_\infty)$ is independent of ε and λ.

Now, fix $\lambda \in (0, \lambda_0)$ and $f \in L^1 \cap L^\infty$. For $\varepsilon \in (0, 1]$ set $y_\varepsilon := J_\lambda^\varepsilon(f)$.

Then, since $\beta_\varepsilon(y_\varepsilon) \in H^1$ and $G_\varepsilon \beta_\varepsilon(y_\varepsilon) \in L^2$ by (2.28),

$$\begin{aligned} &(y_\varepsilon, \beta_\varepsilon(y_\varepsilon))_2 + \lambda(G_\varepsilon(\beta_\varepsilon(y_\varepsilon)), \beta_\varepsilon(y_\varepsilon))_2 \\ &\qquad = \lambda \int_{\mathbb{R}^d} (D_\varepsilon b_\varepsilon^*(y_\varepsilon)) \cdot \nabla\beta_\varepsilon(y_\varepsilon) dr + (f, \beta_\varepsilon(y_\varepsilon))_2. \end{aligned} \tag{2.36}$$

Setting

$$\widetilde{\Psi}_\varepsilon(r) := \int_0^r b_\varepsilon^*(\tau)\beta_\varepsilon'(\tau)d\tau,\ r \in \mathbb{R}, \tag{2.37}$$

by Hypotheses (iii), (iv) we have

$$0 \le \widetilde{\Psi}_\varepsilon(r) \le \frac{1}{2}\,|b|_\infty(|\beta'|_\infty + 1)r^2,\ \forall r \in \mathbb{R},$$

and hence, since $y_\varepsilon \in H^1$, the right hand side of (2.36) is equal to

$$-\lambda \int_{\mathbb{R}^d} \text{div}\ D_\varepsilon \widetilde{\Psi}_\varepsilon(y_\varepsilon)dx + (f, \beta_\varepsilon(y_\varepsilon))_2,$$

which, because $(y_\varepsilon, \beta_\varepsilon(y_\varepsilon))_2 \ge 0$ and $H^1 \subset H^\Psi \subset H^s$ (see (2.1)), by (2.11) and Hypothesis (iv) implies that

$$\begin{aligned}\lambda|\beta_\varepsilon(y_\varepsilon)|^2_{H^{\Psi,\varepsilon}} &\le \frac{1}{2}\,\lambda|b|_\infty(|\beta'|_\infty + 1)\left|(\text{div}\ D)^- + |D|\right|_\infty |y_\varepsilon|_2^2 \\ &+\frac{1}{2}\,|\beta_\varepsilon(y_\varepsilon)|_2^2 + \frac{1}{2}\,|f|_2^2.\end{aligned}$$

Since $|\beta_\varepsilon(r)| \le (\text{Lip}(\beta) + 1)|r|,\ r \in \mathbb{R}$, by (2.35) we obtain

$$\sup_{\varepsilon\in(0,1]} |\beta_\varepsilon(y_\varepsilon)|^2_{H^{\Psi,\varepsilon}} \le C, \tag{2.38}$$

for some $C \in (0, \infty)$. Obviously, we have for all $u \in H^\Psi$ ($\subset \dot{H}^\Psi$) and all $\varepsilon \in (0, 1]$

$$|\Psi(-\Delta)^{\frac{1}{2}}u|_2^2 \le |\Psi(\varepsilon I - \Delta)^{\frac{1}{2}}u|_2^2 \le |\Psi(-\Delta)u|_2^2 + \Psi(\varepsilon)|u|_2^2, \tag{2.39}$$

where we used the sub linearity of Ψ in the second step.

Hence, we conclude from (2.35) and (2.38) that (along a subsequence) as $\varepsilon \to 0$,

$$\begin{aligned}\beta_\varepsilon(y_\varepsilon) &\to z \text{ weakly in } H^\Psi,\ \text{hence strongly in } L^2_{\text{loc}}(\mathbb{R}^d) \text{ (by (2.1))},\\ G_\varepsilon(\beta_\varepsilon(y_\varepsilon)) &\to \Psi(-\Delta)z \text{ in } S',\\ y_\varepsilon &\to y \text{ weakly in } L^2 \text{ and weakly}^* \text{ in } L^\infty.\end{aligned}$$

The second statement follows, because $G_\varepsilon(\varphi) \to \Psi(-\Delta)\varphi$ in L^2 for all $\varphi \in S$. Hence (selecting another subsequence, if necessary), $\beta(y_\varepsilon) \to z$, a.e. Since

β^{-1} is continuous, it follows that $y_\varepsilon \to \beta^{-1}(z) = y$, a.e., and, therefore, $z = \beta(y) \in H^\Psi$. Furthermore, we have

$$b^*_\varepsilon(y_\varepsilon) \to b^*(y) \text{ weakly in } L^2.$$

Recalling that y_ε solves (2.10), we can let $\varepsilon \to 0$ in (2.10) to find that

$$y + \lambda\Psi(-\Delta)\beta(y) + \lambda\,\mathrm{div}(Db^*(y)) = f \text{ in } S'. \tag{2.40}$$

Since $\beta \in \mathrm{Lip}(\mathbb{R})$, the operator $(A_0, D(A_0))$ defined in (1.11) is closed in L^1. If y is as in (2.40), we define

$$J_\lambda(f) := y \in D(A_0),\ \ \lambda \in (0, \lambda_0).$$

Then (2.8) holds and it follows by (2.23) (which, as mentioned earlier, holds, in fact, for all $\lambda > 0$) and Fatou's lemma that for $f_1, f_2 \in L^1 \cap L^\infty$

$$|J_\lambda(f_1) - J_\lambda(f_2)|_1 \le |f_1 - f_2|_1. \tag{2.41}$$

Hence J_λ extends continuously to all of L^1, still satisfying (2.41) for all $f_1, f_2 \in L^1$. Then it follows by the closedness of $(A_0, D(A_0))$ on L^1 that $J_\lambda(f) \in D(A_0)$ and that it solves (2.40) for all $f \in L^1$.

Clearly, (2.6) and (2.7) follow from (2.16) and (2.33), respectively.

Hence, Lemma 2.2 is proved except for (2.4) and (2.5). Equation (2.4) is obvious, since by (2.2) it is equivalent to

$$(I + \lambda_1 A_0)J_{\lambda_2}(f) = \frac{\lambda_1}{\lambda_2}\,f + \left(1 - \frac{\lambda_1}{\lambda_2}\right) J_{\lambda_2} f,$$

or, equivalently,

$$(I + \lambda_2 A_0)J_{\lambda_2}(f) = f.$$

Now, let us prove (2.5). We may assume that $f \in L^1 \cap L^\infty$ and set $y := J_\lambda(f)$. Let $\psi_n \in C_0^\infty(\mathbb{R}^d)$, $\psi_n \uparrow 1$, as $n \to \infty$, $\lim\limits_{n\to\infty} \nabla\psi_n = 0$ on $\mathbb{R}^d$, with $\sup\limits_n |\nabla\psi_n|_\infty < \infty$. Define

$$\varphi_n := (I + \Psi(-\Delta))^{-1}\psi_n = g_1^\Psi * \psi_n,\ \ n \in \mathbb{N},$$

where g_ε^Ψ is as in the Appendix. Then, we have by (A.2)

$$\begin{gathered}\varphi_n \uparrow 1,\ \ \nabla\varphi_n \to 0 \text{ on } \mathbb{R}^d \ \text{ as } n \to \infty,\\ \sup_n(|\varphi_n|_\infty + |\nabla\varphi_n|_\infty) < \infty,\ \ \varphi_n \in L^1 \cap H^{2s},\ n \in \mathbb{N}.\end{gathered} \tag{2.42}$$

Furthermore,

$$\Psi(-\Delta)\varphi_n = \psi_n - (I+\Psi(-\Delta))^{-1}\psi_n \in L^1 \cap L^\infty,$$

are bounded in L^∞ and, as $n \to \infty$, $\Psi(-\Delta)\varphi_n \to 0$ $dx -$ a.e. Hence,

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} \Psi(-\Delta)\varphi_n\, \beta(y)dx = 0. \tag{2.43}$$

Consequently, since $\beta(y) \in L^1$, $y \in D(A)$ with $A_0 y \in L^1$, we have

$$\begin{aligned}
&\int_{\mathbb{R}^d} A_0 y\, dx = \lim_{n\to\infty} \int_{\mathbb{R}^d} \varphi_n A_0 y\, dx \\
&= -\int_{\mathbb{R}^d} \beta(y)dx + \lim_{n\to\infty} {}_S\langle \varphi_n, (I+\Psi(-\Delta))\beta(y) + \operatorname{div}(Db^*(y))\rangle_{S'} \\
&= -\int_{\mathbb{R}^d} \beta(y)dx + \lim_{n\to\infty} \int_{\mathbb{R}^d} (I+\Psi(-\Delta))\varphi_n\, \beta(y)dx + \lim_{n\to\infty} \int_{\mathbb{R}^d} \nabla\varphi_n \cdot Db^*(y)dx,
\end{aligned}$$

which by (2.42) and (2.43) is equal to zero. Hence, integrating the equation $y + \lambda A_0 y = f$ over $\mathbb{R}^d$, (2.5) follows. □

Now, define the operator A by

$$\begin{aligned}
D(A) &:= J_\lambda(L^1)\ (\subset D(A_0)), \\
A(y) &:= A_0(y),\ y \in D(A).
\end{aligned} \tag{2.44}$$

It is easily seen that $J_\lambda(L^1)$ is independent of $\lambda \in (0,\lambda_0)$ and that

$$J_\lambda = (I+\lambda A)^{-1},\ \lambda \in (0,\lambda_0).$$

Therefore, we have

Lemma 2.3. *Under Hypotheses* (i)–(iv)*, the operator A is m-accretive in L^1 and $(I+\lambda A)^{-1} = J_\lambda$, $\lambda \in (0,\lambda_0)$. Moreover, if $\beta \in C^\infty(\mathbb{R})$, then $\overline{D(A)} = L^1$.*

(Here, $\overline{D(A)}$ is the closure of $D(A)$ in L^1.) Indeed, if $\beta \in C^\infty(\mathbb{R})$, then by assumption (ii)

$$A_0(\varphi) = \Psi(-\Delta)\beta(\varphi) + \operatorname{div}(Db(\varphi)\varphi) \in L^1,\ \forall \varphi \in C_0^\infty(\mathbb{R}^d),$$

and so $\overline{D(A)} = L^1$, as claimed.

Then, by the Crandall & Liggett theorem (see, e.g., [1], p. 131), the Cauchy problem (1.10) has, for each $u_0 \in \overline{D(A)}$, a unique mild solution $u = u(t, u_0) \in C([0,\infty); L^1)$ and $S(t)u_0 = u(t, u_0)$ is a C_0-semigroup of contractions on L^1, that is,

$$
\begin{gathered}
|S(t)u_0 - S(t)\bar{u}_0|_1 \le |u_0 - \bar{u}_0|_1,\ \forall u_0, \bar{u}_0 \in \overline{D(A)},\\
S(t+\tau)u_0 = S(t, S(\tau)u_0),\ \forall t, \tau > 0;\ u_0 \in \overline{D(A)},\\
\lim_{t\to 0} S(t)u_0 = u_0 \ \text{ in } L^1(\mathbb{R}^d).
\end{gathered}
$$

Moreover, by (2.7) and the exponential formula

$$
S(t)u_0 = \lim_{n\to\infty} \left(J^n_{\frac{t}{n}}\right)^{-n} u_0,\ \forall\, t \ge 0,
$$

it follows that

$$
|S(t)u_0|_\infty \le e^{\gamma t} |u_0|_\infty,\ \forall t \ge 0,
$$

where $\gamma = (1 + ||D| + (\operatorname{div} D)^-|^{\frac{1}{2}}_\infty)$. Hence, if $u_0 \in L^1 \cap L^\infty$, then $S(t)u_0 \in L^\infty((0,T)\times\mathbb{R}^d)$, $T > 0$. Furthermore, by (2.6) if $u_0 \ge 0$, then $S(t)u_0 \ge 0$.

Let us show now that $u = S(t)u_0$ is a Schwartz distributional solution, that is, (1.14) holds. By (1.12), we have

$$
\begin{aligned}
&\int_0^\infty dt \left(\int_{\mathbb{R}^d} \varphi(t,x)(u_h(t,x) - u_h(t-h,x)\right) dx\\
&\qquad + \int_{\mathbb{R}^d} (\varphi(t,x)\Psi(-\Delta)\beta(u_h(t,x)) - \nabla_x\varphi(t,x)\cdot D(x) b^*(u_h((x)))dx) = 0,\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall \varphi \in C_0^\infty([0,\infty)\times\mathbb{R}^d).
\end{aligned}
$$

This yields

$$
\begin{aligned}
&\frac{1}{h}\int_0^\infty dt \left(\int_{\mathbb{R}^d} u_h(t,x)(\varphi(t+h,x) - \varphi(t,x))\right) dx\\
&\qquad + \int_{\mathbb{R}^d} (\beta(u_h(t,x))\Psi(-\Delta)\varphi(t,x) - \nabla_x\varphi(t,x)\cdot D(x) b^*(u_h(t,x))dx)\\
&\qquad + \frac{1}{h}\int_0^h dt \int_{\mathbb{R}^d} u_0(x)\varphi(t,x)dx = 0,\ \forall \varphi \in C_0^\infty([0,\infty)\times\mathbb{R}^d).
\end{aligned}
$$

Taking into account that, by (1.12) and assumptions (i)–(iii), $\beta(u_h) \to \beta(u)$, $b^*(u_h) \to b^*(u)$ in $C([0,T]; L^1)$ as $h \to 0$ for each $t > 0$, we get that (1.14) holds.

We have, therefore, the following existence result for equation (1.1).

Theorem 2.4. *Assume* $s \in \left(\frac{1}{2}, 1\right)$ *and that Hypotheses* (i)–(v) *hold. Then, there is a* C_0*-semigroup of contractions* $S(t) : L^1 \to L^1$, $t \geq 0$, *such that for each* $u_0 \in \overline{D(A)}$, *which is* L^1 *if* $\beta \in C^\infty(\mathbb{R})$, $u(t, u_0) = S(t)u_0$ *is a mild solution to* (1.1). *Moreover, if* $u_0 \geq 0$, *a.e. on* $\mathbb{R}^d$,

$$u(t, u_0) \geq 0, \quad \text{a.e. on } \mathbb{R}^d,\ \forall\, t \geq 0, \tag{2.45}$$

$$\int_{\mathbb{R}^d} u(t, u_0)(x)dx = \int_{\mathbb{R}^d} u_0(x)dx, \quad \forall\, t \geq 0. \tag{2.46}$$

Moreover, u *is a distributional solution to* (1.1) *on* $[0, \infty) \times \mathbb{R}^d$. *Finally, if* $u_0 \in L^1 \cap L^\infty$, *then all above assertions remain true, if we drop the assumption* $\beta \in \mathrm{Lip}(\mathbb{R})$ *from Hypothesis* (i), *and additionally we have that* $u \in L^\infty((0, T) \times \mathbb{R}^d)$, $T > 0$.

3 The uniqueness of distributional solutions

In general, the mild solution u given by Theorem 2.4 is not unique because the operator A constructed in Lemma 2.3 dependends on the approximating operators $A_\varepsilon y \equiv \Psi(\varepsilon I - \Delta)\beta_\varepsilon(y) + \mathrm{div}(D_\varepsilon b_\varepsilon(y)y)$ and so $u = S(t)u_0$ may be viewed as a *viscosity-mild* solution to (1.1). However, as seen here, this mild solution is even unique in the much larger class of distributional solutions under the following hypotheses on β, b and D:

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

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

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

Namely, we have

Theorem 3.1. *Assume that* $0 < T < \infty$, $d \geq 2$, *and that Hypotheses* (j)–(jjj) *and* (v) *hold. 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* (1.14)*) 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 $y_1 \equiv y_2$. *If* $D \equiv 0$, *then Hypothesis* (j) *can be relaxed to*

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

Proof. (The proof is similar to that of Theorem 3.1 in [6], but it has to be adapted substantially.) Replacing, if necessary, the functions β and b by

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

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

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

$$\beta', b' \in C_b(\mathbb{R}),\ \ \beta' > \alpha_2 \in (0,\infty) \tag{3.2}$$

and, therefore, we have

$$\alpha_1|\beta(r)-\beta(\bar r)| \ \ \ge\ \ |b^*(r)-b^*(\bar r)|,\ \ \forall\, r,\bar r \in \mathbb{R}, \tag{3.3}$$

$$(\beta(r)-\beta(\bar r))(r-\bar r) \ \ \ge\ \ \alpha_3|\beta(r)-\beta(\bar r)|^2,\ \forall\, r,\bar r \in \mathbb{R}, \tag{3.4}$$

where $b^*(r) = b(r)r$, $\alpha_1 \ge 0$, and $\alpha_3 := |\beta'|_\infty^{-1}$. We set

$$\begin{gathered} \Phi_\varepsilon(y) = (\varepsilon I + \Psi(-\Delta))^{-1}y,\ \forall\, y \in L^2, \\ z = y_1 - y_2,\ w = \beta(y_1)-\beta(y_2),\ b^*(y_i) \equiv b(y_i)y_i,\ i=1,2. \end{gathered} \tag{3.5}$$

As seen in the Appendix, the operator Φ_ε is well defined. Moreover, it follows that $\Phi_\varepsilon : L^p \to L^p$, $\forall p \in [1,\infty]$ and

$$\varepsilon|\Phi_\varepsilon(y)|_p \le |y|_p,\ \ \forall y \in L^p,\ \varepsilon > 0. \tag{3.6}$$

We have

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

We set

$$z_\varepsilon = z * \theta_\varepsilon,\ \ w_\varepsilon = w * \theta_\varepsilon,\ \ \zeta_\varepsilon = (D(b^*(y_1)-b^*(y_2))) * \theta_\varepsilon, \tag{3.7}$$

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 $z_\varepsilon, w_\varepsilon, \zeta_\varepsilon, \Psi(-\Delta)w_\varepsilon, \operatorname{div}\zeta_\varepsilon \in L^2(0,T;L^2)$ and we have

$$(z_\varepsilon)_t+\Psi(-\Delta)w_\varepsilon+\operatorname{div}\zeta_\varepsilon=0 \text{ in } \mathcal{D}'(0,T;L^2). \tag{3.8}$$

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

$$(\Phi_\varepsilon(z_\varepsilon))_t=-\Psi(-\Delta)\Phi_\varepsilon(w_\varepsilon)-\operatorname{div}\Phi_\varepsilon(\zeta_\varepsilon)=0 \text{ in } \mathcal{D}'(0,T;L^2). \tag{3.9}$$

By (3.8), (3.9) it follows that $(z_\varepsilon)_t=\frac{d}{dt}\,z_\varepsilon$, $(\Phi_\varepsilon(z))_t=\frac{d}{dt}\,\Phi_\varepsilon(z_\varepsilon)\in L^2(0,T;L^2)$. 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 $[0,T]\ni t\mapsto \Phi_\varepsilon(z_\varepsilon(t))\in L^2$ are absolutely continuous. Moreover, it follows by (3.6) and (3.9) that

$$\Phi_\varepsilon(z_\varepsilon),\Phi_\varepsilon(w_\varepsilon)\in L^2(0,T;L^2). \tag{3.10}$$

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

$$\begin{aligned} h'_\varepsilon(t)&=2(z_\varepsilon(t),(\Phi_\varepsilon(z_\varepsilon(t)))_t)_2 \\ &=2(\varepsilon\Phi_\varepsilon(w_\varepsilon(t))-w_\varepsilon(t)-\operatorname{div}\Phi_\varepsilon(\zeta_\varepsilon(t)),z_\varepsilon(t))_2 \\ &=2\varepsilon(\Phi_\varepsilon(z_\varepsilon(t)),w_\varepsilon(t))_2+2(\nabla\Phi_\varepsilon(z_\varepsilon(t)),\zeta_\varepsilon(t))_2 \\ &\quad -2(z_\varepsilon(t),w_\varepsilon(t))_2,\ \text{a.e. } t\in(0,T). \end{aligned} \tag{3.11}$$

By (3.9)–(3.11) it follows that $t\to h_\varepsilon(t)$ has an absolutely continuous dt-version on $[0,T]$ which we shall consider from now on. Since, by (3.4), we have

$$(z_\varepsilon(t),w_\varepsilon(t))_2\ge\alpha_3|\,|w(t)|*\theta_\varepsilon|_2^2+\gamma_\varepsilon(t), \tag{3.12}$$

$$\gamma_\varepsilon(t):=(z_\varepsilon(t),w_\varepsilon(t))_2-(z(t),w(t))_2, \tag{3.13}$$

we get, by (3.3) and (3.10),

$$\begin{aligned} 0\ &\le h_\varepsilon(t)\le h_\varepsilon(0+)+2\varepsilon\int_0^t(\Phi_\varepsilon(z_\varepsilon(s)),w_\varepsilon(s))_2ds-2\alpha_2\int_0^t|w_\varepsilon(s)|_2^2ds \\ &+2\alpha_1|D|_\infty\int_0^t|\nabla\Phi_\varepsilon(z_\varepsilon(s))|_2|w_\varepsilon(s)|_2ds+2\int_0^t|\gamma_\varepsilon(s)|ds,\ \forall t\in[0,T]. \end{aligned} \tag{3.14}$$

Moreover, since $z\in L^\infty((0,T)\times\mathbb{R}^d)$, by (3.6) we have

$$\varepsilon|\Phi_\varepsilon(z_\varepsilon(t))|_\infty\le|z_\varepsilon(t)|_\infty\le|z(t)|_\infty,\quad \text{a.e. } t\in(0,T). \tag{3.15}$$

As $t\to\Phi_\varepsilon(z_\varepsilon(t))$ has an L^2 continuous version on $[0,T]$, there exists $f\in L^2$ such that $\lim\limits_{t\to 0}\Phi_\varepsilon(z_\varepsilon(t))=f$ in L^2. Furthermore, for every $\varphi\in C_0^\infty(\mathbb{R}^d)$, $s\in(0,T)$,

$$0\le h_\varepsilon(s)\le|\Phi_\varepsilon(z_\varepsilon(s))-f|_2|z_\varepsilon(s)|_2+|f-\varphi|_2|z_\varepsilon(s)|_2+|(\varphi*\theta_\varepsilon,z(s))_2|,$$

and so, by (3.1),

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

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

$$
h_\varepsilon(0+)=0. \tag{3.16}
$$

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

$$
\varepsilon\Phi_\varepsilon(z_\varepsilon(t))+\Psi(-\Delta)\Phi_\varepsilon(z_\varepsilon(t))=z_\varepsilon(t), \tag{3.17}
$$

we get that, for a.e. $t\in(0,T)$,

$$
\varepsilon|\Phi_\varepsilon(z_\varepsilon(t))|_2^2+|(\Psi(-\Delta))^{\frac12}\Phi_\varepsilon(z_\varepsilon(t))|_2^2=(z_\varepsilon(t),\Phi_\varepsilon(z_\varepsilon(t)))_2=h_\varepsilon(t), \tag{3.18}
$$

$$
\varepsilon|(\Phi_\varepsilon(z_\varepsilon(t),w_\varepsilon(t)))_2|\le\varepsilon|\Phi_\varepsilon(z_\varepsilon(t))|_\infty|w_\varepsilon(t)|_1\le|z(t)|_\infty|w(t)|_1. \tag{3.19}
$$

By (3.17), we have

$$
\mathcal{F}(\Phi_\varepsilon(z_\varepsilon(t)))=(\varepsilon+\Psi(|\xi|^2))^{-1}\mathcal{F}(z_\varepsilon(t)). \tag{3.20}
$$

Therefore, by Parseval's formula,

$$
\begin{aligned}
|\nabla\Phi_\varepsilon(z_\varepsilon(t))|_2^2 &= \int_{\mathbb{R}^d}\frac{|\mathcal{F}(z_\varepsilon(t))(\xi)|^2|\xi|^2}{(\varepsilon+\Psi(|\xi|^2))^2}\,d\xi,\quad \forall t\in(0,T),\\
h_\varepsilon(t) &= \int_{\mathbb{R}^d}\frac{|\mathcal{F}(z_\varepsilon(t))(\xi)|^2}{\varepsilon+\Psi(|\xi|^2)}\,d\xi,\qquad \forall t\in(0,T).
\end{aligned}
$$

Then, by (1.9) this yields for some $C\in(0,\infty)$ independent of ε

$$
\begin{aligned}
|\nabla\Phi_\varepsilon(z_\varepsilon(t))|_2^2 \le CR^{2(1-s)}&\int_{[|\xi|\le R]}\frac{|\mathcal{F}(z_\varepsilon(t))(\xi)|^2}{\varepsilon+\Psi(|\xi|^2)}\,d\xi\\
&+C\int_{[|\xi|\ge R]}|\mathcal{F}(z_\varepsilon(t))(\xi)|^2|\xi|^{2(1-2s)}d\xi\\
&\le CR^{2(1-s)}h_\varepsilon(t)+CR^{2(1-2s)}|z_\varepsilon(t)|_2^2,
\end{aligned} \tag{3.21}
$$

$\forall t\in(0,T),\ R>0$, because $2s\ge 1$.

We shall now prove that

$$\lim_{\varepsilon\to 0}\varepsilon(\Phi_\varepsilon(z_\varepsilon(t)),w_\varepsilon(t))_2=0,\ \text{ a.e. }\ t\in(0,T). \tag{3.22}$$

By (3.20) and (1.9) it follows for some $C\in(0,\infty)$ independent of ε

$$\begin{aligned}|(\Phi_\varepsilon(z_\varepsilon(t)),w_\varepsilon(t))_2|&=|(\mathcal{F}(\Phi_\varepsilon(z_\varepsilon(t)),\overline{\mathcal{F}}(w_\varepsilon(t)))_2|\\&\le C\int_{\mathbb{R}^d}\frac{|\mathcal{F}(z_\varepsilon(t))(\xi)|\,|\mathcal{F}(w_\varepsilon(t))(\xi)|}{\varepsilon+|\xi|^{2s}}\,d\xi\\&\le C\left(\int_{\mathbb{R}^d}\left|\frac{\mathcal{F}(z_\varepsilon(t))}{\varepsilon+|\xi|^{2s}}\right|^2d\xi\right)^{\frac12}|w_\varepsilon(t)|_2,\end{aligned}$$

and since

$$\frac{\mathcal{F}(z_\varepsilon(t))}{\varepsilon+|\xi|^{2s}}=\mathcal{F}((\varepsilon I+(-\Delta)^s)^{-1}z_\varepsilon(t)),\ t\in(0,T),$$

this yields

$$|(\Phi_\varepsilon(z_\varepsilon(t)),w_\varepsilon(t))_2|\le C|(\varepsilon I+(-\Delta)^s)^{-1}z_\varepsilon(t)|_2|w(t)|_2. \tag{3.23}$$

On the other hand, for each $f\in L^2(\mathbb{R}^d)$, by [6, Appendix] we have

$$(\varepsilon I+(-\Delta)^s)^{-1}f(x)=\int_{\mathbb{R}^d}g^s_\varepsilon(x-\xi)f(\xi)d\xi, \tag{3.24}$$

where

$$g^s_\varepsilon(x)=\int_0^\infty e^{-\varepsilon\tau}d\tau\int_0^\infty\frac{e^{-\frac{|x|^2}{4r}}}{(4\pi r)^{\frac d2}}\ \eta^s_\tau(dr),$$

and $(\eta^s_\tau)_{\tau\ge0}$ is the one-sided stable semigroup of order $s\in\left(\frac12,1\right)$.

By (A.4), (A.7) and (A.10) in [6], we have

$$\varepsilon\int_{\mathbb{R}^d}g^s_\varepsilon(x)dx=1, \tag{3.25}$$

$$g^s_\varepsilon\in L^\infty(\mathbb{R}^d\setminus B_R(0)),\ \forall R>0,\ \ g^s_\varepsilon(x)=\varepsilon^{\frac{d-2s}{2s}}\ g^s_1(\varepsilon^{\frac{1}{2s}}\,x), \tag{3.26}$$

where $B_R(0)$ is the ball of radius R around the origin in $\mathbb{R}^d$.

Then, by (3.24)–(3.26), we have via the Young inequality

$$
\begin{aligned}
\varepsilon|(\varepsilon I+(-\Delta)^s)^{-1}z_\varepsilon(t)|_2 &= \varepsilon|g^s_\varepsilon * z_\varepsilon(t)|_2 \le \varepsilon|g^s_\varepsilon * z_\varepsilon(t)|_\infty^{\frac12}|g^s_\varepsilon * z_\varepsilon(t)|_1^{\frac12}\\
&\le \varepsilon^{\frac12}|z_\varepsilon(t)|_1^{\frac12}\sup_{x\in\mathbb{R}^d}\left(\int_{\mathbb{R}^d} g^s_\varepsilon(x-\xi)|z_\varepsilon(t)(\xi)|d\xi\right)^{\frac12}\\
&\le \varepsilon^{\frac{d}{4s}}\sup_{x\in\mathbb{R}^d}\left(\int_{\mathbb{R}^d} g^s_1\left(\varepsilon^{\frac{1}{2s}}(x-\xi)\right)|z_\varepsilon(t)(\xi)|d\xi\right)^{\frac12}|z(t)|_1^{\frac12}\\
&\le C_\delta\,\varepsilon^{\frac{d}{4s}}|z(t)|_1\\
&\quad+\varepsilon^{\frac{d}{4s}}|z(t)|_\infty^{\frac12}\left(\int_{[\varepsilon^{\frac{1}{2s}}|x-\xi|\le\delta]} g^s_1\left(\varepsilon^{\frac{1}{2s}}(x-\xi)\right)d\xi\right)^{\frac12}|z(t)|_1^{\frac12}<\infty,
\end{aligned}
\tag{3.27}
$$

where $C_\delta=\sup\{g^s_1(\xi);\,|\xi|>\delta\}$. Now, letting first $\varepsilon\to0$ and then $\delta\to0$, (3.22) follows by (3.23) and (3.25).

Then, by (3.19) it follows that

$$
\lim_{\varepsilon\to0}\varepsilon\int_0^t(\Phi_\varepsilon(z_\varepsilon(s)),w_\varepsilon(s))_2ds=0,\ t\in[0,T]. \tag{3.28}
$$

Next, by (3.14), (3.16) and (3.21), we have

$$
\begin{aligned}
0\le h_\varepsilon(t) &\le 2\varepsilon\int_0^t|(\Phi_\varepsilon(z_\varepsilon(r)),w_\varepsilon(r))_2|dr-2\alpha_3\int_0^t|w_\varepsilon(r)|_2^2dr\\
&\quad+2\alpha_1|D|_\infty\int_0^t|\nabla\Phi_\varepsilon(z_\varepsilon(r))|_2|w_\varepsilon(r)|_2dr+2\int_0^t|\gamma_\varepsilon(r)|dr\\
&\le \eta_\varepsilon(t)+2\alpha_1|D|_\infty C^{\frac12}\int_0^t\left(R^{1-s}h_\varepsilon^{\frac12}(r)+R^{1-2s}|z_\varepsilon(r)|_2\right)|w_\varepsilon(r)|_2dr\\
&\quad-2\alpha_3\int_0^t|w_\varepsilon(r)|_2^2dr,\ \forall\,t\in[0,T],\ R>0,
\end{aligned}
$$

where

$$
\eta_\varepsilon(t):=2\varepsilon\int_0^t|(\Phi_\varepsilon(z_\varepsilon(r)),w_\varepsilon(r))_2|dr+2\int_0^t|\gamma_\varepsilon(r)|dr.
$$

This yields

$$
\begin{aligned}
0\le h_\varepsilon(t)\le\eta_\varepsilon(t)+2\alpha_1|D|_\infty C^{\frac12}\bigg(R^{2(1-s)}\lambda\int_0^t h_\varepsilon(r)dr+\int_0^t\Big(R^{1-2s}|z_\varepsilon(r)|_2^2\\
+\Big(\frac{1}{4\lambda}+R^{1-2s}\Big)|w_\varepsilon(r)|_2^2\Big)dr\bigg)-2\alpha_3\int_0^t|w_\varepsilon(r)|_2^2dr,\ \forall\lambda>0,\ R>0.
\end{aligned}
\tag{3.29}
$$

Taking into account that, by (3.2),

$$|z(t)|_2 \le \alpha_2^{-1}|w(t)|_2, \quad \forall t \in (0,T), \tag{3.30}$$

we have

$$|z_\varepsilon(t)|_2 \le \alpha_2^{-1}|w_\varepsilon(t)|_2 + \nu_\varepsilon(t),\ \forall t \in (0,T), \tag{3.31}$$

where $\nu_\varepsilon(t) := \alpha_2^{-1}|w(t) - w_\varepsilon(t)|_2$. Then, for

$$\widetilde{\eta}_\varepsilon(t) := \eta_\varepsilon(t) + \int_0^t \nu_\varepsilon(r)dr,$$

we get for λ, $R > 0$, large enough

$$0 \le h_\varepsilon(t) \le \widetilde{\eta}_\varepsilon(t) + C\int_0^t h_\varepsilon(r)dr, \quad \text{for } t \in [0,T], \tag{3.32}$$

where $C > 0$ is independent of ε and $\lim\limits_{\varepsilon\to 0} \widetilde{\eta}_\varepsilon(t) = 0$ for all $t \in [0,T]$.

By (3.32), it follows that

$$0 \le h_\varepsilon(t) \le \widetilde{\eta}_\varepsilon(t)\exp(Ct),\ \forall\, t \in [0,T]. \tag{3.33}$$

This implies that $h_\varepsilon(t) \to 0$ as $\varepsilon \to 0$ for every $t \in [0,T]$, hence by (3.18) the left hand side of (3.17) converges to zero in S'. Thus, $0 = \lim\limits_{\varepsilon\to 0} z_\varepsilon(t) = z(t)$ in S' for a.e. $t \in (0,T)$, which implies $y_1 \equiv y_2$. If $D \equiv 0$, we see by (3.14) and (3.16) that $0 \le h_\varepsilon(t) \le \eta_\varepsilon(t)$, $\forall\, t \in (0,T)$, and so by (3.28) the conclusion follows without invoking that $\beta' > 0$, which was only used to have (3.30). □

Remark 3.2. It should be noted that Theorem 3.1 is compatible with Theorem 2.4 because under Hypotheses (i)–(v) there is an $L^1(0,T;\mathbb{R}^d) \cap L^\infty(0,T;\mathbb{R}^d)$ distributional solution to (1.1).

Similarly as Theorem 3.1, one also obtains linearized uniqueness for equation (1.14).

Theorem 3.3. *Under the assumptions of Theorem* 3.1*, let* $T > 0$*,* $u \in L^\infty((0,T)\times\mathbb{R}^d)$ *and let* $y_1, y_2 \in L^\infty((0,T)\times\mathbb{R}^d)$ *with* $y_1 - y_2 \in L^1((0,T)\times\mathbb{R}^d) \cap L^\infty(0,T;L^2)$ *be two distributional solutions to the equation*

$$\begin{aligned} &y_t + \Psi(-\Delta)\left(\frac{\beta(u)}{u}y\right) + \operatorname{div}(yDb(u)) = 0 \text{ on } (0,T)\times\mathbb{R}^d,\\ &y(0) = u_0, \end{aligned} \tag{3.34}$$

where u_0 *is a measure of finite variation on* $\mathbb{R}^d$ *and* $\frac{\beta(0)}{0} := \beta'(0)$*. If* (3.1) *holds, then* $y_1 \equiv y_2$*.*

Proof. The proof is essentially the same as that of Theorem 3.1 and, therefore, it will be sketched only. We note first that, by Hypotheses (j)–(jjj),

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

and, by (3.2)–(3.5),

$$|Db(u)|_\infty \le C_1 \left|\frac{\beta(u)}{u}\right|_\infty \le C_2 \left|\frac{\beta(u)}{u}\right| \ge \alpha_2,\ \text{a.e. in } (0,T)\times\mathbb{R}^d.$$

If $z = y_1 - y_2,\ w = \frac{\beta(u)}{u}(y_1 - y_2)$, we have therefore

$$wz \ge \alpha_2 |w|^2, \qquad \text{a.e. on } (0,T)\times\mathbb{R}^d, \tag{3.35}$$

$$|Db(u)z| \le C_2 |w|, \qquad \text{a.e. on } (0,T)\times\mathbb{R}^d. \tag{3.36}$$

We have

$$z_t + \Psi(-\Delta)w + \operatorname{div}(Db(u)z) = 0 \text{ in } (0,T)\times\mathbb{R}^d,$$

and this yields (see (3.8))

$$(z_\varepsilon)_t + \Psi(-\Delta)w_\varepsilon + \operatorname{div}\zeta_\varepsilon = 0 \text{ in } (0,T)\times\mathbb{R}^d, \tag{3.37}$$

where $z_\varepsilon, w_\varepsilon$ are as in (3.7) and $\zeta_\varepsilon = (D(b(u))z) * \theta_\varepsilon$. If Φ_ε is given by (3.5), by (3.37) we get (3.9) and, if $h_\varepsilon(t) = (\Phi_\varepsilon(z_\varepsilon(t), z_\varepsilon(t)))$, then (3.11) follows and so, by (3.35) we get also in this case the estimates (3.12) and (3.14). From now on the proof is exactly the same as that of Theorem 3.1. Namely, one gets that $h_\varepsilon(0_+) = 0$ and also that (3.21) and (3.28) hold. Finally, one gets (3.29) and, taking into account (3.30)–(3.31), one obtains that (3.33) holds and so $z \equiv 0$, as claimed. □

4 Applications to corresponding nonlinear martingale problems

In this section we fix $T \in (0,\infty)$ and need the following additional hypotheses on Ψ:

(vi) $\displaystyle\int_1^\infty \log s\,\mu(ds) < \infty.$

This is e.g. fulfilled if μ is as in (1.8).

4.1 Existence

Assume that Hypotheses (i)–(vi) hold and consider the nonlocal Kolmogorov operator corresponding to (1.1), i.e.,

$$K_{u(t)}f(x) = \frac{\beta(u(t,x))}{u(t,x)}(\Psi(-\Delta)f)(x) + b(u(t,x))D(x)\cdot\nabla f(x),\ x\in\mathbb{R}^d,\quad (4.1)$$

where $f \in C_c^2 := C_c^2(\mathbb{R}^d)$ and u is the solution to (1.1) from Theorem 2.4 with initial condition $u_0 \in L^1 \cap L^\infty$.

By Theorem 13.6 in [22] it follows that, for all $f \in C_c^2$,

$$\begin{aligned}(\Psi(-\Delta)f)(x) &= \int_{(0,\infty)}\int_{\mathbb{R}^d}(f(x)-f(x+z))\frac{1}{\sqrt{4\pi t}}\,e^{-\frac{|z|^2}{4t}}\,dz\mu(dt)\\ &= \int_{\mathbb{R}^d}(f(x)-f(x+z))\nu(|z|)dz,\end{aligned}\quad (4.2)$$

where

$$\nu(r) = \int_{(0,\infty)}\frac{1}{\sqrt{4\pi t}}\,e^{-\frac{r^2}{4t}}\,\mu(dt),\ r\in(0,\infty).$$

Then, since $\frac{\beta(u)}{u} \in L^\infty((0,T)\times\mathbb{R}^d)$, by (vi) it is easily seen that $K_{u(t)}$ is a Kolmogorov operator of the type considered in Section 1.2 in [21], which satisfies condition (1.18) in [21]. Thus, by Theorem 1.5 (”*superposition principle*”) in [21] and Remark 1.6 in [21], we get

Theorem 4.1. *Under Hypotheses* (i)–(vi)*, there exists a probability measure* $\mathbb{P}$ *on the Skorohod space* $D([0,T];\mathbb{R}^d)$*, which is a solution to the martingale problem corresponding to* $(K_{u(t,\cdot)}, C_c^2)$ *in the sense of Definition* 1.3 *in* [21] *with one dimensional time marginal densities given by* $u(t,x)$, $t\in[0,T]$, $x\in\mathbb{R}^d$*, i.e. for the canonical process* X_t, $t\in[0,T]$*, defined by* $X_t(w)=w(t)$, $w\in D([0,T];\mathbb{R}^d)$*, we have*

$$(\mathbb{P}\circ X_t^{-1})(dx) = u(t,x)dx,\ t\in[0,T].\quad (4.3)$$

Remark 4.2. The probability measure $\mathbb{P}$ in Theorem 4.1 is called a solution to the nonlinear martingale problem corresponding to $(K_\square, C_c^2)$, because it solves the linear martingale problem corresponding to $(K_{\mathcal{L}_{X_t}}, C_c^2)$, where $\mathcal{L}_{X_t}$ is its own one dimensional time marginal law density and $K_{\mathcal{L}_{X_t}}$ is defined as in (4.1) with $\mathcal{L}_{X_t}$ replacing $u(t,\cdot)$, $t\in[0,T]$. We refer to the pioneering work [14], where such nonlinear martingale problems were studied for local Kolmogorov operators.

4.2 Uniqueness

Theorem 4.3. *Assume that Hypotheses* (j)–(jjj),(v),(vi) (*respectively,* (j)′,(jj), (jjj),(v),(vi) *if* $D \equiv 0$) *hold. Let* $\mathbb{P}$, $\widetilde{\mathbb{P}}$ *be probability measures on* $D([0,T];\mathbb{R}^d)$ *such that their time marginals,* $\mathbb{P} \circ X_t^{-1}$, $\widetilde{\mathbb{P}} \circ X_t^{-1}$ *have densities* $\mathcal{L}_{X(t)}$, *and* $\widetilde{\mathcal{L}}_{X_t}$ *respectively, w.r.t. Lebesgue measure for all* $t \in [0,T]$ *such that*

$$((t,x) \to \mathcal{L}_{X_t}(x)),\ ((t,x) \to \widetilde{\mathcal{L}}_{X_t}(x)) \in L^\infty((0,T) \times \mathbb{R}^d). \tag{4.4}$$

If $\mathbb{P}$ *and* $\widetilde{\mathbb{P}}$ *are solutions to the nonlinear martingale problem* $(K_\square, C_c^2)$, *i.e.* (*see Remark* 4.2), *they are solutions to the linear martingale problems corresponding to* $(K_{\mathcal{L}_{X_t}}, C_c^2)$, $(K_{\mathcal{L}_{\widetilde{X}_t}}, C_c^2)$, *respectively, in the sense of* [21], *Definition* 1.3, *then* $\mathbb{P} = \widetilde{\mathbb{P}}$.

Proof. Clearly, by Dynkin's formula, both

$$\mu_t(dx) := \mathcal{L}_{X_t}(x)dx \ \text{ and } \ \widetilde{\mu}_t(dx) := \mathcal{L}_{\widetilde{X}_t}(x)dx,\ t \in [0,T],$$

solve the Fokker–Planck equation (1.14) with the same initial condition $u_0(dx) := u_0(x)dx$, hence satisfy (3.1) with $y_1(t) := \mathcal{L}_{X_t}$ and $y_2(t) := \mathcal{L}_{\widetilde{X}_t}$. Hence, by Theorem 3.1

$$\mathcal{L}_{X_t} = \mathcal{L}_{\widetilde{X}_t},\ \text{for all } t \geq 0,$$

since $t \mapsto \mathcal{L}_{X_t}(x)dx$ and $t \mapsto \mathcal{L}_{\widetilde{X}_t}(x)dx$ are both narrowly continuous and are probability measures for all $t \geq 0$, so both are in $L^\infty(0,T;L^1 \cap L^\infty) \subset L^\infty(0,T;L^2)$.

Now, fixing $\mathcal{L}_{X_t}$ from above, consider the linear Fokker–Planck equation

$$\begin{aligned} &v_t + \Psi(-\Delta)\left(\frac{\beta(\mathcal{L}_{X_t})}{\mathcal{L}_{X_t}}\, v\right) + \operatorname{div}(Db(u)v) = 0, \\ &v(0,x) = u_0(x), \end{aligned} \tag{4.5}$$

again in the weak (distributional) sense analogous to (1.14). Then, by Theorem 3.3 we conclude that $\mathcal{L}_{X_t}$, $t \in [0,T]$, is the unique solution to (4.5) in $L^\infty(0,T;L^1 \cap L^\infty)$. Clearly, both $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ solve the (linear) martingale problem with initial condition $u_0(dx) := u_0(x)dx$ corresponding to $(K_{\mathcal{L}_{X_t}}, C_c^2)$. Since the above is true for all $u_0 \in L^1 \cap L^\infty$, and also holds when we consider (1.1), resp. (4.5), with start in $s > 0$ instead of zero, it follows by exactly the same arguments as in the proof of Lemma 2.12 in [23] that $\mathbb{P} = \widetilde{\mathbb{P}}$. $\square$

Theorem 4.4. *For $s \in [0,\infty)$ and $\zeta \in \mathcal{Z} := \{\zeta \equiv \zeta(x)dx \mid \zeta \in L^1 \cap L^\infty$, $\zeta \geq 0$, $|\zeta|_1 = 1\}$, let $\mathbb{P}_{(s,\zeta)}$ denote the solution to the nonlinear martingale problem corresponding to $(K_\square, C_c^2)$ with the initial condition ζ at the initial time s from Theorems* 4.1 *and* 4.3*. Then, $\mathbb{P}_{(s,\zeta)}$, $(s,\zeta) \in [0,\infty) \times \mathcal{Z}$, form a nonlinear Markov process in the sense of Definition* 2.1 *in* [18]*, i.e. in the sense of McKean* [16]*.*

Proof. The assertion follows from Corollary 3.8 in [18] (see also Example (iii) in Section 4.2 of [18] for the special case with $\Psi(r) := r^s$, $s \in \left(\frac{1}{2}, 1\right)$). □

Remark 4.5. Equation (4.3) in Theorem 4.1 says that our solution u of (1.1) from Theorem 2.4 is the one dimensional time marginal law density of a cadlàg nonlinear Markov process. This realizes McKean's vision formulated in [16] for solutions to nonlinear parabolic PDEs, namely to identify the solutions to the latter as one-dimensional time marginal law densities of a nonlinear Markov process. So, our results show that this is indeed also possible for nonlocal PDEs of type (1.1).

Appendix: Representation and properties of the integral kernel of $(\varepsilon I + \Psi(-\Delta))^{-1}$

Let $\varepsilon > 0$. We have, for $u \in L^2,\ \xi \in \mathbb{R}^d$,

$$\begin{aligned}
\mathcal{F}((\varepsilon I + \Psi(-\Delta))^{-1}u)(\xi) &= \frac{1}{\varepsilon + \Psi(|\xi|^2)}\ \mathcal{F}u(\xi) \\
&= \int_0^\infty e^{-\varepsilon t}\ e^{-t\Psi(|\xi|^2)} dt\, \mathcal{F}u(\xi) \\
&\overset{(1.7)}{=\!=} \int_0^\infty e^{-\varepsilon t} \int_0^\infty e^{-r|\xi|^2} \eta_t^\Psi(dr) dt\, \mathcal{F}u(\xi) \\
&= (2\pi)^{\frac{d}{2}} \int_0^\infty e^{-\varepsilon t} \int_0^\infty \mathcal{F}(p_r)(\xi) \eta_t^\Psi(dr) dt\, \mathcal{F}u(\xi),
\end{aligned}$$

where

$$p_r(x) := \frac{1}{(4\pi r)^{\frac{d}{2}}}\ e^{-\frac{1}{4r}|x|^2},\ x \in \mathbb{R}^d.$$

Hence, defining

$$g_\varepsilon^\Psi(x) := \int_0^\infty e^{-\varepsilon t} \int_0^\infty p_r(x) \eta_t^\Psi(dr) dt,\ x \in \mathbb{R}^d, \tag{A.1}$$

we have

$$(\varepsilon I + \Psi(-\Delta))^{-1} u = g_\varepsilon^\Psi * u.$$

Since $\eta_t^\Psi,\ t \geq 0$, are probability measures, we have

$$\varepsilon \int_{\mathbb{R}^d} g_\varepsilon^\Psi dx = 1, \quad \forall \varepsilon > 0. \tag{A.2}$$

Acknowledgements. This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project ID 317210226 – SFB 1283 and by a grant of the Ministry of Research, Innovation and Digitization, CNCS–UEFISCDI project PN-III-P4-PCE-2021-0006, within PNCDI III. A part of this work was done during a very pleasant stay of the third named author at the University of Madeira as a guest of the second named author.

References

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

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

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

[4] Barbu, V., Röckner, M., *The evolution to equilibrium of solutions to nonlinear Fokker-Planck equations*, *Indiana Univ. Math. J.*, 72 (1) (2023), 89-131.

[5] Barbu, V., Röckner, M., Uniqueness for nonlinear Fokker-Planck equations and for McKean–Vlasov SDEs: the degenerate case, *J. Funct. Anal.*, 285 (4), 2023.

[6] Barbu, V., Röckner, M., Nonlinear Fokker–Planck equations with fractional Laplacian and McKean–Vlasov SDEs with Lévy noise, arXiv:2210.05612v2 [math.PR].

[7] 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.

[8] Carillo, J.A., Entropy solutions for nonlinear degenerate problems, *Archives Rat. Mech. Anal.*, 147 (1999), 269-361.

[9] Carmona, R., Delarue, F., *Probabilistic Theory of Mean Field Games with Applications*, I–II, Springer, 2017.

[10] De Pablo, A., Quiros, F., Rodriguez, A., Vasquez, J.L., A general fractional porous medium equation, *Comm. Pure Appl. Math.*, vol. 65 (9) (2012), 1242-1284.

[11] De Pablo, A., Quiros, F., Rodriguez, A., Vasquez, J.L., A fractional porous medium equation, *Adv. Math.*, 226 (2) (2010), 1378-1409.

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

[13] Fukushima, M., Oshima, Y., Takeda, M., *Dirichlet Forms and Symmetric Markov Processes*, de Gruyter, 2011, x+489 pp.

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

[15] Ma, Zhi Ming, Röckner, M., Introduction to the theory of (nonsymmetric) Dirichlet forms, *Universitext. Springer Verlag, Berlin*, 1992, vi+209 pp.

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

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

[18] Rehmeier, M., Röckner, On nonlinear Markov processes in the sense of McKean, arXiv:2212.12424v2 [math.AP]

[19] Ren, P., Röckner, M., Wang, F.Y., Linearization of nonlinear Fokker–Planck equations and applications, *J. Differential Equations*, 322 (2022), 1-37.

[20] Röckner, M., Wu, W., Xie, Y., Stochastic generalized porous media equations over σ-finite measure spaces with non-continuous diffusivity function, arXiv:2107.09878v2 (2021).

[21] Röckner, M., Xie, L., Zhang, X., Superposition principle for nonlocal Fokker–Planck–Kolmogorov operators, *Probab. Theory Rel. Fields*, 178 (3-4) (2020), 699-733.

[22] Schilling, R., Song, R., Vondraček, Z., *Bernstein Functions*, de Gruyter, 2012.

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

[24] Vásquez, J.L., Nonlinear diffusion with fractional Laplacian operators. In *Nonlinear partial differential equations. The Abel Symposium 2010*, pp. 271-298. Abel Symposia, 7. Berlin. Heidelberg, 2012. doi:10.10007/978-3-642-25361-4_15.