Weighted L^1-semigroup approach for nonlinear Fokker–Planck equations and generalized Ornstein–Uhlenbeck processes

Marco Rehmeier*

November 13, 2023

Abstract

For the nonlinear Fokker–Planck equation

$$\partial_t u = \Delta\beta(u) - \nabla\Phi \cdot \nabla\beta(u) - \mathrm{div}_\varrho\big(D(x)b(u)u\big), \quad (t,x) \in (0,\infty) \times \mathbb{R}^d,$$

where $\varrho = \exp(-\Phi)$ is the density of a finite Borel measure and $\nabla\Phi$ is unbounded, we construct mild solutions with bounded initial data via the Crandall–Liggett semigroup approach in the weighted space $L^1(\mathbb{R}^d, \mathbb{R}; \varrho dx)$. By the superposition principle, we lift these solutions to weak solutions to the corresponding McKean–Vlasov SDE, which can be considered a model for generalized nonlinear perturbed Ornstein–Uhlenbeck processes. Finally, for these solutions we prove the nonlinear Markov property in the sense of McKean.

Keywords: Nonlinear Fokker–Planck equation, mild solution, weighted Sobolev space, McKean–Vlasov stochastic differential equation, nonlinear Markov Process
2020 MSC: 35Q84, 35K55, 46B25, 60H30, 60J25

Declaration of interest: none.

1 Introduction

We study the nonlinear parabolic Fokker–Planck equation on $\mathbb{R}^d$, $d \geq 1$,

$$\partial_t u = \Delta\beta(u) - \nabla\Phi \cdot \nabla\beta(u) - \mathrm{div}_\varrho\big(D(x)b(u)u\big), \quad (t,x) \in (0,\infty) \times \mathbb{R}^d, \tag{1.1}$$

where $\varrho : \mathbb{R}^d \to \mathbb{R}$ is the density of a finite Borel measure, $\beta : \mathbb{R} \to \mathbb{R}$ is increasing, $D : \mathbb{R}^d \to \mathbb{R}^d$ and $b : \mathbb{R} \to \mathbb{R}$ are bounded, $\nabla\Phi = \nabla \log \varrho$, and div_ϱ denotes the dual operator to the gradient in $L^2(\mathbb{R}^d, \mathbb{R}; \varrho dx)$ (please see Hypothesis 1 in Section 2 for the precise assumptions). Our aim is twofold: First, we apply the well-known Crandall-Liggett semigroup approach in the weighted space $L^1(\mathbb{R}^d, \mathbb{R}; \varrho dx)$ to construct mild solutions $u \in C\big([0,\infty), L^1(\mathbb{R}^d, \mathbb{R}; \varrho dx)\big)$ to (1.1) (see Definition 2.2). Secondly, we show that these solutions, when started from probability density initial data, are the curves of $1D$-time marginals of probabilistically weak solutions to the associated nonlinear McKean–Vlasov stochastic differential equation (SDE)

$$dX_t = \left[D(X_t)b\big(v(t,X_t)\varrho^{-1}(X_t)\big) - \frac{\beta(v(t,X_t)\varrho^{-1}(X_t))}{v(t,X_t)\varrho^{-1}(X_t)}\nabla\Phi(X_t)\right]dt + \sqrt{2\frac{\beta(v(t,X_t)\varrho^{-1}(X_t))}{v(t,X_t)\varrho^{-1}(X_t)}}dB_t, \tag{1.2}$$

$\mathcal{L}(X_t) = v(t,x)dx.$

*Faculty of Sciences, Scuola Normale Superiore Pisa, Italy. E-Mail: mrehmeier@math.uni-bielefeld.de

Here $\mathcal{L}(X)$ denotes the distribution of a random variable X, and B is an $\mathbb{R}^d$-Brownian motion. We also show that these solutions to (1.2) constitute a nonlinear Markov process in the sense of McKean, see [11, 12].

Related results. Existence, uniqueness, asymptotic behavior and probabilistic representation of solutions to nonlinear Fokker–Planck equations, in particular of type

$$\partial_t u = \Delta\beta(u) - \mathrm{div}\,\big(B(x,u)u\big), \tag{1.3}$$

where $B : \mathbb{R}^{d+1} \to \mathbb{R}$ is bounded, have frequently been investigated in recent years, see for instance [3, 2, 5, 6] and the references therein. In [2], for $B(x,r) = D(x)b(r)$ and under appropriate assumptions on the coefficients, the Crandall–Liggett semigroup approach in $L^1 = L^1(\mathbb{R}^d, \mathbb{R}; dx)$ is applied to construct an L^1-contraction semigroup $S(t), t \geq 0$, of mild solutions $t \mapsto S(t)(u_0)$, where the initial data u_0 are from L^1. In fact, in the same paper, this result is even extended to finite Radon measures as initial data, and an $L^1 - L^\infty$-regularization result is obtained, i.e. solutions, even those equal to degenerate measures at $t = 0$, are $L^1 \cap L^\infty$-valued for $t > 0$. Since these solutions preserve non-negativity and total mass, starting from probability initial data yields solution curves of probability measure densities. By the Ambrosio–Figalli–Trevisan superposition principle (see [14],[3]), such probability solutions are the $1D$-time marginals of solutions to the McKean–Vlasov SDE

$$dX_t = D(X_t)b(u(t,X_t))dt + \sqrt{2\frac{\beta(u(t,X_t))}{u(t,X_t)}}dB_t.$$

Solutions to the latter equation, at least if D is a gradient vector field, are called nonlinear distorted Brownian motion, see [13]. We would also like to mention further recent existence result obtained via the L^1-semigroup approach, for instance [4, 7], where the authors solve equations of type (1.3), but Δ is replaced by a nonlocal operator, e.g. a fractional Laplacian.

Weighted semigroup approach. In the present paper, we initiate a similar program for (1.1), but, in contrast to all previously mentioned works, in the weighted space $L^1(\mathbb{R}, \mathbb{R}; \varrho dx)$. Our main result can be summarized as follows (please see Theorem 2.4 for the precise statement): If Hypothesis 1 below is satisfied, for any $u_0 \in L^1(\mathbb{R}^d, \mathbb{R}; \varrho dx) \cap L^\infty$ there is a mild solution $S(t)u_0, t \geq 0$, to (1.1), and $(S(t))_{t\geq 0}$ is a semigroup of $L^1(\mathbb{R}^d, \mathbb{R}; \varrho dx)$-contractions. The proof is inspired by the one in [2]. One advantage of our approach is that, compared to the standard approach, it allows to consider nonlinear unbounded first-order perturbations, which are not of divergence–type. Indeed, a main example for $\nabla\Phi$ is $\Phi(x) = \frac{-|x|^2}{2}$, $x \in \mathbb{R}^d$. Equation (1.1) can be treated by the semigroup approach in $L^1(\mathbb{R}^d, \mathbb{R}; \varrho dx)$, since the operator $Lf := \Delta f - \nabla\Phi \cdot \nabla f$ is symmetric on $L^2(\mathbb{R}^d, \mathbb{R}; \varrho)$. In fact, ϱdx is an infinitesimally invariant measure for L, i.e. $L^*\varrho dx = 0$, where L^* denotes the formal dual operator of L. The pair $(L, \varrho dx)$ can be considered a generalization of (Δ, dx). In this general setting we lose, however, very nice mapping properties of the fundamental solution to Δ in Marcinkiewicz spaces. Hence, so far we are only able to prove mild existence for initial data u_0 from $L^1(\mathbb{R}^d, \mathbb{R}; \varrho dx)$ which are also bounded. We can remove the boundedness restriction on u_0 when β is globally Lipschitz, see Proposition 2.5. The case of more general initial data, e.g. degenerate finite Borel measures, is not addressed in the present paper.

Generalized Ornstein–Uhlenbeck processes. As said before, a second aspect of this work is the probabilistic representation of solutions to (1.1). Since our mild solutions $t \mapsto S(t)u_0$ are also distributional solutions (see Def.(2.3)), by the aforementioned superposition principle, we obtain weak solutions to (1.2) with $1D$-time marginals $(S(t)u_0)_{t\geq 0}$. Solutions to (1.2) can be considered *generalized nonlinear perturbed Ornstein–Uhlenbeck processes*, since for $D = 0$, $\beta(r) = \frac{\sigma^2}{2}r$ and $\Phi(x) = -\frac{|x|^2}{2}$, the SDE becomes the classical Ornstein–Uhlenbeck equation. We stress that both (1.1) and (1.2) have nonlinear, irregular coefficients (of so-called Nemytskii-type, i.e. the dependence of the coefficients on a measure is pointwise via its density w.r.t. Lebesgue measure).

Finally, we prove that the laws of the solutions to (1.2) with $1D$-time marginals $(S(t)u_0)_{t\geq 0}$ constitute a *nonlinear Markov process*. Such processes were suggested by McKean in 1966 [11] and were recently revisited in [12]. McKean's vision was to represent solutions to nonlinear PDEs as the $1D$-time marginals of stochastic processes satisfying a nonlinear Markov property. Our results in Section 4 are a contribution to his vision for equation (1.1), which to the best of our knowledge has not been considered in this direction before.

Organization. The paper is organized as follows. In Section 2, we state our assumptions on the coefficients, recall the notion of mild and distributional solutions to (1.1) and state the new existence results, Theorem 2.4 and Proposition 2.5. Moreover, we state the key lemma towards these results, Lemma 2.6. Section 3 is devoted to the proof of this lemma. In Section 4, we lift our solutions to (1.1) to solutions to (1.2) and prove the nonlinear Markov property.

Notation. We set $\mathbb{N}_0 := \{0, 1, 2, ...\}$. The usual Euclidean norm and scalar product on $\mathbb{R}^k$, $k \geq 1$, are $|\cdot|$ and " $\cdot$ ". For an $\mathbb{R}$-valued function f, we write $f^+ := \max(f, 0)$ and $f^- := -\min(f, 0)$. If $(A, D(A))$, is an operator, $R(A)$ denotes its range. If X is a topological space, we denote its Borel σ-algebra by $\mathcal{B}(X)$.

For $\varrho : \mathbb{R}^d \to \mathbb{R}_+$ measurable and $p \in [1, \infty)$, we denote the usual spaces of p-fold ϱdx-integrable vector fields from $\mathbb{R}^d$ to $\mathbb{R}^k$ by $L^p(\mathbb{R}^d, \mathbb{R}^k; \varrho)$, and their usual norms by $|\cdot|_{p,\varrho}$. If $k = 1$, we simply write $L^p(\varrho)$, and $L^p(\mathbb{R}^d, \mathbb{R}^k)$ (L^p, for $k = 1$), if $\varrho \equiv 1$. In the Hilbert space case $p = 2$, the scalar product is denoted $\langle \cdot, \cdot \rangle_{2,\varrho}$. For the corresponding local spaces, we always suppress the dependence on ϱ, since we will only consider the case $0 < \varrho \leq 1$. For the same reason, for $p = \infty$ we write $L^\infty(\mathbb{R}^d, \mathbb{R}^k)$ (L^∞, for $k = 1$) independent of the measure ϱdx, and $|\cdot|_\infty$ for its norm (which is independent of ϱ). $(W^{m,p}(\varrho), |\cdot|_{W^{m,p}(\varrho)})$ are the usual Sobolev spaces of m-times weakly differentiable functions $g : \mathbb{R}^d \to \mathbb{R}$ with p-fold ϱdx-integrable derivatives of order up to m. If $\varrho \equiv 1$, we simply write $W^{m,p}$. For $p = 2$, we write $H^m(\varrho)$ and H^m instead. For the latter spaces, we denote by $H^{-m}(\varrho)$ and H^{-m} their topological dual spaces, respectively.

For $m \in \mathbb{N}_0 \cup \{\infty\}$, we denote the spaces of m-times differentiable functions $g : U \subseteq \mathbb{R}^k \to \mathbb{R}$ (with bounded or compactly supported derivatives up to order m) by $C^m(U)$ ($C^m_b(U), C^m_c(U)$, respectively). For $k = d$ and $U = \mathbb{R}^d$, we simply write C^m, C^m_b, C^m_c. More generally, for a topological space X we let $C(U, X)$ denote the space of continuous maps $g : U \to X$. The usual space of distributions on $\mathbb{R}^d$ (in duality with C^∞_c) is $\mathcal{D}'$.

$\mathcal{M}_b$ is the set of non-negative, finite Borel measures on $\mathbb{R}^d$, and $\mathcal{P}$ its subset of Borel probability measures. The weak topology on $\mathcal{M}_b$ is the initial topology of the maps $\mu \mapsto \int_{\mathbb{R}^d} \varphi \, d\mu$, $\varphi \in C_b$. We also set

$$\mathcal{P}_\infty(\varrho) := \left\{ u \in L^1(\varrho) \cap L^\infty, u \geq 0, \int_{\mathbb{R}^d} u \, \varrho dx = 1 \right\}.$$

Finally, we stress that by f^{-1} we usually denote the inverse of a map f, but for the case $f = \varrho$ with ϱ as fixed at the beginning of the next section, we mean $\varrho^{-1} := \frac{1}{\varrho}$.

2 Mild solutions to nonlinear Fokker–Planck equations via weighted semigroup approach

Throughout, we fix an integer $d \geq 1$. Let $\Phi : \mathbb{R}^d \to \mathbb{R}$, $\varrho := e^{-\Phi}$ and consider the nonlinear Fokker–Planck equation (1.1) with Cauchy initial condition $u(0) = u_0$, where $u_0 : \mathbb{R}^d \to \mathbb{R}$. Here, $-\text{div}_\varrho$ denotes the dual operator to the gradient ∇ in $L^2(\varrho)$, i.e. for $F \in L^2(\mathbb{R}^d, \mathbb{R}^d; \varrho)$, $\text{div}_\varrho F \in H^{-1}(\varrho)$ is defined via

$$-\int_{\mathbb{R}^d} (\text{div}_\varrho F) g \, \varrho dx = \int_{\mathbb{R}^d} F \cdot \nabla g \, \varrho dx, \quad \forall g \in H^1(\varrho).$$

A heuristic integration by parts on the right-hand side suggests $\text{div}_\varrho F = \text{div}\, F + F \cdot \nabla(\log \varrho)$, where div denotes the usual divergence. The latter equality holds in $L^2(\varrho)$, if all terms are well-defined in $L^2(\varrho)$. Note that ϱdx is Fomin differentiable, if $\nabla(\log \varrho) = -\nabla \Phi \in L^1(\mathbb{R}^d, \mathbb{R}^d; \varrho)$. We work under the following assumptions.

Hypothesis 1

(H1) $\Phi \in C^2$ is non-negative and convex with $\lim_{|x|\to\infty} \Phi(x) = \infty$, and $\nabla\Phi \in L^1(\varrho)$.

(H2) $\beta \in C^2(\mathbb{R})$, $\beta'(r) > 0$ for $r \neq 0$, and $\beta(0) = 0$.

(H3) $b \geq 0$, $b \in C_b(\mathbb{R})$.

(H4) $D \in L^\infty(\mathbb{R}^d, \mathbb{R}^d)$, $\mathrm{div}_\varrho D \in L^2_{\mathrm{loc}}$, $(\mathrm{div}_\varrho D)^- \in L^\infty$.

Remark 2.1. *(i) By (H1), we have $\varrho \in C^2$, and, since (H1) also implies the existence of $a \in \mathbb{R}$ and $b > 0$ such that $\Phi(x) \geq a + b|x|$, one has $\int_{\mathbb{R}^d} \varrho\, dx < \infty$, i.e. ϱdx is a finite measure.*

(ii) Moreover, since (H1) implies $0 < \varrho \leq 1$ and $\inf_{x\in K} \varrho(x) > 0$ for all compacts $K \subseteq \mathbb{R}^d$, the embedding $W^{m,p} \subseteq W^{m,p}(\varrho)$ is continuous for all $m \in \mathbb{N}_0$, $p \in [1,\infty)$.

(iii) $\nabla\Phi \in L^1(\mathbb{R}^d, \mathbb{R}^d; \varrho)$ is equivalent to $\nabla\varrho \in L^1(\mathbb{R}^d, \mathbb{R}^d)$.

We will construct *generalized solutions* to (1.1) in the following sense.

Definition 2.2. (i) Let X be a real Banach space, $(\tilde{A}, D(\tilde{A}))$ an operator $\tilde{A} : D(\tilde{A}) \subseteq X \to X$, and $u_0 \in X$. A continuous function $u : [0,\infty) \to X$ is called *generalized (*or *mild) solution* to the Cauchy problem

$$\frac{d}{dt}u + \tilde{A}u = 0, \quad u(0) = u_0, \tag{2.1}$$

if there is $h_0 > 0$ such that for all $0 < h < h_0$ there is a step function u_h, defined by the implicit finite difference scheme

$$\begin{aligned} &u_h(0) := u_h^0 := u_0, \\ &u_h(t) := u_h^i, \forall t \in ((i-1)h, ih], i \in \mathbb{N}, \\ &u_h^i \in D(\tilde{A}), u_h^i + h\tilde{A}u_h^i = u_h^{i-1}, \end{aligned} \tag{2.2}$$

such that

$$u(t) = \lim_{h\to 0} u_h(t) \text{ in } X, \text{ locally uniformly in } t.$$

(ii) u is a *generalized (mild) solution to* (1.1) with initial datum $u_0 \in L^1(\varrho)$, if u is a solution in the sense of (i) with $X = L^1(\varrho)$ and $(\tilde{A}, D(\tilde{A}) = (A_0, D(A_0))$, where $(A_0, D(A_0))$ is defined in (2.7) below.

If $(\tilde{A}, D(\tilde{A}))$ is m-accretive, then for each $u_0 \in \overline{D(\tilde{A})}$ (where $\overline{D(\tilde{A})}$ denotes the closure of $D(\tilde{A})$ in X), there exists a unique generalized solution $u = u(u_0)$ to (2.1), see [9, 1], the approximations u_h are given by

$$u_h(t) = (I + h\tilde{A})^{-i}u_0, \quad t \in ((i-1)h, ih],$$

and the exponential formula

$$u(t) = \lim_{n\to\infty} \left(I + \frac{t}{n}\tilde{A}\right)^{-n} u_0$$

holds locally uniformly in t. In order to relate mild solutions u to (1.1) to solutions to the corresponding McKean–Vlasov SDE (1.2), we need to show that u is also a *distributional solution* to (1.1) in the following sense.

Definition 2.3. A measurable curve $\nu : [0,\infty) \to \mathcal{M}_b$ is a *distributional solution* to (1.1) with initial datum $\nu_0 \in \mathcal{M}_b$, if $\nu(t) = v(t)dx$ such that $v, \beta(v\varrho^{-1})\varrho \in L^1_{\mathrm{loc}}([0,\infty) \times \mathbb{R}^d; dxdt)$ and for all $\varphi \in C_c^\infty$ there is a set of full dt-measure $T_\varphi \subseteq (0,\infty)$ such that for all $t \in T_\varphi$

$$\int_{\mathbb{R}^d} \varphi\, v(t)dx = \int_{\mathbb{R}^d} \varphi\, d\nu_0 + \int_0^t \int_{\mathbb{R}^d} \frac{\beta(v(s)\varrho^{-1})}{v(s)\varrho^{-1}}\Delta\varphi + \left(b(v(s)\varrho^{-1})D - \frac{\beta(v(s)\varrho^{-1})}{v(s)\varrho^{-1}}\nabla\Phi\right) \cdot \nabla\varphi\, d\nu(s)ds. \tag{2.3}$$

We call ν a *probability solution*, if $\nu(t)$ is a probability measure for dt-a.e. $t > 0$.

The local integrability assumptions from the previous definition are fulfilled, if $v(t) = u(t)\varrho$, where u is the mild solution from Theorem 2.4. If $v : t \mapsto \nu(t)$ is vaguely continuous, one can choose $T_\varphi = (0, \infty)$ for all $\varphi \in C_c^\infty$. The following theorem is our main existence result for generalized and distributional solutions to (1.1).

Theorem 2.4. *Let Hypothesis 1 be fulfilled. For each $u_0 \in L^1(\varrho) \cap L^\infty$ there exists a generalized solution $u = u(u_0) \in C([0,\infty), L^1(\varrho))$ to (1.1) with initial datum u_0 such that*

$$|u(t)|_\infty \leq \exp\big(|(\mathrm{div}_\varrho D)^- + |D||_\infty^{\frac{1}{2}} t\big)|u_0|_\infty, \quad \forall t \geq 0, \tag{2.4}$$

and $t \mapsto u(t)\varrho dx$ is also a weakly continuous distributional solution to (1.1). Moreover, $t \mapsto S(t)u_0 := u(u_0)(t)$ is a semigroup of $L^1(\varrho)$-contractions on $L^1(\varrho) \cap L^\infty$, i.e. for all $u_0, v_0 \in L^1(\varrho) \cap L^\infty$ and $t, s \geq 0$

$$\begin{aligned} &S(t+s)u_0 = S(t)S(s)u_0, \;\; S(0) = \mathrm{Id}, \\ &|S(t)u_0 - S(t)v_0|_{1,\varrho} \leq |u_0 - v_0|_{1,\varrho}. \end{aligned} \tag{2.5}$$

Furthermore, if $u_0 \in \mathcal{P}_\infty(\varrho)$, then

$$u(t) \in \mathcal{P}_\infty(\varrho), \quad \forall t \geq 0. \tag{2.6}$$

Moreover, if in addition to (H1)-(H4) also (H2') holds, where

(H2') β is Lipschitz continuous,

we get the following stronger result.

Proposition 2.5. *If (H1)-(H4) and (H2') are satisfied, there is a generalized and distributional solution $u = u(u_0)$ to (1.1) for every $u_0 \in L^1(\varrho)$, and the semigroup- and $L^1(\varrho)$-contraction property of $(S(t))_{t\geq 0}$, $S(t)(u_0) = u(u_0)(t)$, from Theorem 2.4 extend to $L^1(\varrho)$. Moreover, if $u_0 \in L^1(\varrho)$ such that $u_0\varrho dx \in \mathcal{P}$, then $S(t)u_0\, \varrho dx \in \mathcal{P}$ for all $t \geq 0$.*

The proofs of Theorem 2.4 and Proposition 2.5 are given in the remainder of the present and the following section. First, we introduce the operators $(L, D(L))$ and $(L_0, D(L_0))$ as follows.

$$L : D(L) \to L^2(\varrho), \quad Lf := \Delta f - \nabla\Phi \cdot \nabla f, \quad D(L) := \{f \in H^2(\varrho) \,|\, \nabla\Phi \cdot \nabla f \in L^2(\varrho)\}.$$

By Lemma 2.9 below, $(L, D(L))$ is symmetric on $L^2(\varrho)$. Since $\Phi \in C^2$, $\Delta - \nabla\Phi \cdot \nabla$ can also be considered a continuous operator L_0,

$$L_0 : D(L_0) := L^1_{\mathrm{loc}} \to \mathcal{D}', L_0 f = \Delta f - \nabla\Phi \cdot \nabla f.$$

Then, $L = L_0$ on $D(L) \subseteq D(L_0)$ in the sense that $Lf = L_0 f$ as elements in $\mathcal{D}'$ for $f \in D(L)$. Consider the nonlinear operator $(A_0, D(A_0))$ in $L^1(\varrho)$

$$\begin{aligned} A_0(f) &:= -\Delta\beta(f) + \nabla\Phi \cdot \nabla\beta(f) + \mathrm{div}_\varrho(Db(f)f) = -L_0\beta(f) + \mathrm{div}_\varrho(Db(f)f), \\ D(A_0) &:= \big\{f \in L^1(\varrho) \cap L^\infty \,|\, -L_0\beta(f) + \mathrm{div}_\varrho(Db(f)f) \in L^1(\varrho) \cap L^\infty\big\}. \end{aligned} \tag{2.7}$$

Note that $f \in D(A_0)$ implies $\beta(f) \in L^1(\varrho) \cap L^\infty$ (since β is locally Lipschitz continuous), and hence $L_0\beta(f) \in \mathcal{D}'$. The following lemma is the key part to prove Theorem 2.4.

Lemma 2.6. *Let $\lambda_0 := [(|(\mathrm{div}_\varrho)^- + |D||_\infty + |(\mathrm{div}_\varrho)^- + |D||_\infty^{\frac{1}{2}})|b|_\infty]^{-1}$. We have*

(i) $R(I + \lambda A_0) = L^1(\varrho) \cap L^\infty$, $\forall 0 < \lambda < \lambda_0$.

(ii) For each $0 < \lambda < \lambda_0$*, there is* $J_\lambda : L^1(\varrho) \cap L^\infty \to D(A_0) \subseteq L^1(\varrho) \cap L^\infty$*,* $J_\lambda f \in (I + \lambda A_0)^{-1} f$ *such that*

$$|J_\lambda f - J_\lambda g|_{1,\varrho} \le |f - g|_{1,\varrho}, \quad \forall f, g \in L^1(\varrho) \cap L^\infty.$$

Moreover, for all $f \in L^1(\varrho) \cap L^\infty$ *and* $0 < \lambda_1, \lambda_2, \lambda < \lambda_0$

$$J_{\lambda_2} f = J_{\lambda_1}\left(\frac{\lambda_1}{\lambda_2} f + \big(1 - \frac{\lambda_1}{\lambda_2}\big) J_{\lambda_2} f\right), \tag{2.8}$$

$$|J_\lambda f|_\infty \le (1 + |(\mathrm{div}_\varrho D)^- + |D||_\infty^{\frac{1}{2}})|f|_\infty, \tag{2.9}$$

and

$$J_\lambda(\mathcal{P}_\infty(\varrho)) \subseteq \mathcal{P}_\infty(\varrho). \tag{2.10}$$

Finally,

$$|J_\lambda g - g|_{1,\varrho} \le C\lambda(|g|_{H^2(\varrho)} + |g|_{W^{2,1}(\varrho)}), \quad \forall g \in C_c^\infty, \tag{2.11}$$

for a constant $C > 0$ *independent of* λ*.*

The proof of Lemma 2.6 is given in the next section. In the rest of the present section, we prove Theorem 2.4 as follows. Define the operator $A : D(A) \subseteq L^1(\varrho) \cap L^\infty \to L^1(\varrho) \cap L^\infty$ via

$$\begin{aligned} A(f) &:= A_0(f), \\ D(A) &:= \big\{J_\lambda f, f \in L^1(\varrho) \cap L^\infty\big\}, \end{aligned} \tag{2.12}$$

for any $0 < \lambda < \lambda_0$. Indeed, by (2.8), $D(A)$ does not depend on λ. Any generalized solution of

$$\frac{d}{dt} u + Au = 0 \tag{2.13}$$

is also a generalized solution to the same equation with A_0 replacing A, and hence a generalized solution to (1.1). By definition, A is a restriction of A_0 such that $R(I + \lambda A) = R(I + \lambda A_0) = L^1(\varrho) \cap L^\infty$ for all $0 < \lambda < \lambda_0$, but in contrast to A_0, A is defined such that $I + \lambda A$ is bijective from $D(A)$ to $L^1(\varrho) \cap L^\infty$. Indeed, we have the following lemma.

Lemma 2.7. *For each* $0 < \lambda < \lambda_0$*,* $I + \lambda A$ *is bijective from* $D(A)$ *to* $L^1(\varrho) \cap L^\infty$ *with inverse* J_λ*, and*

$$|(I + \lambda A)^{-1} f - (I + \lambda A)^{-1} g|_{1,\varrho} \le |f - g|_{1,\varrho}, \quad \forall f, g \in L^1(\varrho) \cap L^\infty.$$

Moreover, (2.8)-(2.10) *hold with* $(I + \lambda A)^{-1}$ *instead of* J_λ*, and* $L^1(\varrho) \cap L^\infty \subseteq \overline{D(A)}$*, where* $\overline{C}$ *denotes the* $L^1(\varrho)$*-closure of a set* $C \subseteq L^1(\varrho)$*.*

Proof of Lemma 2.7. All assertions but the final one are immediate consequences of Lemma 2.6. For the final assertion, letting $\lambda \longrightarrow 0$ in (2.11), the definition of $D(A)$ yields $C_c^\infty \subseteq \overline{D(A)}$, and hence the claim follows from the density of C_c^∞ in $L^1(\varrho) \cap L^\infty$ with respect to $|\cdot|_{1,\varrho}$. □

Then the proof of Theorem 2.4 is obtained as follows.

Proof of Theorem (2.4). The existence of a semigroup $(S(t))_{t \ge 0}$ of $L^1(\varrho)$-contractions on $L^1(\varrho) \cap L^\infty$ such that $t \mapsto S(t)u_0$ is a generalized solution to (2.13) (and hence to (1.1)) with initial datum u_0 for each $u_0 \in L^1(\varrho) \cap L^\infty$ and

$$S(t)u_0 = \lim_{n\to\infty} \left(I + \frac{t}{n} A\right)^{-n} u_0 = \lim_{n\to\infty} (J_{\frac{t}{n}})^n u_0 \text{ in } L^1(\varrho) \text{ locally uniformly in } t \ge 0 \tag{2.14}$$

follows from Lemma 2.7 and the Crandall-Liggett theory for nonlinear evolution equations in Banach spaces, see [1, Thm.4.3]. Indeed, even though Lemma 2.7 does not imply m-accretivity of $(A, D(A))$

in $L^1(\varrho)$, the proof of [1, Thm.4.3.] shows that it still implies the claims above. By (2.9), standard arguments entail (2.4). Moreover, if $u_0 \in \mathcal{P}_\infty(\varrho)$, (2.6) follows from (2.10) and (2.14). Since $t \mapsto u(t)\varrho dx$ is weakly continuous due to the continuity of $t \mapsto u(t)$ in $L^1(\varrho)$, it remains to prove that $t \mapsto u(t)\varrho dx$ is also a distributional solution. This we prove after the proof of Lemma 2.6 at the end of Section 3.

Remark 2.8. *Let us comment on the question of uniqueness of u in the class of generalized and distributional solutions to* (1.1)*.*

(i) As said before, Lemma 2.7 does not yield m-accretivity of $(A, D(A))$ in either $L^1(\varrho)$ or $L^1(\varrho)\cap L^\infty$. Hence the Crandall-Liggett theorem does not yield uniqueness of u in the class of generalized solutions to (1.1)*. u is, however, the unique generalized solution such that the approximations u_h in Definition 2.2 can be chosen so that $\sup_{0<h<h_0} |u_h(t)|_\infty$ is locally bounded in t for some $h_0 > 0$. Indeed, this follows by Lemma 2.7 and an inspection of the uniqueness part of the proof of [1, Thm.4.1]. However, even with m-accretivity of $(A, D(A))$, uniqueness of u would only hold in the class of generalized solutions* with respect to *$(A, D(A))$, which is a restriction of $(A_0, D(A_0))$, but not in the class of* all *generalized solutions.*

(ii) For $\varrho \equiv 1$ and under additional assumptions on the coefficients, it is proven in [5] that A_0 itself is m-accretive. In this case, the generalized solutions are unique. We believe that such a result can be extended to more general weights ϱ, but we do not pursue this question in the present paper.

For the proof of Lemma 2.6, we need the following result.

Lemma 2.9. *(i) Let $v \in D(L)$ and $w \in H^1(\varrho)$. Then $\int_{\mathbb{R}^d} Lv\, w\, \varrho dx = -\int_{\mathbb{R}^d} \nabla v \cdot \nabla w\, \varrho dx$.*

(ii) If either

(a) $v, w \in D(L)$ or

(b) $v \in L^1_{\text{loc}}$ such that $L_0 v \in L^1_{\text{loc}}$ and $w \in C^2_c$, then $\int_{\mathbb{R}^d} w\, L_0 v\, \varrho dx = \int_{\mathbb{R}^d} vLw\, \varrho dx$.

In particular, $(L, D(L))$ is symmetric on $L^2(\varrho)$.

Proof. (i) For $w \in C^1_c$, the claim follows from integration by parts to $(\Delta v)w\varrho$, noting that $w\varrho \in C^1_c$ and $-\nabla v \cdot \nabla(w\varrho) = -\nabla v \cdot (\nabla w - w\nabla\Phi)\varrho$. Since C^1_c is dense in $H^1(\varrho)$, the claim follows.

(ii) For (a), repeat the proof of (i), then reverse the roles of v and w to get

$$\int_{\mathbb{R}^d} w\, L_0 v\, \varrho dx = \int_{\mathbb{R}^d} wLv\, \varrho dx = -\int_{\mathbb{R}^d} \nabla v \cdot \nabla w\, \varrho dx = \int_{\mathbb{R}^d} vLw\, \varrho dx.$$

For (b), as a distribution $L_0 v$ acts on $w\varrho \in C^2_c$ via $\int w\, L_0 v\, \varrho dx$. Also, Δv and $-\nabla\Phi \cdot \nabla v$ are distributions, and the action of their sum applied to $\omega\varrho$ is

$$\int_{\mathbb{R}^d} v([\Delta(w\varrho) + \operatorname{div}(w\varrho\nabla\Phi)]dx = \int_{\mathbb{R}^d} v[\Delta w - \nabla\Phi \cdot \nabla w]\varrho dx = \int_{\mathbb{R}^d} vLw\, \varrho dx.$$

□

3 Proof of Lemma 2.6

Let $\lambda > 0$. The definition of $(A_0, D(A_0))$ entails $R(I + \lambda A_0) \subseteq L^1(\varrho) \cap L^\infty$. Let us prove the reverse inclusion. In order to solve

$$u + \lambda A_0 u = f,$$

we approximate β, D and b as follows. For $\varepsilon > 0$, consider

$$\tilde{\beta}_\varepsilon(r) := \beta_\varepsilon(r) + \varepsilon r := \frac{1}{\varepsilon}\big(r - (I+\varepsilon\beta)^{-1}r\big) + \varepsilon r = \beta\big((I+\varepsilon\beta)^{-1}r\big) + \varepsilon r,\ r \in \mathbb{R}.$$

It is readily seen that $\tilde{\beta}_\varepsilon$ is strictly increasing and, for all $r, t \in \mathbb{R}$,

$$|\tilde{\beta}_\varepsilon(r) - \tilde{\beta}_\varepsilon(t)| \le \big(\varepsilon + 2\varepsilon^{-1}\big)|r-t|, \tag{3.1}$$

$$|\tilde{\beta}_\varepsilon(r) - \tilde{\beta}_\varepsilon(t)| \ge \varepsilon|r-t|, \tag{3.2}$$

$$\tilde{\beta}_\varepsilon(0) = 0, \tag{3.3}$$

i.e. $\tilde{\beta}_\varepsilon : \mathbb{R} \to \mathbb{R}$ is a bijective $(\varepsilon + \frac{2}{\varepsilon})$-Lipschitz map with $\frac{1}{\varepsilon}$-Lipschitz inverse. Set

$$b_\varepsilon(r) := \frac{(b * \rho_\varepsilon)(r)}{1+\varepsilon|r|}, \quad r \in \mathbb{R},$$

and $D_\varepsilon := \eta_\varepsilon D$, where $\rho_\varepsilon(r) := \varepsilon^{-1}\rho(\frac{r}{\varepsilon}), \rho \in C_c^\infty, \rho \ge 0$ is a standard mollifier, and $(\eta_\varepsilon)_{\varepsilon>0} \subseteq C_c^1$ is a family of functions with $0 \le \eta_\varepsilon \le 1$, $|\nabla\eta_\varepsilon| \le 1$ and $\eta_\varepsilon \equiv 1$ on $B_{\varepsilon^{-1}}(0)$ (the Euclidean ball with radius ε^{-1} centered at 0). Consequently, $b^*_\varepsilon(r) := b_\varepsilon(r)r$ is bounded and Lipschitz, and $b_\varepsilon(r) \longrightarrow b(r)$ pointwise as $\varepsilon \to 0$. Moreover, by (H4),

$$\mathrm{div}_\varrho D_\varepsilon = \eta_\varepsilon \mathrm{div}_\varrho D + \nabla\eta_\varepsilon \cdot D \in L^2(\varrho) + L^\infty.$$

Consider the approximate equation

$$u - \lambda L\tilde{\beta}_\varepsilon(u) + \lambda\varepsilon\tilde{\beta}_\varepsilon(u) + \lambda \mathrm{div}_\varrho(D_\varepsilon b^*_\varepsilon(u)) = f. \tag{3.4}$$

To solve it, first let $f \in L^2(\varrho)$, consider in $L^2(\varrho)$

$$(\varepsilon I - L)^{-1}u + \lambda\tilde{\beta}_\varepsilon(u) + \lambda(\varepsilon I - L)^{-1}\mathrm{div}_\varrho(D_\varepsilon b^*_\varepsilon(u)) = (\varepsilon I - L)^{-1}f, \tag{3.5}$$

and define on $L^2(\varrho)$ the maps F_ε, G^1_ε, G^2_ε

$$F_\varepsilon : u \mapsto (\varepsilon I - L)^{-1}u, \quad G^1_\varepsilon : u \mapsto \lambda\tilde{\beta}_\varepsilon(u), \quad G^2_\varepsilon : u \mapsto \lambda(\varepsilon I - L)^{-1}\mathrm{div}_\varrho(D_\varepsilon b^*_\varepsilon(u)),$$

so that (3.5) becomes

$$(F_\varepsilon + G^1_\varepsilon + G^2_\varepsilon)u = (\varepsilon I - L)^{-1}f.$$

By (3.1), G^1_ε is $L^2(\varrho)$-continuous, and (3.2) and the monotonicity of β_ε entail

$$\langle G^1_\varepsilon(u-v), u-v\rangle_{2,\varrho} \ge \lambda\varepsilon|u-v|^2_{2,\varrho}, \quad \forall u, v \in L^2(\varrho).$$

In [10] it is shown that the resolvent set of $(L, D(L))$ contains $(0,\infty)$ and, moreover,

$$|(\varepsilon I - L)^{-1}u|_{2,\varrho} \le \varepsilon^{-1}|u|_{2,\varrho}, \quad \forall u \in L^2(\varrho).$$

In particular, F_ε is continuous on $L^2(\varrho)$. Furthermore, Lemma 2.9 (i) yields

$$\langle F_\varepsilon(u), u\rangle_{2,\varrho} = \varepsilon|(\varepsilon I - L)^{-1}u|^2_{2,\varrho} + |\nabla(\varepsilon I - L)^{-1}u|^2_{2,\varrho}, \quad \forall u \in L^2(\varrho).$$

Since $r \mapsto b^*_\varepsilon(r)$ is Lipschitz and $D_\varepsilon \in L^\infty$, $u \mapsto D_\varepsilon b^*_\varepsilon(u)$ is $L^2(\varrho)$-continuous and hence G^2_ε is $L^2(\varrho)$-continuous as well. Moreover, denoting by $C_\varepsilon > 0$ a constant depending on $|D_\varepsilon|_\infty$ and the Lipschitz constant of b^*_ε, we have

$$\begin{aligned}\langle G^2_\varepsilon(u) - G^2_\varepsilon(v), u-v\rangle_{2,\varrho} &= -\lambda\langle D_\varepsilon(b^*_\varepsilon(u) - b^*_\varepsilon(v)), \nabla(\varepsilon I - L)^{-1}(u-v)\rangle_{2,\varrho} \\ &\ge -C_\varepsilon\lambda|u-v|_{2,\varrho}|\nabla(\varepsilon I - L)^{-1}(u-v)|_{2,\varrho}, \quad \forall u, v \in L^2(\varrho).\end{aligned}$$

Altogether, these estimates lead to

$$\begin{aligned}&\langle F_\varepsilon(u-v)+G^1_\varepsilon(u)-G^1_\varepsilon(v)+G^2_\varepsilon(u)-G^2_\varepsilon(v),u-v\rangle_{2,\varrho}\\&\geq \lambda\varepsilon|u-v|^2_{2,\varrho}+\varepsilon|(\varepsilon I-L)^{-1}(u-v)|^2_{2,\varrho}+|\nabla(\varepsilon I-L)^{-1}(u-v)|^2_{2,\varrho}\\&\quad -C_\varepsilon\lambda|u-v|^2_{2,\varrho}|\nabla(\varepsilon I-L)^{-1}(u-v)|^2_{2,\varrho}\geq \frac{\lambda\varepsilon}{2}|u-v|^2_{2,\varrho},\end{aligned}$$

provided $0<\lambda<\lambda_\varepsilon$ for $\lambda_\varepsilon>0$ sufficiently small, i.e. $F_\varepsilon+G^1_\varepsilon+G^2_\varepsilon$ is coercive on $L^2(\varrho)$. Since the continuity and monotonicity (= accretivity) of $F_\varepsilon+G^1_\varepsilon+G^2_\varepsilon$ on $L^2(\varrho)$ also imply its m-accretivity, it follows by [1, Cor.2.2.] that $F_\varepsilon+G^1_\varepsilon+G^2_\varepsilon$ is surjective (hence bijective) on $L^2(\varrho)$. Therefore, for each $f\in L^2(\varrho)$, (3.5) has a unique solution $u_\varepsilon(f)=u_\varepsilon(\lambda,f)$ in $L^2(\varrho)$. Since for this solution the right-hand side and the first summand on the left-hand side of (3.5) are in $D(L)\subseteq H^2(\varrho)$, and $G^2_\varepsilon(u_\varepsilon)\in H^1(\varrho)$, it follows that $\tilde\beta_\varepsilon(u_\varepsilon)\in H^1(\varrho)$. Since $\tilde\beta_\varepsilon$ has a Lipschitz inverse, we obtain $u_\varepsilon\in H^1(\varrho)$. Thus also $b^*_\varepsilon(u_\varepsilon)\in H^1(\varrho)$. Since

$$\mathrm{div}_\varrho(D_\varepsilon b^*_\varepsilon(u_\varepsilon))=b^*_\varepsilon(u_\varepsilon)\mathrm{div}_\varrho D_\varepsilon+D_\varepsilon\cdot\nabla b^*_\varepsilon(u_\varepsilon)$$

and $b^*_\varepsilon(u_\varepsilon)\in L^2(\varrho)\cap L^\infty$, $\mathrm{div}_\varrho D_\varepsilon\in L^2(\varrho)+L^\infty$, $D_\varepsilon\in L^\infty$, we obtain $\mathrm{div}_\varrho(D_\varepsilon b^*_\varepsilon(u_\varepsilon))\in L^2(\varrho)$, and therefore even $G^2_\varepsilon(u_\varepsilon)\in D(L)$. Hence, all terms of (3.5) are in $D(L)$, and applying $(\varepsilon I-L)$ to both sides shows that u_ε solves (3.4) in $L^2(\varrho)$.

Next, we prove $u_\varepsilon\in L^1(\varrho)$ if $f\in L^1(\varrho)\cap L^2(\varrho)$ and that for $0<\lambda<\lambda_\varepsilon$

$$|u_\varepsilon(\lambda,f_1)-u_\varepsilon(\lambda,f_2)|_{1,\varrho}\leq|f-g|_{1,\varrho},\quad\forall f_1,f_2\in L^1(\varrho).$$

To this end, set $u:=u_1-u_2:=u_\varepsilon(\lambda,f_1)-u_\varepsilon(\lambda,f_2)$ and $f:=f_1-f_2$. Then, by (3.4),

$$u=\lambda L[\tilde\beta_\varepsilon(u_1)-\tilde\beta_\varepsilon(u_2)]-\lambda\varepsilon[\tilde\beta_\varepsilon(u_1)-\tilde\beta_\varepsilon(u_2)]-\lambda\mathrm{div}_\varrho(D_\varepsilon[b^*_\varepsilon(u_1)-b^*_\varepsilon(u_2)])+f.\tag{3.6}$$

Consider, for $\delta>0$, the Lipschitzian function $\chi_\delta:\mathbb{R}\to\mathbb{R}$,

$$\chi_\delta(r):=\begin{cases}1, & r\geq\delta,\\ \frac{r}{\delta}, & |r|<\delta,\\ -1, & r<-\delta,\end{cases}$$

i.e. $\chi_\delta\to$ sign pointwise for $\delta\to0$. Multiplying both sides of (3.6) by $\Lambda_\delta:=\chi_\delta(\tilde\beta_\varepsilon(u_1)-\tilde\beta_\varepsilon(u_2))\in H^1(\varrho)$, integrating with respect to ϱdx and using $\Lambda_\delta(\tilde\beta_\varepsilon(u_1)-\tilde\beta_\varepsilon(u_2))\geq0$ implies

$$\int_{\mathbb{R}^d}u\Lambda_\delta\varrho dx\leq\lambda\int_{\mathbb{R}^d}L[\tilde\beta_\varepsilon(u_1)-\tilde\beta_\varepsilon(u_2)]\Lambda_\delta\,\varrho dx+\int_{\mathbb{R}^d}\Lambda_\delta\mathrm{div}_\varrho(D_\varepsilon[b^*_\varepsilon(u_1)-b^*_\varepsilon(u_2)])\,\varrho dx+\int_{\mathbb{R}^d}f\Lambda_\delta\,\varrho dx.$$

By Lemma 2.9 (i) and since $\chi'_\delta(r)=\mathbb{1}_{\{|r|<\delta\}}\frac{1}{\delta}$, the first integral on the right-hand side of the previous inequality equals $-\frac{1}{\delta}\int_{\{|\tilde\beta_\varepsilon(u_1)-\tilde\beta_\varepsilon(u_2)|<\delta\}}|\nabla[\tilde\beta_\varepsilon(u_1)-\tilde\beta_\varepsilon(u_2)]|^2\varrho dx\leq0$. Moreover, for a constant $c_0>0$,

$$-\int_{\mathbb{R}^d}\Lambda_\delta\mathrm{div}_\varrho(D_\varepsilon[b^*_\varepsilon(u_1)-b^*_\varepsilon(u_2)])\,\varrho dx\leq\frac{c_0}{\delta}\int_{\{|\tilde\beta_\varepsilon(u_1)-\tilde\beta_\varepsilon(u_2)|<\delta\}}|D_\varepsilon||u||\nabla(\tilde\beta_\varepsilon(u_1)-\tilde\beta_\varepsilon(u_2))|\,\varrho dx.$$

Since $\tilde\beta_\varepsilon$ has a Lipschitz inverse, we further bound the right-hand side of the previous estimate by

$$c_0|D_\varepsilon|_{2,\varrho}\bigg(\int_{\{|\tilde\beta_\varepsilon(u_1)-\tilde\beta_\varepsilon(u_2)|<\delta\}}|\nabla(\tilde\beta_\varepsilon(u_1)-\tilde\beta_\varepsilon(u_2))|^2\,\varrho dx\bigg)^{\frac12}\xrightarrow{\delta\to0}0,$$

since $\nabla(\tilde\beta_\varepsilon(u_1)-\tilde\beta_\varepsilon(u_2))=0$ ϱdx-a.s. on $\{|\tilde\beta_\varepsilon(u_1)-\tilde\beta_\varepsilon(u_2)|=0\}$. Thus, altogether we obtain by Fatou's lemma and since $|\Lambda_\delta|\leq1$:

$$\int_{\mathbb{R}^d}|u|\,\varrho dx\leq|f|_{1,\varrho},\tag{3.7}$$

Since $u_\varepsilon(\lambda, 0) \equiv 0$ for all $0 < \lambda < \lambda_\varepsilon$, in particular we have shown $|u_\varepsilon|_{1,\varrho} \leq |f|_{1,\varrho}$ for any $f \in L^1(\varrho) \cap L^2(\varrho)$. By (3.1), then also $\tilde{\beta}_\varepsilon(u_\varepsilon) \in L^1(\varrho)$. Consequently, (3.4) implies $-L\tilde{\beta}_\varepsilon(u_\varepsilon) + \mathrm{div}_\varrho(D_\varepsilon b^*_\varepsilon(u_\varepsilon)) \in L^1(\varrho)$, and in this sense (3.4) holds in $L^1(\varrho)$.

Now let $f \in L^1(\varrho)$ and consider the operator $(A_\varepsilon, D(A_\varepsilon))$

$$A_\varepsilon u := -\Delta\tilde{\beta}_\varepsilon(u) + \nabla\Phi \cdot \nabla\tilde{\beta}_\varepsilon(u) + \varepsilon\tilde{\beta}_\varepsilon(u) + \mathrm{div}_\varrho(D_\varepsilon b^*_\varepsilon(u)),$$

$$D(A_\varepsilon) := \{u \in L^1(\varrho) : \ A_\varepsilon u \in L^1(\varrho)\},$$

so that we may rewrite (3.4) as

$$u + \lambda A_\varepsilon u = f. \tag{3.8}$$

Let $(f_n)_{n\in\mathbb{N}} \subseteq L^1(\varrho) \cap L^2(\varrho)$ such that $f_n \longrightarrow f$ in $L^1(\varrho)$ as $n \to \infty$. By (3.7), the corresponding solutions $u_n := u_\varepsilon(\lambda, f_n)$ to (3.4) converge to some $u \in L^1(\varrho)$ as $n \to \infty$, and hence $A_\varepsilon(u_n)$ converges in $L^1(\varrho)$ to some g. Since $\tilde{\beta}_\varepsilon(u_n) \in D(L)$, we have $\tilde{\beta}_\varepsilon(u_\varepsilon), \Delta\tilde{\beta}_\varepsilon(u_\varepsilon), \nabla\Phi \cdot \nabla\tilde{\beta}_\varepsilon(u_\varepsilon) \in L^1_{\mathrm{loc}}$, and the Lipschitz continuity of b^*_ε and $\tilde{\beta}_\varepsilon$ yields for all $\varphi \in C^2_c$

$$\begin{aligned}\int_{\mathbb{R}^d} (A_\varepsilon u_n)\,\varphi\,\varrho dx &= \int_{\mathbb{R}^d} \tilde{\beta}_\varepsilon(u_n)\big(-L\varphi + \varepsilon\varphi\big) - D_\varepsilon b^*_\varepsilon(u_n)\cdot\nabla\varphi\,\varrho dx\\ &\xrightarrow{n\to\infty} \int_{\mathbb{R}^d} \tilde{\beta}_\varepsilon(u)\big(-L\varphi + \varepsilon\varphi\big) - D_\varepsilon b^*_\varepsilon(u)\cdot\nabla\varphi\,\varrho dx\\ &\qquad = {}_{\mathcal{D}'}\langle -\Delta\tilde{\beta}_\varepsilon(u) + \nabla\Phi\cdot\nabla\tilde{\beta}_\varepsilon(u) + \varepsilon\tilde{\beta}_\varepsilon(u) + \mathrm{div}_\varrho(D_\varepsilon b^*_\varepsilon(u)), \varphi\varrho\rangle_{\mathcal{D}},\end{aligned}$$

where ${}_{\mathcal{D}'}\langle F, f\rangle_{\mathcal{D}}$ denotes the dual pairing of a second-order distribution F and $f \in C^2_c$. Since $\varrho \in C^2(\mathbb{R}^d)$ and $\varrho > 0$, one has $\{\varphi\varrho; \varphi \in C^2_c\} = C^2_c$. Consequently, $u \in D(A_\varepsilon)$ and $g = A_\varepsilon u$, i.e. u solves (3.8). Denoting this (in general non-unique) solution to (3.8) by $u_\varepsilon(\lambda, f)$, (3.7) implies for all $0 < \lambda < \lambda_\varepsilon$

$$|u_\varepsilon(\lambda, f) - u_\varepsilon(\lambda, g)|_{1,\varrho} \leq |f - g|_{1,\varrho}, \quad \forall f, g \in L^1(\varrho). \tag{3.9}$$

By [1, Prop.3.3], this implies that for each $\varepsilon > 0$ equation (3.8) has a solution $u_\varepsilon(\lambda, f) \in D(A_\varepsilon)$ for all $f \in L^1(\varrho)$ for *all* $\lambda > 0$, (3.9) holds for all $\lambda > 0$, and

$$u_\varepsilon(\lambda_2, f) = u_\varepsilon\left(\lambda_1, \frac{\lambda_1}{\lambda_2}f + (1 - \frac{\lambda_1}{\lambda_2})u_\varepsilon(\lambda_2, f)\right), \quad \forall 0 < \lambda_1, \lambda_2 < \infty, f \in L^1(\varrho). \tag{3.10}$$

Moreover, similarly to the proof of equality (2.29) in [7], one proves $u_\varepsilon(\lambda, f) \in L^1(\varrho) \cap L^2(\varrho)$ for all $f \in L^1(\varrho) \cap L^2(\varrho)$ and all $\lambda > 0$. Since we know by the above construction that $u_\varepsilon(\lambda, f) \in H^1(\varrho)$ if $\lambda < \lambda_\varepsilon$ and $f \in L^1(\varrho) \cap L^2(\varrho)$, (3.10) implies $u_\varepsilon(\lambda, f) \in H^1(\varrho)$ for such f for *all* $\lambda > 0$.

From now on, let $f \in L^1(\varrho) \cap L^\infty$ and let $u_\varepsilon = u_\varepsilon(\lambda, f) \in H^1(\varrho) \cap L^1(\varrho)$ be the corresponding solution to (3.8). For $0 < \lambda < \lambda_0$, where

$$\lambda_0 = \left[\big(|(\mathrm{div}_\varrho D)^- + |D||_\infty + |(\mathrm{div}_\varrho D)^- + |D||_\infty^{\frac{1}{2}}\big)|b|_\infty\right]^{-1},$$

we are now going to prove

$$\sup_{\varepsilon>0} |u_\varepsilon|_\infty \leq |f|_\infty(1 + |(\mathrm{div}_\varrho D)^- + |D||_\infty^{\frac{1}{2}}) \tag{3.11}$$

as follows. Setting $M_\varepsilon := |(\mathrm{div}_\varrho D_\varepsilon)^-|_\infty^{\frac{1}{2}}|f|_\infty$, one has, since $(\mathrm{div}_\varrho D_\varepsilon)^- \leq |(\mathrm{div}_\varrho D)^- + |D||_\infty$,

$$\begin{aligned}&u_\varepsilon - |f|_\infty - M_\varepsilon - \lambda L\big(\tilde{\beta}_\varepsilon(u_\varepsilon) - \tilde{\beta}_\varepsilon(|f|_\infty + M_\varepsilon)\big) + \lambda\varepsilon\big(\tilde{\beta}_\varepsilon(u_\varepsilon) - \tilde{\beta}_\varepsilon(|f|_\infty + M_\varepsilon)\big)\\ &\quad + \lambda\mathrm{div}_\varrho\big(D_\varepsilon\big(b^*_\varepsilon(u_\varepsilon) - b^*_\varepsilon(|f|_\infty + M_\varepsilon)\big)\big) \leq f - |f|_\infty - M_\varepsilon + \lambda(\mathrm{div}_\varrho D_\varepsilon)^- b^*_\varepsilon(|f|_\infty + M_\varepsilon) \leq 0.\end{aligned}$$

Multiplying the above inequality by $\chi_\delta([u_\varepsilon - |f|_\infty + M_\varepsilon]^+) \in H^1(\varrho)$, integrating with respect to ϱdx and using

$$\chi_\delta([u_\varepsilon - |f|_\infty - M_\varepsilon]^+)\lambda\varepsilon\big(\tilde{\beta}_\varepsilon(u_\varepsilon) - \tilde{\beta}_\varepsilon(|f|_\infty + M_\varepsilon)\big) \geq 0$$

as well as

$$-\lambda \int_{\mathbb{R}^d} \chi_\delta([u_\varepsilon - |f|_\infty - M_\varepsilon]^+) L\big(\tilde{\beta}_\varepsilon(u_\varepsilon) - \tilde{\beta}_\varepsilon(|f|_\infty + M_\varepsilon)\big)\, \varrho dx \geq 0,$$

we find

$$\int_{\mathbb{R}^d} \chi_\delta([u_\varepsilon - |f|_\infty - M_\varepsilon]^+)(u_\varepsilon - |f|_\infty - M_\varepsilon)\, \varrho dx \leq \frac{\lambda}{\delta} \int_{C_\delta} D_\varepsilon \big(b_\varepsilon^*(u_\varepsilon) - b_\varepsilon^*(|f|_\infty + M_\varepsilon)\big) \cdot \nabla u_\varepsilon\, \varrho dx,$$

where $C_\delta := \{|f|_\infty + M_\varepsilon \leq u_\varepsilon \leq |f|_\infty + M_\varepsilon + \delta\}$. The right-hand side of the previous inequality is bounded from above by

$$\frac{c_\varepsilon \lambda}{\delta} |D_\varepsilon|_{2,\varrho} \bigg(\int_{C_\delta} |u_\varepsilon - |f|_\infty - M_\varepsilon|^2 |\nabla u_\varepsilon|^2\, \varrho dx \bigg)^{\frac{1}{2}} \leq c_\varepsilon \lambda |D_\varepsilon|_{2,\varrho} \bigg(\int_{C_\delta} |\nabla u_\varepsilon|^2\, \varrho dx \bigg)^{\frac{1}{2}} \xrightarrow{\delta \to 0} 0,$$

where the final convergence follows from $u_\varepsilon \in H^1(\varrho)$ and $\mathbb{1}_{C_\delta} \nabla u_\varepsilon \xrightarrow{\delta \to 0} 0$ ϱdx-a.s. Consequently,

$$\int_{\mathbb{R}^d} [u_\varepsilon - |f|_\infty - M_\varepsilon]^+\, \varrho dx \leq \liminf_{\delta \to 0} \int_{\mathbb{R}^d} \chi_\delta([u_\varepsilon - |f|_\infty - M_\varepsilon]^+)(u_\varepsilon - |f|_\infty - M_\varepsilon)\, \varrho dx = 0,$$

which implies

$$u_\varepsilon \leq |f|_\infty + M_\varepsilon = |f|_\infty (1 + |(\mathrm{div}_\varrho D_\varepsilon)^-|_\infty^{\frac{1}{2}}) \leq |f|_\infty (1 + |(\mathrm{div}_\varrho D)^- + |D||_\infty^{\frac{1}{2}}).$$

Similarly, one obtains

$$u_\varepsilon \geq -|f|_\infty (1 + |(\mathrm{div}_\varrho D)^- + |D||_\infty^{\frac{1}{2}}),$$

and so (3.11) follows.

For the remainder of the proof, we fix $0 < \lambda < \lambda_0$, $f \in L^1(\varrho) \cap L^\infty$. By (3.9) and (3.11), $(u_\varepsilon)_{\varepsilon > 0}, u_\varepsilon = u_\varepsilon(\lambda, f)$, is uniformly bounded in each $L^p(\varrho)$, $p \in [1, \infty]$. In particular, there exists an $L^2(\varrho)$-weakly convergent subsequence (for simplicity again denoted $(u_\varepsilon)_{\varepsilon > 0}$) and some $u = u(\lambda, f) \in L^2(\varrho)$ such that

$$u_\varepsilon \xrightarrow{\varepsilon \to 0} u \quad L^2(\varrho) - \text{weakly}. \tag{3.12}$$

Moreover, the uniform L^∞-bound for $\{u_\varepsilon\}_{\varepsilon > 0}$ also gives the ϱdx-a.s. estimate

$$|\tilde{\beta}_\varepsilon(u_\varepsilon)| \leq (\varepsilon + C_{\beta,f}) |u_\varepsilon|,$$

where $C_{\beta,f} \in [0, \infty)$ is the Lipschitz constant of β on $[-|f|_\infty, |f|_\infty]$. Consequently, $\{\tilde{\beta}_\varepsilon(u_\varepsilon)\}_{\varepsilon > 0}$ is uniformly $L^p(\varrho)$-bounded for any $p \in [1, \infty]$ as well, so there is η such that for all $q \in (1, \infty)$ $\tilde{\beta}_\varepsilon(u_\varepsilon) \xrightarrow{\varepsilon \to 0} \eta$ $L^q(\varrho)$-weakly. In fact, also $u_\varepsilon \longrightarrow u$ pointwise ϱdx-a.s., as $\varepsilon \to 0$, as the following argument shows. Considering (3.8) with $u_\varepsilon \in H^1(\varrho) \cap L^1(\varrho) \cap L^\infty$ instead of u, multiplying with u_ε, integrating with respect to ϱdx, using Lemma 2.9 (i), the monotonicity of $\tilde{\beta}_\varepsilon$, Young's inequality, and $b_\varepsilon^*(u_\varepsilon) \in H^1(\varrho)$ gives

$$|u_\varepsilon|_{L^2(\varrho)}^2 + 2\lambda \int_{\mathbb{R}^d} (\beta_\varepsilon)'(u_\varepsilon) |\nabla u_\varepsilon|^2\, \varrho dx \leq |f|_{L^2(\varrho)}^2 + 2\lambda \int_{\mathbb{R}^d} D_\varepsilon b_\varepsilon^*(u_\varepsilon) \cdot \nabla u_\varepsilon\, \varrho dx. \tag{3.13}$$

Defining the non-negative function ψ,

$$\psi(r) := \int_0^r b_\varepsilon^*(s)\, ds, \quad r \in \mathbb{R},$$

we rewrite and estimate the final integral on the right-hand side of the previous inequality as

$$-2\lambda \int_{\mathbb{R}^d} (\mathrm{div}_\varrho D_\varepsilon) \psi(u_\varepsilon)\, \varrho dx \leq \lambda |b|_\infty |u_\varepsilon|_\infty |u_\varepsilon|_{1,\varrho} |(\mathrm{div}_\varrho D)^- + |D||_\infty. \tag{3.14}$$

Note that the right-hand side of this estimate is bounded by a finite constant C, which is independent of $\varepsilon > 0$. Setting $g_\varepsilon(r) := (I + \varepsilon\beta)^{-1}(r), r \in \mathbb{R}$, one has

$$\beta'_\varepsilon(r) \geq \frac{\beta'}{1+\beta'}\big(g_\varepsilon(r)\big)g'_\varepsilon(r)^2,$$

and introducing $a(r) := \int_0^r \frac{\beta'}{1+\beta'}(s)ds, r \in \mathbb{R}$, one has

$$|\nabla a(g_\varepsilon(u_\varepsilon))|^2 \leq |\nabla u_\varepsilon|^2 \frac{\beta'}{1+\beta'}\big(g_\varepsilon(u_\varepsilon)\big)(g_\varepsilon)'(u_\varepsilon)^2 \leq |\nabla u_\varepsilon|^2 \beta'_\varepsilon(u_\varepsilon),$$

and thus we obtain from (3.13),(3.14):

$$|u_\varepsilon|^2_{L^2(\varrho)} + 2\lambda \int |\nabla a(g_\varepsilon(u_\varepsilon))|^2 \varrho dx \leq |f|^2_{L^2(\varrho)} + C.$$

Since also

$$|a(g_\varepsilon(r))| \leq |g_\varepsilon(r)| = |(I+\varepsilon\beta)^{-1}(r)| \leq |r|,$$

$\{a(g_\varepsilon(u_\varepsilon))\}_{\varepsilon>0}$ is uniformly bounded in $H^1(\varrho)$, thus compact in L^2_{loc}, so there is $\Lambda \in L^2_{\text{loc}}$ such that for a suitable subsequence, the L^2_{loc}-convergence

$$a(g_\varepsilon(u_\varepsilon)) \xrightarrow{\varepsilon\to 0} \Lambda$$

holds, and on a further subsequence this convergence also holds almost surely (we restrict our attention to this subsequence without changing notation). Using the *strict* monotonicity of $r \mapsto \beta(r)$, it follows that $r \mapsto a(r)$ is strictly increasing, hence invertible, so that

$$g_\varepsilon(u_\varepsilon) \xrightarrow{\varepsilon\to 0} a^{-1}(\Lambda)$$

ϱdx-a.s. This yields $\varepsilon\beta(g_\varepsilon(u_\varepsilon)) \xrightarrow{\varepsilon\to 0} 0$ ϱdx-a.s., and hence

$$u_\varepsilon = g_\varepsilon(u_\varepsilon) + \varepsilon\beta(g_\varepsilon(u_\varepsilon)) \xrightarrow{\varepsilon\to 0} a^{-1}(\Lambda) \quad \varrho dx - \text{a.s.}.$$

Therefore, for u from (3.12) we have $u = a^{-1}(\Lambda)$ a.s. and, by (3.11), we conclude $u \in L^\infty$ with

$$|u|_\infty \leq |f|_\infty(1 + |(\text{div}_\varrho D)^- + |D||_\infty^{\frac{1}{2}}),$$

and $u_\varepsilon \xrightarrow{\varepsilon\to 0} u$ in L^p_{loc} for each $p \in [1,\infty)$. Moreover, by Fatou's lemma and the boundedness of $\{u_\varepsilon\}_{\varepsilon>0}$ in each $L^p(\varrho)$, we infer $u \in L^p(\varrho)$ for each $p \in [1,\infty]$. Furthermore, it is easily seen that $\{\tilde{\beta}_\varepsilon\}_{\varepsilon\in(0,1)}$ is locally equicontinuous and hence, since β is locally Lipschitz, the convergence

$$\tilde{\beta}_\varepsilon(u_\varepsilon) \xrightarrow{\varepsilon\to 0} \beta(u) \tag{3.15}$$

holds ϱdx-a.s. and in L^p_{loc}, $p \geq 1$. Hence $\eta = \beta(u)$.

Letting $\varepsilon \to 0$ in (3.8), it follows that $A_\varepsilon u_\varepsilon$ converges in L^p_{loc}, $p \in [1,\infty)$, to some limit $\Psi \in L^1(\varrho) \cap L^\infty$. Moreover, by (3.15), the following convergences hold in $\mathcal{D}'$:

$$\Delta\tilde{\beta}_\varepsilon(u_\varepsilon) \xrightarrow{\varepsilon\to 0} \Delta\beta(u), \quad \nabla\Phi\cdot\nabla\tilde{\beta}_\varepsilon(u_\varepsilon) \xrightarrow{\varepsilon\to 0} \nabla\Phi\cdot\nabla\beta(u), \quad \varepsilon\tilde{\beta}_\varepsilon(u_\varepsilon) \xrightarrow{\varepsilon\to 0} 0.$$

Moreover, since $(b^*_\varepsilon)_{\varepsilon>0}$ is locally equicontinuous and $D_\varepsilon \longrightarrow D$ ϱdx-a.s. with $|D_\varepsilon| \leq |D|$, also

$$\text{div}_\varrho(D_\varepsilon b^*_\varepsilon(u_\varepsilon)) \xrightarrow{\varepsilon\to 0} \text{div}_\varrho(Db^*(u)) \text{ in } \mathcal{D}'.$$

Therefore, the identity

$$\Psi = L^1_{\text{loc}} - \lim_{\varepsilon\to 0} A_\varepsilon u_\varepsilon = -\Delta\beta(u) + \nabla\Phi\cdot\nabla\beta(u) + \operatorname{div}_\varrho(Db^*(u)) = A_0(u)$$

holds in $L^1(\varrho)\cap L^\infty$. In particular, $u\in D(A_0)$ and

$$u + \lambda A_0 u = f$$

holds in $L^1(\varrho)$. Set

$$J_\lambda : L^1(\varrho)\cap L^\infty \to D(A_0),\quad f\mapsto J_\lambda f := u = u(\lambda, f).$$

By the L^1_{loc}-convergence $u_\varepsilon \xrightarrow{\varepsilon\to 0} u$, one obtains (2.8),(2.9) and, by also using (3.9),

$$|J_\lambda f - J_\lambda g|_{1,\varrho} \le |f-g|_{1,\varrho},\quad \forall f,g\in L^1(\varrho)\cap L^\infty.$$

Concerning (2.10), we have for $f\in L^1(\varrho)\cap L^\infty$ and any pair $(\varepsilon,\lambda)\in(0,\infty)\times(0,\lambda_0)$

$$f\ge 0 \text{ a.s.} \implies u_\varepsilon(f)\ge 0 \ \varrho dx - \text{a.s.}$$

Indeed, for $f\ge 0$, multiplying (3.8) with u_ε in place of u by $\chi_\delta(u_\varepsilon^-)$, integrating with respect to ϱdx, and using the monotonicity of $\tilde\beta_\varepsilon$ and $\chi'_\delta\ge 0$ on $[0,\infty)$ gives

$$\int_{\mathbb{R}^d} u_\varepsilon\chi_\delta(u_\varepsilon^-)\,\varrho dx \ge \frac{\lambda}{\delta}\int_{\{u_\varepsilon\in[-\delta,0]\}} D_\varepsilon b_\varepsilon^*(u_\varepsilon)\cdot\nabla u_\varepsilon^-\ \varrho dx \ge -\lambda|D_\varepsilon|_\infty|b|_\infty\int_{\{u_\varepsilon\in[-\delta,0]\}}|\nabla u_\varepsilon^-|\,\varrho dx \xrightarrow{\varepsilon\to 0} 0.$$

Consequently, we find

$$\int_{\{u_\varepsilon\le 0\}} u_\varepsilon\,\varrho dx = \liminf_{\varepsilon\to 0}\int_{\mathbb{R}^d} u_\varepsilon\chi_\delta(u_\varepsilon^-)\,\varrho dx \ge 0,$$

and thus also $u = \lim_{\varepsilon\to 0} u_\varepsilon \ge 0$ ϱdx-a.s. Furthermore, for non-negative $f\in L^1(\varrho)\cap L^\infty$ and $\varphi\in C_c^\infty$, Lemma 2.9 gives

$$\begin{aligned}\int J_\lambda(f)\varphi\,\varrho dx &= \lim_{\varepsilon\to 0}\int_{\mathbb{R}^d} u_\varepsilon\varphi\,\varrho dx\\ &= \int f\varphi\,\varrho dx + \lambda\lim_{\varepsilon\to 0}\left(\int_{\mathbb{R}^d}\beta(u_\varepsilon)(\Delta\varphi - \nabla\Phi\cdot\nabla\varphi)\,\varrho dx + \int_{\mathbb{R}^d} D_\varepsilon b_\varepsilon^*(u_\varepsilon)\cdot\nabla\varphi\,\varrho dx\right)\\ &= \int f\varphi\,\varrho dx + \lambda\int_{\mathbb{R}^d}\beta(J_\lambda f)\big(\Delta\varphi - \nabla\Phi\cdot\nabla\varphi\big)\,\varrho dx + \int_{\mathbb{R}^d} Db^*(J_\lambda f)\cdot\nabla\varphi\,\varrho dx.\end{aligned} \tag{3.16}$$

Since $J_\lambda(f),\beta(J_\lambda f), Db^*(J_\lambda f)\in L^\infty$, choosing $\varphi=\varphi_l$ such that $0\le\varphi_l\le 1$, $\varphi_l(r)=1$ for $|r|\le l$ and $|\nabla\varphi_l|+|\Delta\varphi_l|\longrightarrow 0$ as $l\to\infty$ such that $|\nabla\varphi_l|+|\Delta\varphi_l|$ are bounded in L^∞ uniformly in l and considering (3.16) for $l\to\infty$ gives

$$\int J_\lambda(f)\varrho dx = \int f\varrho dx,$$

where we also used $-\varrho\nabla\Phi = \nabla\varrho\in L^1(\mathbb{R}^d,\mathbb{R}^d)$. Hence, (2.10) holds. Finally, concerning (2.11), let $g\in C_c^\infty$. For all pairs $(\varepsilon,\lambda)\in(0,\infty)\times(0,\lambda_0)$, we have $g\in D(A_\varepsilon)$ and

$$g - \lambda\Delta\tilde\beta_\varepsilon(g) + \lambda\nabla\Phi\cdot\nabla\tilde\beta_\varepsilon(g) + \lambda\varepsilon\tilde\beta_\varepsilon(g) + \operatorname{div}_\varrho(D_\varepsilon b_\varepsilon^*(g)) = g + \lambda A_\varepsilon g.$$

Since the right-hand side is in $L^1(\varrho)\cap L^2(\varrho)$ and $g\in L^2(\varrho)$ with $\tilde\beta_\varepsilon(g)\in D(L)$, we have $g = u_\varepsilon(\lambda, g+\lambda A_\varepsilon g)$, which implies

$$|g - u_\varepsilon(\lambda,g)|_{L^1(\varrho)} \le \lambda|A_\varepsilon g|_{1,\varrho}.$$

Since $\beta \in C^2$, it follows that $\tilde\beta_\varepsilon$, $\tilde\beta'_\varepsilon$ and $\tilde\beta''_\varepsilon$ are locally bounded uniformly in $\varepsilon \in (0,1)$, and hence

$$|A_\varepsilon g|_{1,\varrho} \le |\Delta\tilde\beta_\varepsilon(g)|_{1,\varrho} + |\nabla\Phi\cdot\nabla\tilde\beta_\varepsilon(g)|_{1,\varrho} + \varepsilon|\tilde\beta_\varepsilon(g)|_{1,\varrho} + |\mathrm{div}_\varrho(D_\varepsilon b^*_\varepsilon(g)|_{1,\varrho} \le C(|g|_{W^{2,1}(\varrho)} + |g|_{H^2(\varrho)}),$$

where $C>0$ depends on g,β,Φ,b and D, but not on $\varepsilon\in(0,1)$ or $\lambda>0$. Hence by the L^1_{loc}-convergence of $u_\varepsilon(\lambda,g)$ to $J_\lambda g$, we conclude (2.11), which concludes the proof of Lemma 2.6. □

Now we complete the proof of Theorem 2.4 by showing that the generalized solutions u are also solutions to (1.1) in distributional sense.

Proof of Theorem 2.4 continued. Denote by $u = u(u_0)$ the generalized solution to (1.1) given by Theorem 2.4, i.e. $u(t)$ is the (locally uniform in t) $L^1(\varrho)$-limit as $h\to 0$ of $u_h(t)$, where u_h is the step function from (2.2) (with A as in(2.12) instead of $\tilde A$), with $u_h^i = J_h^{-i}u_0$. We have (setting $u_h(s) := u_0$ for $s\in(-\infty,0)$)

$$u_h(t) + hA(u_h(t)) = u_h(t-h), \quad t>0,$$

in $L^1(\varrho)$, which after multiplying with $\varphi\in C_c^\infty([0,\infty)\times\mathbb{R}^d)$ and integrating with respect to $dt\otimes\varrho dx$ gives

$$\int_0^\infty\int_{\mathbb{R}^d} h^{-1}(u_h(t,x)-u_h(t-h,x))\varphi(t,x) - \beta(u_h(t,x))L\varphi(t,x)\,\varrho dxdt \qquad (3.17)$$
$$-\int_{\mathbb{R}^d} Db^*(u_h(t,x))\cdot\nabla\varphi(t,x)\,\varrho dx = 0.$$

Because the local uniform $L^1(\varrho)$-convergence in t of $u_h(t)$ to $u(t)$ as $h\to 0$ implies, using (2.4), $\sup_{0<h<h_1}|u_h(t)|_\infty \le (1+C)|u_0|_\infty$ for sufficiently small $h_1>0$ and $C>0$ (see also Remark 2.8 (i)), we obtain $\beta(u_h(t))\to\beta(u(t))$ in $L^1(\varrho)$ locally uniformly in t. Moreover,

$$\int_0^\infty\int_{\mathbb{R}^d}\frac{(u_h(t,x)-u_h(t-h,x)}{h}\varphi(t,x)\,\varrho dxdt$$
$$= -\int_0^\infty\int_{\mathbb{R}^d}\frac{\varphi(t+h,x)-\varphi(t,x)}{h}u_h(t,x)\,\varrho dxdt - \int_{-h}^0\int_{\mathbb{R}^d}\frac{\varphi(t+h,x)}{h}u_0(x)\,\varrho dxdt,$$

which, as $h\to 0$, by the locally uniform in t convergence of u_h to u in $L^1(\varrho)$, converges to

$$-\int_0^\infty\int_{\mathbb{R}^d}\partial_t\varphi(t,x)u(t)\,\varrho dxdt - \int_{\mathbb{R}^d}u_0\varphi(0,x)\,\varrho dx.$$

Consequently, letting $h\to 0$ in (3.17) yields

$$\int_0^\infty\int_{\mathbb{R}^d}\bigg(\partial_t\varphi(t,x)+\frac{\beta(u(t,x)}{u(t,x)}L\varphi(t,x)+D(x)b(u(t,x))\cdot\nabla\varphi(t,x)\bigg)u(t,x)\varrho dxdt+\int_{\mathbb{R}^d}u_0\varphi(0,x)\,\varrho dx = 0.$$

Equivalently (see [8, Thm.6.1.2]), $v(t,x) := u(t,x)\varrho(x)$ satisfies for all $t\in(0,\infty)$

$$\int_{\mathbb{R}^d}\varphi(t)\,v(t)dx = \int_{\mathbb{R}^d}\varphi(0)\,v(0)dx$$
$$+\int_0^t\int_{\mathbb{R}^d}\frac{\beta(v(s)\varrho^{-1})}{v(s)\varrho^{-1}}\Delta\varphi + \bigg(Db(v(s)\varrho^{-1}) - \frac{\beta(v(s)\varrho^{-1})}{v(s)\varrho^{-1}}\nabla\Phi\bigg)\cdot\nabla\varphi\,v(s)dxds,$$

i.e. $t\mapsto v(t,x)dx$ is a weakly continuous distributional solution to (1.1) with initial datum $u_0\varrho dx$. This completes the proof of Theorem 2.4. □

Finally, we give the proof of Proposition 2.5.

Proof of Proposition 2.5. It suffices to extend Lemma 2.6 from $L^1(\varrho)\cap L^\infty$ to $L^1(\varrho)$. More precisely, we slightly change the definition of the domain of A_0 to

$$\tilde{D}(A_0) := \big\{ f\in L^1(\varrho) \,|\, -L_0\beta(f)+\mathrm{div}_\varrho(Db(f)f)\in L^1(\varrho)\big\},$$

let $0<\lambda<\lambda_0$ with λ_0 as in Lemma 2.6, and we show: $R(I+\lambda A_0)=L^1(\varrho)$, J_λ extends to $L^1(\varrho)$ such that

$$|J_\lambda f - J_\lambda g|_{1,\varrho} \le |f-g|_{1,\varrho}, \quad \forall f,g\in L^1(\varrho) \tag{3.18}$$

and (2.8) extends to $f\in L^1(\varrho)$. Indeed, once this is proven, all assertions of the proposition but the final one follows directly from the Crandall–Liggett nonlinear semigroup theory, this time applied on the full Banach space $L^1(\varrho)$ (compare the first part of the proof of Theorem 2.4) in Section 2). That the mild solutions are also distributional solutions follows as in the final part of the proof of Theorem 2.4 by noting that here the convergence of $\beta(u_h)$ to $\beta(u)$ locally uniformly in t as $h\to 0$ follows from (H2'). For $f\in L^1(\varrho)$, let $(f_n)_{n\ge1}\subseteq L^1(\varrho)\cap L^\infty$ such that $f_n \longrightarrow f$ in $L^1(\varrho)$ as $n\to\infty$. Then the contraction property of J_λ implies the existence of an $L^1(\varrho)$-limit poin $u=u(\lambda,f)$ of $(J_\lambda f_n)_{n\ge1}$. Since $u_n+\lambda A_0u_n=f_n$, it follows that $(A_0u_n)_{n\ge1}$ has an $L^1(\varrho)$-limit point Ψ. By (H2'), $\beta(u_n)\longrightarrow\beta(u)$ in $L^1(\varrho)$ as $n\to\infty$, hence $L_0\beta(u_n)\longrightarrow L_0\beta(u)$ in $\mathcal{D}'$ as $n\to\infty$. Since, due to the boundedness of b and D, also $Db^*(u_n)\longrightarrow Db^*(u)$ in $L^1(\varrho)$ as $n\to\infty$, we obtain $A_0u_n\longrightarrow -L_0\beta(u)+\mathrm{div}_\varrho(Db^*(u))$ in $\mathcal{D}'$. Hence, $u\in\tilde{D}(A_0)$, $\Psi=A_0u$, and thus $u=u(\lambda,f)\in(I+\lambda A_0)^{-1}f$. Extending J_λ to $L^1(\varrho)$ via $J_\lambda f := u(\lambda,f)$, it is easily seen that (3.18) holds and (2.8) extends to all $f\in L^1(\varrho)$.

Concerning the final assertion, note that $S(t)u_0=\lim_{n\to\infty}(J_{\frac{t}{n}})^n u_0$, let $f\ge0$, $|f|_{1,\varrho}=1$, and set $f_n:=\max(f,n)\in L^1(\varrho)\cap L^\infty$, $n\in\mathbb{N}_0$. We know from the proof of Theorem 2.4 that $f_n\ge0 \implies J_\lambda f_n\ge 0$ and, for such f_n, $|J_\lambda f_n|_{1,\varrho}=|f_n|_{1,\varrho}$. Hence, since $J_\lambda f = L^1(\varrho)-\lim_{n\to\infty}J_\lambda f_n$, we obtain $J_\lambda f\ge 0$ and $|J_\lambda f|_{1,\varrho}=|f|_{1,\varrho}$, which implies $S(t)u_0\,\varrho dx\in\mathcal{P}$, if $u_0\varrho dx\in\mathcal{P}$. □

4 Nonlinear perturbed Ornstein–Uhlenbeck processes

Let Hypothesis 1 be satisfied.

4.1 Existence

In this section, we solve the McKean–Vlasov SDE (1.2) with $1D$-time marginals given by the mild solutions to (1.1). As said in the introduction, equation (1.2) can be regarded a model for generalized nonlinear perturbed Ornstein–Uhlenbeck processes, since for $D=0$, $\Phi=-\frac{|x|^2}{2}$ and $\beta(r)=\frac{\sigma^2}{2}r$, $\sigma>0$, one recovers the classical Ornstein–Uhlenbeck SDE

$$dX_t = -X_t dt + \sigma dB_t.$$

Definition 4.1. A *(probabilistically weak) solution* to (1.2) is a triple consisting of a filtered probability space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\ge0},\mathbb{P})$, an $\mathbb{R}^d$-valued $(\mathcal{F}_t)$-Brownian motion B and an $(\mathcal{F}_t)$-adapted stochastic process $X=(X_t)_{t\ge0}$ on Ω such that $\mathbb{P}\circ X_t^{-1}=v(t,x)dx$,

$$\int_0^T D(X_t)b\big(v(t,X_t)\varrho^{-1}(X_t)\big)-\frac{\beta(v(t,X_t)\varrho^{-1}(X_t))}{v(t,X_t)\varrho^{-1}(X_t)}\nabla\Phi(X_t)dt \text{ and } \int_0^T \frac{\beta(v(t,X_t)\varrho^{-1}(X_t))}{v(t,X_t)\varrho^{-1}(X_t)}dt$$

belong to $L^1(\Omega;\mathbb{P})$ for all $T>0$, and $\mathbb{P}$-a.s.

$$\begin{aligned} X_t = X_0 + \int_0^t D(X_t)b\big(v(t,X_t)\varrho^{-1}(X_t)\big) - \frac{\beta(v(t,X_t)\varrho^{-1}(X_t))}{v(t,X_t)\varrho^{-1}(X_t)}\nabla\Phi(X_t)dt \\ + \int_0^t \sqrt{2\frac{\beta(v(t,X_t)\varrho^{-1}(X_t))}{v(t,X_t)\varrho^{-1}(X_t)}}dB_t, \quad \forall t\ge0. \end{aligned}$$

Let $u_0 \in \mathcal{P}_\infty(\varrho)$, $u = u(u_0)$ be the generalized solution to (1.1) from Theorem 2.4 with $u(0) = u_0$ and consider the curve

$$t \mapsto \nu(t, u_0) := u(t)\varrho dx, \tag{4.1}$$

i.e. ν is a weakly continuous distributional probability solution to (1.1) in the sense of Definition 2.3. We have

$$\int_0^t \int_{\mathbb{R}^d} \left|\frac{\beta(v(s)\varrho^{-1})}{v(s)\varrho^{-1}}\right| + \left|b(v(s)\varrho^{-1})D - \frac{\beta(v(s)\varrho^{-1})}{v(s)\varrho^{-1}}\nabla\Phi\right| v(s)dxds$$
$$\leq \int_0^t \int_{\mathbb{R}^d} \big|\beta(u(s))\big| + \big|b(u)uD\big| + \big|\beta(u(s))\nabla\Phi\big|\,\varrho dxds < \infty,$$

where the second inequality holds due to $u \in L^\infty_{\text{loc}}([0,\infty), L^\infty)$ (which follows from (2.4)), the boundedness of D and b, the local Lipschitz-continuity of β, and $\nabla\Phi \in L^1(\mathbb{R}^d, \mathbb{R}^d; \varrho)$. Hence the *superposition principle* (see [3, Sect.2] and [14]) yields the existence of a probabilistically weak solution $(X_t)_{t\geq 0}$ to (1.2) such that $\mathcal{L}(X_t) = \nu(t, u_0)$, $t \geq 0$.

If also (H2') holds, this existence result extends as follows. By Proposition 2.5, there is a distributional probability solution $\nu(t, u_0) : t \mapsto S(t)u_0\,\varrho dx$ for all u_0 such that $u_0\,\varrho dx \in \mathcal{P}$. Hence, in this case, for any such u_0 there is a weak solution $X = X(u_0)$ to (1.2) with $\mathcal{L}(X_t) = \nu(t, u_0)$, $t \geq 0$.

4.2 Solutions to nonlinear perturbed Ornstein–Uhlenbeck equations as nonlinear Markov processes

In addition to Hypothesis 1, now we also assume

(H5) $b \in C^1(\mathbb{R})$, and for each compact $K \subseteq \mathbb{R}$ there is $\alpha_K \in (0,\infty)$ such that

$$|b'(r)r + b(r)| \leq \alpha_K|\beta'(r)|, \quad \forall r \in K.$$

If $b \in C^1(\mathbb{R})$ and $\beta'(r) > 0$ for all $r \in \mathbb{R}$, the second part of (H5) is automatically satisfied.
In the previous subsection, we showed the following. For any initial datum $\nu_0 = u_0\varrho(x)dx \in \mathcal{P}_0$, where

$$\mathcal{P}_0 := \{\nu \in \mathcal{P} \,|\, \nu = u_0\varrho(x)dx, u_0 \in \mathcal{P}_\infty(\varrho)\},$$

there is a weak solution $X = X(u_0)$ to (1.2) such that its curve of $1D$-time marginals is given by (4.1). By (2.4),(2.5) it follows that $\nu(t, u_0) \in \mathcal{P}_0$, $u \in L^\infty_{\text{loc}}((0,\infty), L^\infty)$, and

$$\nu(t+s, u_0) = \nu(t, u(s)), \quad \forall t, s \geq 0, u_0 \in \mathcal{P}_\infty(\varrho). \tag{4.2}$$

The aim of this subsection is to prove that the path law family $\{\mathbb{P}_{u_0}\}_{u_0 \in \mathcal{P}_\infty(\varrho)}$, $\mathbb{P}_{u_0} := \mathcal{L}(X(u_0))$ constitutes a *nonlinear Markov process* in the sense of [11, 12]. More precisely, we want to show that the following *nonlinear Markov property* holds, compare [12, Def.2.1]. Denote by $\pi_t : C([0,\infty), \mathbb{R}^d) \to \mathbb{R}^d$, $\pi_t(w) = w(t)$, the canonical projections and set $\mathcal{F}_t := \sigma(\pi_s, 0 \leq s \leq t)$, then the nonlinear Markov property is

$$\mathbb{P}_{u_0}(\pi_t \in A|\mathcal{F}_r) = p_{u_0,r,\pi_r}(\pi_{t-r} \in A), \quad \forall 0 \leq r \leq t, u_0 \in \mathcal{P}_\infty(\varrho), A \in \mathcal{B}(\mathbb{R}^d).$$

Here $(p_{u_0,r,y}(dw))_{y\in\mathbb{R}^d}$ is a regular conditional probability kernel from $\mathbb{R}^d$ to $\mathcal{B}(C([0,\infty), \mathbb{R}^d))$ of $\mathbb{P}_{u(r)}[\,\cdot\,|\pi_0 = y]$, $y \in \mathbb{R}^d$, where $u(r) = u(u_0)(r)$. First, we observe

$$\nu_0 \in \mathcal{P}_0, \tilde{\nu}_0 \in \mathcal{P}, \tilde{\nu} \leq C\nu_0 \text{ for some } C \geq 1 \implies \tilde{\nu}_0 \in \mathcal{P}_0.$$

Therefore, by (4.2) and [12, Thm.3.7, Cor.3.8], it suffices to prove that the *linearized Fokker–Planck equations*, obtained by fixing a priori the densities $t \mapsto u(t)\varrho$ in the measure component of the coefficients in the distributional formulation of (1.1), are well-posed in a suitable subclass, see Lemma (4.3) after the following definition.

Definition 4.2. Let $s \geq 0$ and $y \in L^\infty_{\text{loc}}((s,\infty), L^\infty)$. A weakly continuous curve $\zeta : [s,\infty) \to \mathcal{P}$ is a *(distributional) probability solution to the $y\varrho$-linearized version of* (1.1) with initial datum ζ_s, if $\int_s^T \int_{\mathbb{R}^d} |\nabla\Phi| \, d\zeta_t dt < \infty$ for all $T \geq s$, and

$$\int_{\mathbb{R}^d} \varphi \, d\zeta_t = \int_{\mathbb{R}^d} \varphi \, d\zeta_s + \int_s^t \int_{\mathbb{R}^d} \frac{\beta(y)}{y} L\varphi + b(y)D \cdot \nabla\varphi \, d\zeta_s ds, \quad \forall t \geq s, \varphi \in C_c^\infty. \tag{4.3}$$

Equation (4.3) is obtained by a priori replacing $v(s)$ in (2.3) by the fixed curve $s \mapsto y(s)\varrho$. Note that since b and D are bounded, and since $r \mapsto \frac{\beta(r)}{r}$ is locally bounded and $y \in L^\infty_{\text{loc}}((s,\infty), L^\infty)$, the local-global integrability condition for all coefficients except for $\nabla\Phi$ is automatically satisfied, i.e.

$$\int_s^T \int_{\mathbb{R}^d} \Big|\frac{\beta(y)}{y}\Big| + \big|b(y)D\big| \, d\zeta_t dt < \infty, \quad \forall T \geq s,$$

for any measurable curve $t \mapsto \zeta_t \in \mathcal{P}$.

Lemma 4.3. *Let $\nu_0 \in \mathcal{P}_0$, $\nu_0 = u_0\varrho dx$, and $\nu(t,u_0) = u(t)\varrho dx$ be as in* (4.1)*. For each $s \geq 0$, $t \mapsto \nu(t,u_0)$, $t \geq s$, is the unique probability solution in $L^\infty_{\text{loc}}((s,\infty), L^\infty)$ to the $u\varrho$-linearized version of* (1.1) *in the sense of the previous definition with initial datum $\nu(s,u_0)$.*

Proof. The assertion follows analogously to the proof of [5, Thm.4.1]. Indeed, under our assumptions (H1)-(H5), the proof of that result, as well as the proof of its nonlinear version, Theorem 3.2. of the same reference, can be repeated for the operator L and the spaces $H^k(\varrho)$ instead of Δ and H^k, $k \in \mathbb{N}_0$, respectively. Indeed, since our assertion is restricted to probability solutions, we prove uniqueness of solutions in $L^\infty_{\text{loc}}((s,\infty), L^\infty \cap L^1)$, while in [5] uniqueness is even proven in the larger class of solutions in $L^\infty_{\text{loc}}((0,\infty), H^{-1})$. □

Finally, we obtain the desired result.

Proposition 4.4. *There is a nonlinear Markov process $\{\mathbb{P}_{u_0}\}_{u_0 \in \mathcal{P}_\infty(\varrho)}$, which consists of path laws of solutions to* (1.2) *and whose $1D$-time marginals are given by $u(t)\varrho dx$, where $u(t) = S(t)$, $t \geq 0$, is the flow of solutions to* (1.1) *constructed in Theorem 2.4. Moreover, for each $u_0 \in \mathcal{P}_\infty(\varrho)$, $\mathbb{P}_{u_0}$ is the only solution law to* (1.2) *with initial datum $u_0\varrho dx$ and $1D$-marginals $\nu(t,u_0)$, $t \geq 0$.*

Proof. The result follows from 4.2, Lemma 4.3 and [12, Cor.3.8]. □

Acknowledgements. Funded by the German Research Foundation (DFG) - Project number 517982119. The author would like to thank Michael Röckner for valuable discussions, as well as the research group of Franco Flandoli in Pisa for their hospitality.

References

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

[2] V. Barbu and M. Röckner. Solutions for nonlinear Fokker–Planck equations with measures as initial data and McKean-Vlasov equations. *Journal of Functional Analysis*, 280(7):108926, 2021.

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

[4] V. Barbu and M. Röckner. Nonlinear Fokker–Planck equations with fractional Laplacian and McKean-Vlasov SDEs with Lévy-noise. *arXiv preprint 2210.05612*, 2022.

[5] V. Barbu and M. Röckner. Uniqueness for nonlinear Fokker–Planck equations and for McKean-Vlasov SDEs: The degenerate case. *arXiv preprint 2203.00122*, 2022.

[6] V. Barbu and M. Röckner. The evolution to equilibrium of solutions to nonlinear Fokker–Planck equation. *Indiana Univ. Math. J.*, 72(1):89–131, 2023.

[7] V. Barbu, M. Röckner, and J. L. da Silva. Nonlocal, nonlinear Fokker–Planck equations and nonlinear martingale problems. *arXiv preprint 2308.06388*, 2023.

[8] V.I. Bogachev, N.V. Krylov, M. Röckner, and S.V. Shaposhnikov. *Fokker–Planck–Kolmogorov Equations*. Mathematical Surveys and Monographs 207. American Mathematical Society, 2015.

[9] M. G. Crandall and T. M. Liggett. Generation of semi-groups of nonlinear transformations on general banach spaces. *American Journal of Mathematics*, 93(2):265–298, 1971.

[10] G. Da Prato and A. Lunardi. Elliptic operators with unbounded drift coefficients and neumann boundary condition. *Journal of Differential Equations*, 198(1):35–52, 2004.

[11] H. P. McKean. A class of Markov processes associated with nonlinear parabolic equations. *Proceedings of the National Academy of Sciences of the United States of America*, 56(6):1907–1911, 12 1966.

[12] M. Rehmeier and M. Röckner. On nonlinear Markov processes in the sense of McKean. *arXiv preprint 2212.12424*, 2022.

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

[14] D. Trevisan. Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients. *Electron. J. Probab.*, 21:41 pp., 2016.