OCTAV MAYER INSTITUTE OF MATHEMATICS
ROMANIAN ACADEMY, IAŞI BRANCH
Bd. Carol I nr. 10, 700506, Iaşi, Romania

PREPRINT SERIES OF THE
OCTAV MAYER INSTITUTE OF MATHEMATICS

The evolution to equilibrium of solutions to nonlinear Fokker-Plank equation

Viorel Barbu
Octav Mayer Institute of Mathematics of the Romanian Academy, Iaşi, Romania

Michael Röckner
Fakultät für Mathematik, Universität Bielefeld, D33501, Germany
Academy of Mathematics and System Sciences, CAS, Beijing

Nr. 01-2020

The evolution to equilibrium of solutions to nonlinear Fokker-Planck equation

Viorel Barbu* Michael Röckner†‡

Abstract

One proves the H-theorem for mild solutions to a nondegenerate, nonlinear Fokker-Planck equation

$$u_t - \Delta\beta(u) + \mathrm{div}(D(x)b(u)u) = 0,\ t \geq 0,\ x \in \mathbb{R}^d, \tag{1}$$

and under appropriate hypotheses on β, D and b the convergence in $L^1_{\mathrm{loc}}(\mathbb{R}^d)$, $L^1(\mathbb{R}^d)$, respectively, for some $t_n \to \infty$ of the solution $u(t_n)$ to an equilibrium state of the equation for a large set of nonnegative initial data in L^1. These results are new in the literature on nonlinear Fokker-Planck equations arising in the mean field theory and are also relevant to the theory of stochastic differential equations. As a matter of fact, by the above convergence result, it follows that the solution to the McKean–Vlasov stochastic differential equation corresponding to (1), which is a *nonlinear distorted Brownian motion*, has this equilibrium state as its unique invariant measure.

Keywords: Fokker-Planck equation, m-accretive operator, probability density, Lyapunov function, H-theorem, McKean–Vlasov stochastic differential equation, nonlinear distorted Brownian motion.
2010 Mathematics Subject Classification: 35B40, 35Q84, 60H10.

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

1 Introduction

We shall study here the asymptotic behaviour of solutions $u = u(t, x)$ to the nonlinear Fokker-Planck equation

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

under the following hypotheses on the functions $\beta : \mathbb{R} \to \mathbb{R}$, $D : \mathbb{R}^d \to \mathbb{R}^d$ and $b : \mathbb{R} \to \mathbb{R}$, where $1 \le d < \infty$.

(i) $\beta \in C^1(\mathbb{R}),\ \beta(0) = 0,\ \gamma \le \beta'(r) \le \gamma_1,\ \forall r \in \mathbb{R}$, for $0 < \gamma < \gamma_1 < \infty$.

(ii) $b \in C_b(\mathbb{R}) \cap C^1(\mathbb{R})$.

(iii) $D \in C_b(\mathbb{R}^d;\mathbb{R}^d) \cap W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)$.

(iv) $D = -\nabla\Phi$, *where* $\Phi \in C^1(\mathbb{R}^d)$, $\Phi \ge 1$, $\lim\limits_{|x|_d\to\infty} \Phi(x) = +\infty$ *and there exists* $m \in [2,\infty)$ *such that* $\Phi^{-m} \in L^1(\mathbb{R}^d)$.

A typical example is

$$\Phi(x) = C(1+|x|^2)^\alpha,\ x\in\mathbb{R}^d, \tag{1.2}$$

with $\alpha \in \left(0, \frac{1}{2}\right]$.

If (i)–(iv) hold, we prove the existence of solutions given by a nonlinear semigroup $S(t)$, $t > 0$, of contractions in $L^1(\mathbb{R}^d)$ (Theorem 4.1), which is positivity and mass preserving. If, in addition to (i)–(iv), also (v) holds, where

(v) $b(r) \ge b_0 > 0$ for $r \ge 0$,

we prove the convergence of the solutions to equilibrium in $L^1_{\mathrm{loc}}(\mathbb{R}^d)$, while (see Theorem 6.1) the convergence in $L^1(\mathbb{R}^d)$ is proved if, in addition to (i)–(v), the following condition holds

(vi) $\gamma_1\Delta\Phi - b_0|\nabla\Phi|^2 \le 0$.

Examples for Φ satisfying (vi) are such that $\Phi =$ const. (≥ 1) on a ball of radius R_1 around zero and Φ behaves like Φ in (1.2) outside a ball around zero of radius $R_2 > R_1$, where R_1 and R_2 are properly chosen depending on γ_1 and b_0.

Equation (1.1), where u is a probability density, is known in the literature as the nonlinear Fokker-Planck equation (NFPE) and it is relevant in the kinetic theory of statistical mechanics as a generalized mean field Smoluchowski equation for the case where the diffusion and transport coefficients depend on the density u. (See [13], [18]–[20] [26].) The case of the classical Smoluchowski equation is recovered for $b \equiv 1$ and $\beta(r) \equiv r$. In the case where the first order part in (1.1) is given by a vector field independent of the spatial variable x, the existence and uniqueness of a kinetic, respectively generalized entropic, solution to (1.1) in $L^1(\mathbb{R}^d)$ was proved in [14]. In this paper, we give an existence and uniqueness result for (1.1) in the sense of mild solutions in $L^1(\mathbb{R}^d)$, i.e., given as a nonlinear semigroup $S(t)$, $t > 0$, in $L^1(\mathbb{R}^d)$ (see Proposition 2.2). Its proof is different from that in [14] and, though it has an intrinsic interest in itself, it is used subsequently to prove our main result about convergence to equilibrium and existence of a unique stationary solution to (1.1). In [6] (see, also, [4], [5]), a more general NFPE of the form

$$u_t - \sum_{i,j=1}^{d} D_{ij}^2(a_{ij}(x,u)u) + \operatorname{div}(b(x,u)u) = 0 \tag{1.3}$$

was studied under appropriate assumptions on $a_{ij} : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ and $b : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$. In the latter case, it is shown that, if u_0 is a probability density, the distributional mild solution u to (1.3) is the probability density of the law $\mathcal{L}_{X(t)}$ of the (probabilistically) weak solution to the McKean–Vlasov stochastic differential equation (SDE)

$$dX(t) = b(X(t), u(t, X(t)))dt + \sqrt{2}\,\sigma(X(t), u(t, X(t)))dW(t), \tag{1.4}$$

where $\sigma\sigma^{\perp} = \frac{1}{2}\,(a_{ij})_{i,j=1}^{d}$ and $X(0)$ has law $u_0 dx$, where $dx =$ the Lebesgue measure on $\mathbb{R}^d$.

In the special case (1.1), SDE (1.4) reduces to

$$dX(t) = D(X(t))b(u(t, X(t)))dt + \frac{1}{\sqrt{2}} \left(\frac{\beta(u(t, X(t)))}{u(t, X(t))}\right)^{\frac{1}{2}} dW(t), \tag{1.5}$$

which, since $D = -\nabla\Phi$, is a nonlinear analogue of the SDE for the classical distorted Brownian motion, where $\beta = id$ and $b \equiv const$. Hence, its solution $X(t)$, $t \geq 0$, can be considered as a nonlinear distorted Brownian motion.

One of our motivations is to apply our asymptotic results *to find an invariant* (*probability*) *measure for the nonlinear distorted Brownian motion* *on* $\mathbb{R}^d$. So, Theorems 6.1 and 6.4 solve this problem and this is one of the main contributions of this work. Condition (vi) requires a certain balance between the strength of the (in general nonlinear) diffusion coefficient β' and the strength of the nonlinear drift coefficient b in terms of the "potential" Φ. Without the additional condition (vi), there is in general no equilibrium on $L^1(\mathbb{R}^d)$ for equation (1.1). Just consider the linear case $\beta = id$ and $D \equiv 0$, so the case where (1.1) is the heat equation. Hence, as in the linear case, we need a big enough "negative" drift. Condition (vi) is, however, not optimal, because for the Fokker-Planck equation associated to the classical Ornstein-Uhlenbeck process on $\mathbb{R}^d$, it does not hold, though the standard Gaussian measure is its equilibrium measure.

We would like to mention here another special case of (1.1), namely with $\beta(u) = u^m$, $m > 1$, $b \equiv const.$ and $D(x) = x$, which is not covered by our results, but was deeply analyzed in [12]. In this case, the equilibrium is given through an explicit formula and the decay rate in L^1-distance is calculated in [12]. So, the approach is completely different from ours which is to prove the so-called H-theorem (see below) to show convergence of solutions to a unique equilibrium of (1.1) in $L^1(\mathbb{R}^d)$ as $t \to \infty$. A general result combining [12], the linear case and ours including convergence rates is still to be proved and will be subject to our future study. As explained in detail in [6, Section 2], the nonlinear Fokker-Planck equation (1.1) is a (very singular) special case (called "Nemytskii type") of a general nonlinear Fokker-Planck-Kolmogorov equation in the sense of Section 6.7(iii) in [8] and of [22], [23], where the solutions are measure-valued and the coefficients depend on these solutions. There is a number of papers where existence of and convergence to equilibria are studied (see, e.g., [9] and [17] and the references therein). However, in these papers the dependence of the coefficients on the measures is assumed to be linear or Lipschitz continuous in weighted variation norm, which is never fulfilled in our Nemytskii-type case. So, these results do not apply here.

The main objective of this work is to study the asymptotic behaviour of a solution $t \to u(t)$ for $t \to \infty$ and prove the so called H-theorem for the NFPE (1.1), that is, prove the existence of a Lyapunov function $V : D(V) \subset L^1_{\text{loc}}(\mathbb{R}^d) \to \mathbb{R}$ for (1.1) and prove, for a certain class of $u_0 \in L^1$, $u_0 \geq 0$, the ω-limit set

$$\omega(u_0) = \left\{ w = \lim_{n\to\infty} u(t_n) \text{ in } L^1_{\text{loc}}(\mathbb{R}^d),\ \{t_n\} \to \infty \right\} \tag{1.6}$$

is nonempty. This is proved in Sections 4 and 5 and under assumptions (i)–(v). Moreover, if (vi) also holds, we shall prove in Section 6 that $\omega(u_0)$ reduces to a single element u_∞, which is a stationary solution to (1.1). Furthermore, u_∞ is a probability density, if so is u_0. As a consequence, $u_\infty dx$ is an invariant measure for SDE (1.5), i.e., if $u_0 = u_\infty$, then *the nonlinear distorted Brownian motion $X(t)$, $t \geq 0$, has the law $u_\infty dx$, $\forall t \geq 0$.*

The H-theorem amounts to saying that the function

$$V(u) = -S[u] + E[u],\ u \in L^1(\mathbb{R}^d), \tag{1.7}$$

where S is the entropy of the system and E is the mean field energy, is a Lyapunov function for (1.1), that is, monotonically decreasing in time on the solutions to (1.1). In our case,

$$S[u] = \int_{\mathbb{R}^d} \eta(u(x))dx,\ E(u) = \int_{\mathbb{R}^d} \Phi(x)u(x)dx, \tag{1.8}$$

where

$$\eta(r) = -\int_0^r d\tau \int_\tau^1 \frac{\beta'(s)}{sb(s)}\, ds,\ r \geq 0.$$

This form of the Lyapunov theorem comes from the classical H-theorem and is consistent with the Boltzman thermodynamics (see, e.g., [13]–[19], [26]), in which case $\beta' \equiv b \equiv const.$, so S in (1.8) reduces to the classical Boltzmann-Gibbs entropy.

In the literature on NFPE arising in the mean field theory, the H-theorem is often invoked, but in most cases its proof is formal because, in general, the NFPE (1.1) has not a classical solution and so the computation is not rigorous. By our knowledge, this paper contains the first rigorous mathematical result on the H-theorem for NFPE. In fact, here the basic functional space for the well-posedness is $L^1(\mathbb{R}^d)$ and, in general, the space of the maximal spatial regularity for u is the Sobolev space $W^{1,q}(\mathbb{R}^d)$, $1 < q \leq \frac{d}{d-2}$, (which happens in the special case of the porous media equation $b \equiv 0$, $a_{ij}(u)u \equiv \delta_{ij}\beta(u)$). This low regularity precludes the classical argument involving regular Lyapunov functions. However, the situation is different for linear FPE where, in the last decades, many convergence results to equilibrium were obtained. We refer to the monographs [2], [28] and, e.g., to [1], [24], as well as the references therein.

Let us now explain the structure of the paper. The first part is concerned with the well-posedness of NFPE (1.1) in $L^1(\mathbb{R}^d)$ via the theory of nonlinear semigroups of contractions in $L^1(\mathbb{R}^d)$, i.e., the construction of such a semigroup $S(t)$, $t > 0$, so that $t \mapsto S(t)u_0$ a continuous function $u : [0, \infty) \to L^1(\mathbb{R}^d)$ given as the limit of the finite difference scheme associated with (1.1) (the so called *mild* solution). Moreover, u is obtained as the limit in $L^1(\mathbb{R}^d)$ of the smooth solutions $\{u_\varepsilon\}_{\varepsilon>0}$ to an approximating equation associated with (1.1). The corresponding result given in Proposition 2.1 is not essentially new since, as mentioned earlier, a similar existence result was previously established in [4], [5], [6], [14]. However, we have developed here a semigroup approach to NFPE (1.1) necessary for the treatment of the asymptotic behaviour of solutions. In fact, in the second part of the work we shall prove under assumptions (i)–(v) the H-theorem for (1.1) (Theorem 4.1). The ω-limit set is a singleton $\{u_\infty\}$ and the invariant measure of the solution $X(t)$, $t \geq 0$, of SDE (1.5) if, additionally, the balance (vi) holds (Theorem 6.1). A main point to prove the latter is to show that $S(t)$ is also a contraction on the weighted L^1 space with the potential Φ from condition (iv) as its weight (see Lemma 6.2).

Finally, we prove that the equilibrium u_∞ from Theorem 6.1 is indeed the unique solution of the stationary version of (1.1) in the sense of distributions (Theorem 6.4) and, as a consequence, that the stationary nonlinear distorted Brownian motion is unique in law (Theorem 6.5).

Notation. For $p \in [1, \infty)$, $L^p(\mathbb{R}^d)$ – simply denoted L^p, is the space of all Lebesgue p-summable functions on $\mathbb{R}^d$. The norm in L^p is denoted by $|\cdot|_p$. Similarly, if $\mathcal{O}$ is a Lebesgue measurable set, $L^p(\mathcal{O})$ is the space of all p-summable functions on $\mathcal{O}$. By $L^p_{\rm loc}(\mathbb{R}^d)$ we denote the space of Lebesgue measurable functions $u : \mathbb{R}^d \to \mathbb{R}$ which are in $L^p(\mathcal{O})$ for every bounded measurable subset $\mathcal{O} \subset \mathbb{R}^d$. ($L^p_{\rm loc}$ is endowed with a standard locally convex metrizable topology.) The scalar product of L^2 is denoted by $\langle \cdot, \cdot \rangle_2$. If $\mathcal{O}$ is an open subset of $\mathbb{R}^d$, we denote by $\mathcal{D}'(\mathcal{O})$ the space of Schwartz distributions on $\mathcal{O}$ and by $W^{1,p}(\mathcal{O})$ the Sobolev space $\{u \in L^p(\mathcal{O}),\ D_i u \in L^p(\mathcal{O})$ for $i = 1, ...d\}$, where $D_i = \frac{\partial}{\partial x_i}$ is taken in the sense of Schwartz distributions. We set also $H^k(\mathcal{O}) = W^{k,2}(\mathcal{O})$, $k \in \mathbb{N}$. We denote the Euclidean norm of $\mathbb{R}^d$ by $|\cdot|_d$ or $|\cdot|$, if there is no possible confusion, and by $C_b(\mathbb{R})$ and $C_b(\mathbb{R}^d, \mathbb{R}^d)$ the spaces of continuous and bounded functions from $\mathbb{R}$ to itself and, respectively, from $\mathbb{R}^d$ to $\mathbb{R}^d$. By $C^1(\mathbb{R})$ we denote the space of continuously differentiable real valued functions.

2 Existence of mild solutions for NFPE (1.1)

Consider in the space $L^1 = L^1(\mathbb{R}^d)$ the operator $A_0 : D(A_0) \subset L^1 \to L^1$, defined by

$$\begin{aligned} A_0 u &= -\Delta\beta(u) + \operatorname{div}(Db(u)u),\ \forall u \in D(A_0), \\ D(A_0) &= \{u \in L^1;\ -\Delta\beta(u) + \operatorname{div}(Db(u)u) \in L^1\}. \end{aligned} \tag{2.1}$$

Here, the differential operators Δ and div are taken in the sense of Schwartz distributions, i.e., in $\mathcal{D}'(\mathbb{R}^d)$. Obviously, the operator $(A_0, D(A_0))$ is closed on L^1.

By hypotheses (i)–(iii), we see that $\beta(u), Dub(u) \in L^1$, $\forall u \in L^1$, and so $-\Delta\beta(u), \operatorname{div}(Dub(u)) \in \mathcal{D}'(\mathbb{R}^d)$ for all $u \in L^1$.

Proposition 2.1 *Assume that hypotheses* (i)–(iv) *hold. Then,*

$$R(I + \lambda A_0) = L^1,\ \forall \lambda > 0, \tag{2.2}$$

and there is an operator $J_\lambda : L^1 \to L^1$ *such that*

$$J_{\lambda_2}(u) = J_{\lambda_1}\left(\frac{\lambda_1}{\lambda_2}\, u + \left(1 - \frac{\lambda_1}{\lambda_2}\right) J_{\lambda_2}(u)\right),\ 0 < \lambda_1,\ \lambda_2 < \infty, \tag{2.3}$$

$$J_\lambda(u) = (I + \lambda A_0)^{-1} u,\quad \forall u \in L^1,\ \lambda > 0, \tag{2.4}$$

$$|J_\lambda(u) - J_\lambda(v)|_1 \le |u - v|_1,\ \forall \lambda > 0,\ u, v \in L^1. \tag{2.5}$$

Furthermore,

$$\overline{D(A)} = L^1, \tag{2.6}$$

where "——" denotes the closure in L^1 *and* A *is the operator defined by formula* (2.9) *below. Moreover, there exists* $\lambda_0 > 0$ *such that, for all* $\lambda \in (0, \lambda_0)$,

$$\int_{\mathbb{R}^d} J_\lambda(u_0) dx = \int_{\mathbb{R}^d} u_0(x) dx,\ \forall u_0 \in L^1, \tag{2.7}$$

$$J_\lambda(u_0) \ge 0,\quad a.e.\ in\ \mathbb{R}^d\ if\ u_0 \ge 0,\ a.e.\ in\ \mathbb{R}^d. \tag{2.8}$$

The proof of Proposition 2.1 will be given in Section 3.

We note that $J_\lambda(L^1) = J_\lambda(L^1)$, $\forall \lambda > 0$. We are lead to introduce the operator $A : D(A) \subset L^1 \to L^1$,

$$Au = A_0 u,\ \forall u \in D(A) = J_\lambda(L^1),\ \forall \lambda > 0. \tag{2.9}$$

By (2.2)-(2.6), it follows that A *is* m*-accretive in* L^1. This means (see, e.g. [1], p. 97) that $|u - v + \lambda(Au - Av)|_1 \geq |u - v|_1,\ \forall u, v \in D(A),\ \lambda > 0$, and $R(I + \lambda A) = L^1$, $\forall \lambda > 0$ (equivalently, for some $\lambda > 0$).

We note that

$$(I + \lambda A)^{-1} u = J_\lambda(u),\ \forall u \in L^1,\ \lambda > 0. \tag{2.10}$$

Consider now the Cauchy problem associated with A, that is,

$$\begin{aligned} &\frac{du}{dt} + Au = 0,\ t \geq 0, \\ &u(0) = u_0. \end{aligned} \tag{2.11}$$

A continuous function $u : [0, \infty) \to L^1$ is said to be a *mild solution to equation* (2.11) if

$$u(t) = \lim_{h \to 0} u_h(t) \text{ in } L^1,\ \forall t \geq 0, \tag{2.12}$$

uniformly on compacts of $[0, \infty)$, where $u_h^0 = u_0$, and

$$u_h(t) = u_h^i,\ t \in [ih, (i+1)h),\ i = 0, 1, ..., \tag{2.13}$$

$$u_h^{i+1} + hAu_h^{i+1} = u_h^i,\ i = 0, ... \tag{2.14}$$

Since A is m-accretive, we have by the Crandall & Liggett theorem (see, e.g., [3], p. 141) the following existence result for problem (2.11).

Proposition 2.2 *Under hypotheses* (i)–(iv)*, there is a unique mild solution* u *to equation* (2.11)*. Moreover, for every* $u_0 \in \overline{D(A)} = L^1$*, one has, for all* $t \geq 0$*,*

$$u(t) = \lim_{n \to \infty} \left(I + \frac{t}{n} A \right)^{-n} u_0 \tag{2.15}$$

uniformly on bounded intervals of $[0, \infty)$ *in the strong topology in* L^1*. One also has that*

$$\int_{\mathbb{R}^d} u(t, x) dx = \int_{\mathbb{R}^d} u_0(x) dx,\ \forall t \geq 0, \tag{2.16}$$

$$u(t, x) \geq 0,\ \textit{a.e. on } (0, \infty) \times \mathbb{R}^d \textit{ if } u_0 \geq 0,\ \textit{a.e. in } \mathbb{R}^d. \tag{2.17}$$

The function u will be called the *mild solution* to NFPE (1.1).

In particular, it follows by (2.16), (2.17) that, for each $t \geq 0$, $u(t,\cdot)$ is a probability density if so is u_0.

We note that (2.16)–(2.17) follow by (2.7)-(2.8) and (2.15).

The map $t \to S(t)u_0$ is a continuous semigroup of contractions on L^1, that is,

$$S(t)u_0 = u(t) = \lim_{n\to\infty} \left(I + \frac{t}{n}\,A\right)^{-n} u_0,\ \forall t \geq 0, \tag{2.18}$$

$$S(t+s)u_0 = S(t)S(s)u_0,\ \forall t,s \geq 0,\ u_0 \in L^1, \tag{2.19}$$

$$\lim_{t\to 0} S(t)u_0 = u_0 \text{ in } L^1, \tag{2.20}$$

$$|S(t)u_0 - S(t)\bar{u}_0|_1 \leq |u_0 - \bar{u}_0|_1,\ \forall t \geq 0,\ u_0, \bar{u}_0 \in L^1. \tag{2.21}$$

If

$$\mathcal{P} = \left\{ u \in L^1;\ u \geq 0,\ \int_{\mathbb{R}^d} u(x)dx = 1 \right\}, \tag{2.22}$$

we see by (2.16)–(2.18) that

$$S(t)(\mathcal{P}) \subset \mathcal{P},\ \forall t \geq 0. \tag{2.23}$$

Since, for every i and h the function $u_h^{i+1} \in D(A)$ is a solution to (2.14) in the sense of distributions, i.e. in the space $\mathcal{D}'(\mathbb{R}^d)$, it follows also that the mild solution u to (2.11) is a solution to NFPE (1.1) in the sense of Schwartz distributions on $(0,\infty)\times\mathbb{R}^d$, that is,

$$\int_0^\infty \int_{\mathbb{R}^d} (u\varphi_t + \beta(u)\Delta\varphi + Db(u)u\cdot\nabla\varphi)dx\,dt = 0,\ \forall \varphi \in \mathcal{D}((0,\infty)\times\mathbb{R}^d), \tag{2.24}$$

where $\mathcal{D}((0,\infty)\times\mathbb{R}^d)$ is the space of infinitely differentiable functions on $(0,\infty)\times\mathbb{R}^d$ with compact support.

It should be emphasized, however, that the solution u to NFPE (1.1) exists and is unique in the class of mild solutions corresponding to the operator A and not in the space of Schwartz distributions on $(0,\infty)\times\mathbb{R}^d$.

We consider the following subspace of L^1

$$\mathcal{M} \;=\; \left\{ u \in L^1; \int_{\mathbb{R}^d} \Phi(x)|u(x)|dx < \infty \right\}$$

with the norm

$$\|u\| = \int_{\mathbb{R}^d} \Phi(x)|u(x)|dx, \ \forall u \in \mathcal{M}. \tag{2.25}$$

It turns out that the semigroup $S(t)$ leaves invariant $\mathcal{M}$. More precisely, we prove in Section 3:

Proposition 2.3 *Assume that* (i)–(iv) *hold. Then*

$$\|S(t)u_0\| \le \|u_0\| + \rho t|u_0|_1, \ \ \forall u_0 \in \mathcal{M}, \tag{2.26}$$

where $\rho = (m+1)|\Delta\Phi|_\infty\gamma_1 + |b|_\infty(1+m)^2|D|_\infty^2$.

Remark 2.4 Propositions 2.1–2.3 remain valid if, in addition to hypotheses (i)–(iv), we assume, instead of (iv),

(iv)′ $D_0 = \sup\limits_{x\in\mathbb{R}^d} |D(x)\cdot x| < \infty$,

but we have to replace $\mathcal{M}$ by

$$\mathcal{M}_2 = \left\{ u \in L^1 : \int_{\mathbb{R}^d} |x|^2|u(x)|dx < \infty \right\}$$

with the norm

$$\|u\|_2 = \int_{\mathbb{R}^d} |x|^2|u(x)|dx,$$

and we have to replace ρ in Proposition 2.3 by $\widetilde{\rho} := 2(d\gamma_1 + D_0|b|_\infty)$ (see Remark 3.3 below). The assumption (iv), in particular that D is the negative of the gradient of a positive function, becomes, however, important for Sections 4–6 below, i.e., to prove the H-Theorem.

3 Proof of Propositions 2.1 and 2.3

As mentioned earlier, one can derive Proposition 2.1 from similar results established in [5], [6]. However, for later use we shall prove it by a constructive regularization technique already developed in the above works. Namely, we define, for each $\varepsilon > 0$, the operator $(A_0)_\varepsilon : D((A_0)_\varepsilon) \subset L^1 \to L^1$ defined by

$$(A_0)_\varepsilon u = -\Delta(\beta(u)) + \varepsilon\beta(u) + \mathrm{div}(D_\varepsilon b_\varepsilon^*(u)), \tag{3.1}$$

$$D((A_0)_\varepsilon) = \{u \in L^1, \ -\Delta(\beta(u)) + \varepsilon\beta(u) + \mathrm{div}(D_\varepsilon b_\varepsilon^*(u)) \in L^1\}. \tag{3.2}$$

Here Δ and div are taken in the sense of Schwartz distributions and

$$b_\varepsilon \equiv b * \rho_\varepsilon, \quad b^*_\varepsilon(r) \equiv \frac{b_\varepsilon(r)r}{1+\varepsilon|r|}, \ r \in \mathbb{R}, \tag{3.3}$$

where $\rho_\varepsilon(r) \equiv \frac{1}{\varepsilon}\ \rho\left(\frac{r}{\varepsilon}\right)$, $\rho \in C^\infty_0(\mathbb{R})$, $\rho \geq 0$, is a standard mollifier.

Moreover,

$$D_\varepsilon = -\nabla\Phi_\varepsilon, \quad \Phi_\varepsilon(x) \equiv \frac{\Phi(x)}{(1+\varepsilon\Phi(x))^m}.$$

Then $\Phi_\varepsilon \in L^2$, since $m \geq 2$, and

$$D_\varepsilon = D(1+\varepsilon\Phi)^{-m} - m\varepsilon\Phi D(1+\varepsilon\Phi)^{-(m+1)} \tag{3.4}$$

and, therefore, by (iv)

$$\begin{gathered} D_\varepsilon \in C_b(\mathbb{R}^d) \cap W^{1,\infty}(\mathbb{R}^d) \cap L^1(\mathbb{R}^d;\mathbb{R}^d) \\ |D_\varepsilon|_\infty \leq (1+m)\,|D|_\infty, \ \lim_{\varepsilon\to 0} D_\varepsilon(x) = D(x), \ \forall x \in \mathbb{R}^d, \\ \varepsilon^m|D_\varepsilon| \leq (1+m)|D|_\infty \Phi^{-m}, \ \forall \varepsilon > 0. \end{gathered} \tag{3.5}$$

We also note that $b^*_\varepsilon, b_\varepsilon$ are bounded and Lipschitz and that, for $\varepsilon \to 0$,

$$b^*_\varepsilon(r) \to b(r)r \ \text{ uniformly on compacts.} \tag{3.6}$$

Obviously, the operator $((A_0)_\varepsilon, D((A_0)_\varepsilon))$ is closed on L^1.

Lemma 3.1 *Under hypotheses* (i)–(iii)*, the operator* $(A_0)_\varepsilon$ *satisfies relations* (2.2)-(2.8)*. Assume further that* (iv) *holds. Then, there is* $\lambda_0 > 0$ *independent of* $f \in L^1$ *such that, for all* $\lambda \in (0, \lambda_0)$*,*

$$\lim_{\varepsilon\to 0}(I+\lambda(A_0)_\varepsilon)^{-1}f = J_\lambda f \quad \textit{in } L^1, \ \forall f \in L^1. \tag{3.7}$$

As in the case of the operator A, we define

$$A_\varepsilon u = (A_0)_\varepsilon u, \ \forall u \in D(A_\varepsilon) = (I+(A_0)_\varepsilon)^{-1}L^1. \tag{3.8}$$

Then, *Lemma* 3.1 *implies that* A_ε *is m-accretive in* L^1 *and* (2.2)-(2.8) *hold.* Moreover, by (3.7) it follows that, if (iv) holds, then

$$\lim_{\varepsilon\to 0}(I+\lambda A_\varepsilon)^{-1}f = J_\lambda(f) \ \text{ in } L^1, \ \forall f \in L^1, \ \text{for } \lambda \in (0,\lambda_0). \tag{3.9}$$

Proof of Lemma 3.1. For the m-accretivity, it should be proved that, for all $\lambda > 0$, $\varepsilon > 0$, we have

$$R(I + \lambda(A_0)_\varepsilon) = L^1, \tag{3.10}$$

$$|(I + \lambda(A_0)_\varepsilon)^{-1} f_1 - (I + \lambda A_\varepsilon)^{-1} f_2|_1 \le |f_1 - f_2|_1, \tag{3.11}$$

for all $f_1, f_2 \in L^1$. To this end, we fix first $f \in L^2 \cap L^1$ and consider the equation $u + \lambda A_\varepsilon u = f$, that is,

$$u - \lambda\Delta(\beta(u)) + \varepsilon\lambda\beta(u) + \lambda \operatorname{div}(D_\varepsilon b^*_\varepsilon(u)) = f \text{ in } \mathcal{D}'(\mathbb{R}^d). \tag{3.12}$$

To solve equation (3.12), we consider the equation

$$(\varepsilon I - \Delta)^{-1} u + \lambda\beta(u) + \lambda(\varepsilon I - \Delta)^{-1} \operatorname{div}(D_\varepsilon b^*_\varepsilon(u)) = (\varepsilon I - \Delta)^{-1} f \text{ in } L^2. \tag{3.13}$$

Clearly, a solution of (3.13) satisfies (3.12) in L^2. We set

$$\begin{aligned} F_\varepsilon(u) &= (\varepsilon I - \Delta)^{-1} u,\ G(u) = \lambda\beta(u),\ u \in L^2, \\ G_\varepsilon(u) &= \lambda(\varepsilon I - \Delta)^{-1}(\operatorname{div}(D_\varepsilon b^*_\varepsilon(u))),\ u \in L^2, \end{aligned} \tag{3.14}$$

and note that F_ε and G are accretive and continuous in L^2.

We also have by assumptions (ii)–(iii) that G_ε is continuous in L^2 and

$$\begin{aligned} &\int_{\mathbb{R}^d} (G_\varepsilon(u) - G_\varepsilon(\bar{u}))(u - \bar{u}) dx \\ &\quad = -\lambda \int_{\mathbb{R}^d} D_\varepsilon(b^*_\varepsilon(u) - b^*_\varepsilon(\bar{u})) \cdot \nabla(\varepsilon I - \Delta)^{-1}(u - \bar{u})) dx \\ &\quad \ge -C_\varepsilon\lambda |u - \bar{u}|_2 |\nabla(\varepsilon I - \Delta)^{-1}(u - \bar{u})|_2,\ \forall u, \bar{u} \in L^2(\mathbb{R}^d), \end{aligned} \tag{3.15}$$

for some $C_\varepsilon > 0$. Moreover, we have

$$\int_{\mathbb{R}^d} (\varepsilon I - \Delta)^{-1} u u\, dx = \varepsilon |(\varepsilon I - \Delta)^{-1} u|_2^2 + |\nabla(\varepsilon I - \Delta)^{-1} u|_2^2,\ \forall u \in L^2. \tag{3.16}$$

By (3.13)–(3.16), we see that, for $u^* = u - \bar{u}$, we have

$$\begin{aligned} &(F_\varepsilon(u^*) + G_\varepsilon(u) - G_\varepsilon(\bar{u}) + G(u) - G(\bar{u}), u^*)_2 \\ &\quad \ge \lambda\gamma |u^*|_2^2 + |\nabla(\varepsilon I - \Delta)^{-1} u^*|_2^2 + \varepsilon |(\varepsilon I - \Delta)^{-1} u^*|_2^2 \\ &\quad -C_\varepsilon \lambda |u^*|_2 |\nabla(\varepsilon I - \Delta)^{-1} u^*|_2^2. \end{aligned}$$

This implies that $F_\varepsilon+G_\varepsilon+G$ is accretive and coercive on L^2 for $\lambda<\lambda_\varepsilon$, where λ_ε is sufficiently small. Since this operator is continuous and accretive, it follows that it is m-accretive and, therefore, surjective (because it is coercive). Hence, for each $f\in L^2\cap L^1$ and $\lambda<\lambda_\varepsilon$, equation (3.13) has a unique solution $u_\varepsilon\in L^2$.

Since $u_\varepsilon\in L^2$, $b^*_\varepsilon(r)\le C_\varepsilon|r|$, $r\in\mathbb{R}$, and $D_\varepsilon\in L^\infty$, by (3.12) we see that $\beta(u_\varepsilon)\in H^1(\mathbb{R}^d)$, whence by (i) we have

$$u_\varepsilon\in H^1(\mathbb{R}^d). \tag{3.17}$$

Multiplying (3.12) by u_ε and $\beta(u_\varepsilon)$, respectively, and integrating over $\mathbb{R}^d$ we get after some calculation that, for $\lambda<\lambda_0$ with λ_0 small enough,

$$|u_\varepsilon|_2^2+\lambda|\nabla\beta(u_\varepsilon)|_2^2+\lambda|\nabla u_\varepsilon|_2^2+\varepsilon\lambda|\beta(u_\varepsilon)|_2^2\le C_{\lambda_0}|f|_2^2, \tag{3.18}$$

where C_{λ_0} is independent of ε.

We denote by $u_\varepsilon(f)\in H^1(\mathbb{R}^d)$ the solution to (3.13) for $f\in L^2\cap L^1$ and prove that

$$|u_\varepsilon(f_1)-u_\varepsilon(f_2)|_1\le|f_1-f_2|_1,\ \forall f_1,f_2\in L^1\cap L^2. \tag{3.19}$$

Here is the argument. We set $u=u_\varepsilon(f_1)-u_\varepsilon(f_2)$, $f=f_1-f_2$. By (3.12), we have, for $u_i=u_\varepsilon(f_i)$, $i=1,2$,

$$\begin{aligned}u-\lambda\Delta(\beta(u_1)-\beta(u_2))+\varepsilon\lambda(\beta(u_1)-\beta(u_2))\\ +\lambda\,\mathrm{div}(D_\varepsilon(b^*_\varepsilon(u_1)-b^*_\varepsilon(u_2)))=f\ \text{ in } L^2.\end{aligned} \tag{3.20}$$

Proceeding as in [6] (see, also, [15]), we consider the Lipschitzian function $\mathcal{X}_\delta:\mathbb{R}\to\mathbb{R}$,

$$\mathcal{X}_\delta(r)=\begin{cases}1 & \text{for } r\ge\delta,\\ \dfrac{r}{\delta} & \text{for } |r|<\delta,\\ -1 & \text{for } r<-\delta,\end{cases} \tag{3.21}$$

where $\delta>0$. We set

$$F_\varepsilon=\lambda\nabla(\beta(u_1)-\beta(u_2))-\lambda D_\varepsilon(b^*_\varepsilon(u_1)-b^*_\varepsilon(u_2))$$

and rewrite (3.20) as

$$u=\mathrm{div}\ F_\varepsilon-\varepsilon\lambda(\beta(u_1)-\beta(u_2))+f. \tag{3.22}$$

By (3.17), it follows that $F_\varepsilon \in L^2(\mathbb{R}^d)$. We set $\Lambda_\delta = \mathcal{X}_\delta(\beta(u_1) - \beta(u_2))$.

Since $\Lambda_\delta \in H^1(\mathbb{R}^d)$, it follows that $\Lambda_\delta \mathrm{div}\, F_\varepsilon \in L^1$ and so, by (3.22), we have

$$\begin{aligned}\int_{\mathbb{R}^d} u\Lambda_\delta dx = &-\int_{\mathbb{R}^d} F_\varepsilon \cdot \nabla\Lambda_\delta dx\\ &-\varepsilon\lambda\int_{\mathbb{R}^d}(\beta(u_1)-\beta(u_2))\Lambda_\delta dx + \int_{\mathbb{R}^d} f\Lambda_\delta dx\\ &= -\int_{\mathbb{R}^d}(F_\varepsilon\cdot\nabla(\beta(u_1)-\beta(u_2))\mathcal{X}'_\delta(\beta(u_1)-\beta(u_2))dx\\ &-\varepsilon\lambda\int_{\mathbb{R}^d}(\beta(u_1)-\beta(u_2))\mathcal{X}_\delta(\beta(u_1)-\beta(u_2))dx + \int_{\mathbb{R}^d} f\Lambda_\delta dx.\end{aligned} \tag{3.23}$$

We set

$$\begin{aligned}I^1_\delta &= \int_{\mathbb{R}^d} D_\varepsilon(b^*_\varepsilon(u_1)-b^*_\varepsilon(u_2))\cdot\nabla\Lambda_\delta dx\\ &= \int_{\mathbb{R}^d} D_\varepsilon(b^*_\varepsilon(u_1)-b^*_\varepsilon(u_2))\cdot\nabla(\beta(u_1)-\beta(u_2))\mathcal{X}'_\delta(\beta(u_1)-\beta(u_2))dx\\ &= \frac{1}{\delta}\int_{[|\beta(u_1)-\beta(u_2)|\le\delta]} D_\varepsilon(b^*_\varepsilon(u_1)-b^*_\varepsilon(u_2))\cdot\nabla(\beta(u_1)-\beta(u_2))dx.\end{aligned} \tag{3.24}$$

Since $|D_\varepsilon|_d \in L^\infty \cap L^2$ and by assumption (i)

$$|b^*_\varepsilon(u_1)-b^*_\varepsilon(u_2)| \le \mathrm{Lip}(b^*_\varepsilon)|u_1-u_2| \le \gamma\ \mathrm{Lip}(b^*_\varepsilon)|\beta(u_1)-\beta(u_2)|,$$

it follows that

$$\begin{aligned}&\lim_{\delta\to0}\frac{1}{\delta}\int_{[|(\beta(u_1)-\beta(u_2))|\le\delta]}|D_\varepsilon(b^*_\varepsilon(u_1)-b^*_\varepsilon(u_2))\cdot\nabla(\beta(u_1)-\beta(u_2))|dx\\ &\le \gamma\ \mathrm{Lip}(b^*_\varepsilon)|D_\varepsilon|_2\lim_{\delta\to0}\left(\int_{[|\beta(u_1)-\beta(u_2)|\le\delta]}|\nabla(\beta(u_1)-\beta(u_2))|^2dx\right)^{\frac12} = 0.\end{aligned}$$

This yields

$$\lim_{\delta\to0} I^1_\delta = 0, \tag{3.25}$$

because $\nabla(\beta(u_1)-\beta(u_2))(x) = 0$, a.e. on $[x\in\mathbb{R}^d; \beta(u_1(x))-\beta(u_2(x))=0]$. On the other hand, since $\mathcal{X}'_\delta \ge 0$, we have

$$\int_{\mathbb{R}^d}\nabla(\beta(u_1)-\beta(u_2))\cdot\nabla(\beta(u_1)-\beta(u_2))\mathcal{X}'_\delta(\beta(u_1)-\beta(u_2))\,dx \ge 0. \tag{3.26}$$

By (3.23)–(3.26), since $|\Lambda_\delta| \le 1$, we get

$$\lim_{\delta\to 0}\int_{\mathbb{R}^d} u\mathcal{X}_\delta(\beta(u_1)-\beta(u_2))dx \le \int_{\mathbb{R}^d}|f|\,dx$$

and, since $u\mathcal{X}_\delta(\beta(u_1)-\beta(u_2)) \ge 0$ and $\mathcal{X}_\delta \to$ sign as $\delta \to 0$, by Fatou's lemma this yields

$$|u|_1 \le |f|_1, \tag{3.27}$$

as claimed.

Next, for f arbitrary in L^1, consider a sequence $\{f_n\} \subset L^2$ such that $f_n \to f$ strongly in L^1. Let $\{u^n_\varepsilon\} \subset L^1 \cap L^2$ be the corresponding solutions to (3.13) for $0 < \lambda < \lambda_\varepsilon$. We have, for all $m, n \in \mathbb{N}$,

$$u^n_\varepsilon - u^m_\varepsilon + \lambda((A_0)_\varepsilon u^n_\varepsilon - (A_0)_\varepsilon u^m_\varepsilon) = f_n - f_m. \tag{3.28}$$

Taking into account (3.27)), we obtain by the above equation that

$$|u^n_\varepsilon - u^m_\varepsilon|_1 \le |f_n - f_m|_1,\ \forall n, m \in \mathbb{N}.$$

Hence, for $n \to \infty$, we have

$$u^n_\varepsilon \to u_\varepsilon(f) \text{ in } L^1.$$

Now, (3.28) implies that $(A_0)_\varepsilon u^n_\varepsilon \to v$ in L^1. Since $((A_0)_\varepsilon, D((A_0)_\varepsilon))$ is closed on L^1, we conclude that $u_\varepsilon(f) \in D((A_0)_\varepsilon)$ and that

$$u_\varepsilon(f) + \lambda(A_0)_\varepsilon u_\varepsilon(f) = f, \tag{3.29}$$

which proves (3.10) for $\lambda < \lambda_\varepsilon$. Moreover, by (3.27), we have

$$|u_\varepsilon(f_1) - u_\varepsilon(f_2)|_1 \le |f_1 - f_2|_1,\ \forall f_1, f_2 \in L^1, \tag{3.30}$$

which proves (3.11) for $\lambda < \lambda_\varepsilon$. By Proposition 3.3 in [3], p. 99, it follows that

$$R(1 + \lambda(A_0)_\varepsilon) = L^1,\ \forall \lambda > 0,$$

and, therefore, (3.10)-(3.11) hold for all $\lambda > 0$ if $f \in L^1$. We also have

$$\int_{\mathbb{R}^d}(I+\lambda(A_0)_\varepsilon)^{-1}f\,dx = \int_{\mathbb{R}^d} f\,dx - \varepsilon\lambda\int_{\mathbb{R}^d}\beta((I+\lambda(A_0)_\varepsilon)^{-1}f)dx, \quad \forall f \in L^1,\ \lambda > 0, \tag{3.31}$$

and there exists $\widetilde{\lambda}_0$ independent of ε such that, for all $\lambda \in (0, \widetilde{\lambda}_0)$,

$$(I + \lambda(A_0)_\varepsilon)^{-1} f \geq 0, \text{ a.e. in } \mathbb{R}^d \text{ if } f \geq 0, \text{ a.e. in } \mathbb{R}^d. \tag{3.32}$$

(The latter follows by multiplying (3.12), where $u = u_\varepsilon$, with sign u_ε^- and integrating over $\mathbb{R}^d$.)

Next, we show (3.7). Fix $\lambda < \lambda_0 = \min(\lambda_0, \widetilde{\lambda}_0)$ and let $f \in L^1 \cap L^2$. If $u_\varepsilon = u_\varepsilon(f)$, by (3.18), it follows that $\{u_\varepsilon\}$ is bounded in $H^1(\mathbb{R}^d)$ and $\{\beta(u_\varepsilon)\}$ is bounded in $H^1(\mathbb{R}^d)$. Clearly, $u_\varepsilon(f) = 0$ if $f \equiv 0$, hence (3.30) implies that $\{u_\varepsilon\}$ is bounded in L^1. Hence, along a subsequence, again denoted $\{\varepsilon\} \to 0$, we have

$$\begin{aligned} u_\varepsilon &\longrightarrow u && \text{weakly in } H^1(\mathbb{R}^d), \text{ strongly in } L^2_{\text{loc}}(\mathbb{R}^d), \\ \beta(u_\varepsilon) &\longrightarrow \beta(u) && \text{weakly in } H^1(\mathbb{R}^d) \text{ and strongly in } L^2_{\text{loc}}(\mathbb{R}^d), \\ \Delta\beta(u_\varepsilon) &\longrightarrow \Delta\beta(u) && \text{weakly in } H^{-1}(\mathbb{R}^d), \end{aligned} \tag{3.33}$$

and, by hypotheses (ii) and (3.6),

$$b^*_\varepsilon(u_\varepsilon) \quad \longrightarrow \quad b(u)u \quad \text{strongly in } L^2_{\text{loc}}(\mathbb{R}^d). \tag{3.34}$$

This yields

$$D_\varepsilon b^*_\varepsilon(u_\varepsilon) \to Db(u)u \text{ strongly in } L^2_{\text{loc}}(\mathbb{R}^d). \tag{3.35}$$

Passing to the limit in (3.12), we obtain

$$u - \lambda\Delta\beta(u) + \lambda \operatorname{div}(Db(u)u) = f \text{ in } \mathcal{D}'(\mathbb{R}^d), \tag{3.36}$$

where $u = u(\lambda, f) \in H^1(\mathbb{R}^d)$. By (3.30) and (3.33), it follows via Fatou's lemma that

$$|u(\lambda, f_1) - u(\lambda, f_2)|_1 \leq |f_2 - f_2|_1, \ \forall f_1, f_2 \in L^2 \cap L^1, \tag{3.37}$$

and hence (since $u(\lambda, f) = 0$ if $f \equiv 0$) $u_1(\lambda, f), u_2(\lambda, f) \in L^1 \cap L^2$, if $f \in L^1 \cap L^2$.

In particular, $u(\lambda, f) \in D(A_0)$ and

$$u(\lambda, f) + \lambda A_0 u(\lambda, f) = f, \tag{3.38}$$

and so (2.2) follows. We define $J_\lambda : L^1 \to L^1$ as $J_\lambda(f) = u(\lambda, f)$ and by (3.30), it follows (2.5). Clearly, by (3.33),

$$u_\varepsilon \to u \text{ in } L^1_{\text{loc}}, \tag{3.39}$$

for $0 < \lambda < \lambda_0$. To prove that (3.39) in fact holds in L^1, we shall prove first the following lemma, which has an intrinsic interest.

Lemma 3.2 *Assume that hypotheses* (i)–(iv) *hold and let* $u_0 \in \mathcal{M}$. *Then, for all* $\lambda \in (0, \lambda_0)$,

$$\|(I + \lambda(A_0)_\varepsilon)^{-1}u_0\| \le \|u_0\| + \rho_\varepsilon \lambda |u_0|_1, \tag{3.40}$$

where $\rho_\varepsilon = \gamma_1(m+1)|\Delta\Phi|_\infty + \gamma_1 m(m+3)\varepsilon|D|^2_\infty + |b|_\infty(1+m)^2|D|^2_\infty$.

Proof. Let $u_0 \in \mathcal{M}$. If we multiply equation (3.29) by $\varphi_\nu \mathcal{X}_\delta(\beta(u_\varepsilon))$, where $u_\varepsilon = (I + \lambda(A_0)_\varepsilon)^{-1}u_0$, $\varphi_\nu(x) = \Phi_\varepsilon(x)\exp(-\nu\Phi_\varepsilon(x))$ and integrate over $\mathbb{R}^d$, we get, since $\mathcal{X}'_\delta \ge 0$,

$$\begin{aligned}
\int_{\mathbb{R}^d} u_\varepsilon \mathcal{X}_\delta(\beta(u_\varepsilon))\varphi_\nu\,dx &\le -\lambda \int_{\mathbb{R}^d} \nabla\beta(u_\varepsilon)\cdot\nabla(\mathcal{X}_\delta(\beta(u_\varepsilon))\varphi_\nu)dx \\
&\quad +\lambda \int_{\mathbb{R}^d} D_\varepsilon b^*_\varepsilon(u_\varepsilon)\cdot\nabla(\mathcal{X}_\delta(\beta(u_\varepsilon))\varphi_\nu)dx + \int_{\mathbb{R}^d}|u_0|\varphi_\nu dx \\
&\le -\lambda \int_{\mathbb{R}^d} \nabla\beta(u_\varepsilon)\cdot\nabla\varphi_\nu \mathcal{X}_\delta(\beta(u_\varepsilon))dx \\
&\quad +\lambda \int_{\mathbb{R}^d} D_\varepsilon b^*_\varepsilon(u_\varepsilon)\cdot\nabla\beta(u_\varepsilon)\mathcal{X}'_\delta(\beta(u_\varepsilon))\varphi_\nu dx \\
&\quad +\lambda \int_{\mathbb{R}^d} (D_\varepsilon\cdot\nabla\varphi_\nu) b^*_\varepsilon(u_\varepsilon)\mathcal{X}_\delta(\beta(u_\varepsilon))dx + \int_{\mathbb{R}^d}|u_0|\varphi_\nu dx.
\end{aligned} \tag{3.41}$$

Letting $\delta \to 0$, we get as above

$$\begin{aligned}
\int_{\mathbb{R}^d} |u_\varepsilon|\varphi_\nu dx &\le -\lambda \int_{\mathbb{R}^d} \nabla|\beta(u_\varepsilon)|\cdot\nabla\varphi_\nu dx \\
&\quad +\overline{\lim_{\delta\to 0}}\frac{\lambda}{\delta}\int_{[|\beta(u_\varepsilon)|\le\delta]} |D_\varepsilon|\,|b^*_\varepsilon(u_\varepsilon)|\,|\nabla\beta(u_\varepsilon)|\varphi_\nu dx \\
&\quad +\lambda \int_{\mathbb{R}^d} \operatorname{sign} u_\varepsilon b^*_\varepsilon(u_\varepsilon) D_\varepsilon\cdot\nabla\varphi_\nu\,dx + \int_{\mathbb{R}^d}|u_0|\varphi_\nu dx \\
&\le \lambda \int_{\mathbb{R}^d} (|\beta(u_\varepsilon)|\Delta\varphi_\nu + |b^*_\varepsilon(u)|\,|\nabla\Phi_\varepsilon\cdot\nabla\varphi_\nu|)dx + \int_{\mathbb{R}^d}|u_0|\varphi_\nu dx,
\end{aligned} \tag{3.42}$$

since $D_\varepsilon = -\nabla\Phi_\varepsilon$. We have

$$\nabla\varphi_\nu(x) = (\nabla\Phi_\varepsilon - \nu\Phi_\varepsilon\nabla\Phi_\varepsilon)\exp(-\nu\Phi_\varepsilon), \tag{3.43}$$

$$\begin{aligned}
\Delta\varphi_\nu(x) &= (\Delta\Phi_\varepsilon - \nu|\nabla\Phi_\varepsilon|^2 - \nu\Phi_\varepsilon\Delta\Phi_\varepsilon + \nu^2\Phi_\varepsilon|\nabla\Phi_\varepsilon|^2 \\
&\quad -\nu|\nabla\Phi_\varepsilon|^2)\exp(-\nu\Phi_\varepsilon).
\end{aligned} \tag{3.44}$$

Then, letting $\nu \to 0$, we get by (3.42) that

$$\|u_\varepsilon\| \le \|u_0\| + \lambda\gamma_1 |\Delta\Phi_\varepsilon|_\infty |u_0|_1, \ \forall \varepsilon > 0.$$

On the other hand,

$$\begin{aligned}\Delta\Phi_\varepsilon = -\text{div } D_\varepsilon &= (1 - m\varepsilon\Phi(1+\varepsilon\Phi)^{-1})(1+\varepsilon\Phi)^{-m}\Delta\Phi \\ &+ m\varepsilon((m+1)\varepsilon\Phi(1+\varepsilon\Phi)^{-1} - 2)(1+\varepsilon\Phi)^{-(m+1)}|D|^2.\end{aligned} \tag{3.45}$$

Therefore,

$$|\Delta\Phi_\varepsilon|_\infty \le (m+1)|\Delta\Phi|_\infty + m(m+3)\varepsilon|D|^2_\infty,$$

and this, together with (3.5), yields (3.40), as claimed.

Remark 3.3 If, as in Remark 2.4, we replace (iv), $\mathcal{M}$, $\|\cdot\|$ and ρ by (iv)$'$ (see Remark 2.4), $\mathcal{M}_2$, $\|\cdot\|_2$ and $\widetilde{\rho}$, respectively, we can prove a complete analogue of Lemma 3.2 by the same arguments. One only has to replace φ_ν by the function $\widetilde{\varphi}_\nu(x) = |x|^2 e^{-\nu|x|^2}$ in the above proof. Once one has this analogue of Lemma 3.2, the proofs below can easily be adjusted to this case.

Proof of (3.7)**.** By (3.40) and hypothesis (iv), it follows that, if $f \in \mathcal{M}$, then we have, for all $\lambda \in (0, \lambda_0)$ and $\varepsilon \in (0,1)$, $N > 0$,

$$\int_{\{\Phi \ge N\}} |(I + \lambda(A_0)_\varepsilon)^{-1} f| dx \le \frac{1}{N} \|(I + \lambda(A_0)_\varepsilon)^{-1} f\| \le \frac{1}{N}(\|f\| + \rho_1 \lambda |f|_1).$$

Recalling (3.39) and that $\{\Phi \le N\}$ is compact, the latter implies that, if $f \in \mathcal{M} \cap L^2$, then

$$\lim_{\varepsilon \to 0} |u_\varepsilon - u|_1 = 0,$$

i.e.,

$$\lim_{\varepsilon \to 0} (I + \lambda(A_0)_\varepsilon)^{-1} f = (I + \lambda A_0)^{-1} f \text{ in } L^1, \ \forall f \in \mathcal{M} \cap L^2. \tag{3.46}$$

Since $L^2 \cap \mathcal{M}$ is dense in L^1 and $(I + \lambda A_\varepsilon)^{-1}$, $\varepsilon > 0$, are equicontinuous, (3.7) follows.

Proof of Proposition 2.1 (continued). Fix $\lambda \in (0, \lambda_0)$ and let $f \in L^1$. Let $\{f_n\} \in L^2 \cap L^1$ be such that $f_n \to f$ in L^1. If $u_n \in H^1(\mathbb{R}^d)$ is the corresponding solution to (3.38), by (3.37) we have

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

and, therefore, $u_n \to u$ strongly in L^1 as $n \to \infty$. Since, by (3.38), $A_0 u_n \to \frac{1}{\lambda}(f-u)$ and because A_0 is closed in L^1, we infer that u is a solution to (3.38). Hence, for $\lambda \in (0,\lambda_0)$, $R(I+\lambda A_0) = L^1$, and, by (3.37), formula (2.5) follows.

Again by Proposition 3.3. in [3], p. 99, (2.2) and (2.5) follow for all $\lambda > 0$.

Proof of (2.6)**.** Let $\overline{Y}$ be the L^1-closure of $Y = \{u \in D(A_0) \cap H^1 \cap L^\infty; \Delta\beta(u) \in L^2\}$. Then (2.7) follows from the following two claims:

Claim 1. $\overline{Y} = L^1$. It suffices to prove that $C_0^\infty(\mathbb{R}^d)$ is contained in $\overline{Y}$. We fix $f \in C_0^\infty(\mathbb{R}^d)$ and consider the equation

$$u_\varepsilon - \varepsilon\Delta\beta(u_\varepsilon) = f \ \text{ in } \mathcal{D}'(\mathbb{R}^d), \tag{3.47}$$

which, as seen earlier, has for each $\varepsilon > 0$ a unique solution $u_\varepsilon \in H^1(\mathbb{R}^d)$ with $\beta(u_\varepsilon) \in H^2(\mathbb{R}^d)$, satisfying

$$\|u_\varepsilon\|_{H^1}^2 + \|\beta(u_\varepsilon)\|_{H^1}^2 + \varepsilon|\Delta\beta(u_\varepsilon)|_2^2 \le C\|f\|_{H^1},$$

where C is independent of ε. This implies that $u_\varepsilon \to f$ in $L^2_{\rm loc}$ as $\varepsilon \to 0$ (along a subsequence). Since, as it can be seen from the proof of (3.40) in Lemma 3.2,

$$\|u_\varepsilon\| \le C(\|f\| + |f|_1), \ \forall \varepsilon > 0,$$

it follows that $u_\varepsilon \to f$ in L^1 as $\varepsilon \to 0$. On the other hand, by (3.47), we see that $|u_\varepsilon|_\infty \le |f|_\infty$. (This follows in a standard way by multiplying (3.47) with $\text{sign}(u_\varepsilon - |f|_\infty)^+$ and $\text{sign}(u_\varepsilon + |f|_\infty)^-$, respectively, and integrating over $\mathbb{R}^d$.) Hence $u_\varepsilon \in L^\infty$. Let us prove that $\nabla u_\varepsilon \in (L^1)^d$.

We set $v_\varepsilon = \beta(u_\varepsilon)$ and rewrite (3.47) as

$$\beta^{-1}(v_\varepsilon) - \varepsilon\Delta v_\varepsilon = f \ \text{ in } \mathcal{D}'(\mathbb{R}^d).$$

If $w_i = \frac{\partial v_\varepsilon}{\partial x_i}$, we get

$$(\beta^{-1})'(v_\varepsilon)w_i - \varepsilon\Delta w_i = \frac{\partial f}{\partial x_i}, \ i = 1,...,d.$$

If multiply by sign w_i and integrate over $\mathbb{R}^d$, we get

$$\int_{\mathbb{R}^d} (\beta^{-1})'(v_\varepsilon)|w_i|dx \le \int_{\mathbb{R}^d} \left|\frac{\partial f}{\partial x_i}\right| dx, \ \forall i = 1,...,d.$$

Since $(\beta^{-1})'(v_\varepsilon) \geq \frac{1}{\gamma}$, we get $w_i \in L^1$, as claimed.

Now, we see that

$$A_0 u_\varepsilon = -\Delta\beta(u_\varepsilon) + \operatorname{div}(Db(u_\varepsilon)u_\varepsilon) \in L^1,\ \forall \varepsilon > 0,$$

because $-\Delta\beta(u_\varepsilon) = \varepsilon^{-1}(f - u_\varepsilon) \in L^1 \cap L^\infty$ and

$$\operatorname{div}(Db(u_\varepsilon)u_\varepsilon) = (\operatorname{div} D)b(u_\varepsilon)u_\varepsilon + (D \cdot \nabla u_\varepsilon)(b'(u_\varepsilon)u_\varepsilon + b(u_\varepsilon)) \in L^1,$$

because $u_\varepsilon \in H^1(\mathbb{R}^d) \cap L^\infty \cap W^{1,1}(\mathbb{R}^d)$ and while, by hypotheses (ii)–(iii), $\operatorname{div} D \in L^\infty$, $D \in L^\infty(\mathbb{R}^d, \mathbb{R}^d)$ and $b \in C_b \cap C^1$. This means that $u_\varepsilon \in Y$. Hence, $C_0^\infty \subset \overline{Y}$, as claimed.

Claim 2. $Y \subset D(A)$. We first note that (as is easy to see), for small enough $\lambda > 0$, the map $I + \lambda A_0 : D(A_0) \cap H^1 \to L^1$ is injective.

Now, let $u \in Y$, then

$$f = (I + \lambda A_0)u \in L^1 \cap L^2.$$

Hence, by construction, $u \in D(A_0) \cap H^1$. Thus,

$$(I + \lambda A_0)u = f = (I + \lambda A_0)(I + \lambda A_0)^{-1} f$$

and, by the mentioned injectivity, this implies $u = J_\lambda(f)$.

We note that by the same arguments it follows that

$$\overline{D(A_\varepsilon)} = L^1. \tag{3.48}$$

Note that (2.7) is immediate by (3.7), because by (3.12) we see that

$$\int_{\mathbb{R}^d} u_\varepsilon dx = \int_{\mathbb{R}^d} f\, dx,\ \forall \varepsilon > 0.$$

Finally, if $f \geq 0$, then we have $u_\varepsilon \geq 0$ in $\mathbb{R}^d$. Indeed, if we multiply (3.12) by $\mathcal{X}_\delta(u_\varepsilon^-)$ and integrate on $\mathbb{R}^d$, we get

$$\int_{\mathbb{R}^d} u_\varepsilon^- dx \leq 0$$

because, as seen earlier,

$$\lim_{\delta \to 0} \frac{1}{\delta} \int_{\mathbb{R}^d} \mathcal{X}'_\delta(u_\varepsilon^-) \nabla u_\varepsilon \cdot Db(u_\varepsilon) u_\varepsilon dx = 0.$$

This completes the proof of Proposition 2.1.

Proof of Proposition 2.3. By Lemma 3.1 and (3.40) in Lemma 3.2, we have, for $\lambda \in (0, \lambda_0)$, and $\delta > 0$,

$$\|(I + \lambda A)^{-1} u_0\| \le \|u_0\| + \rho\lambda |u_0|_1, \ \forall u_0 \in \mathcal{M}.$$

This yields

$$\|(I + \lambda A)^{-n} u_0\| \le \|u_0\| + n\lambda\rho |u_0|_1, \ \forall n \in \mathbb{N},$$

and so, by (2.15), we get

$$\|S(t)u_0\| \le \|u_0\| + \rho t |u_0|_1, \ \forall t \ge 0, \ u_0 \in \mathcal{M}, \tag{3.49}$$

as claimed.

4 The H-theorem

Let $S(t)$ be the continuous semigroup of contractions defined by (2.18). A (4.11)semicontinuous function $V : L^1 \to (-\infty, \infty]$ is said to be a *Lyapunov function* for $S(t)$ (equivalently, for equations (1.1) or (2.11)) if

$$V(S(t)u_0) \le V(S(s)u_0), \text{ for } 0 \le s \le t < \infty, \ u_0 \in L^1.$$

(See, e.g., [16] and [21].)

In the following, we shall restrict the semigroup to the probability density set $\mathcal{P}$ (see (2.22)). For each $u_0 \in \mathcal{P}$, consider the ω-limit set

$$\omega(u_0) = \{w = \lim S(t_n)u_0 \text{ in } L^1_{\rm loc} \text{ for some } \{t_n\} \to \infty\}.$$

Our aim here is to construct a Lyapunov function for $S(t)$, to prove that $\omega(u_0) \ne \emptyset$ and also that every $u_\infty \in \omega(u_0)$ is an equilibrium state of equation (1.1), that is, $Au_\infty = 0$. To this end, we shall assume that, besides (i)–(iv), hypothesis (v) also holds.

Consider the function $\eta \in C(\mathbb{R})$,

$$\eta(r) = -\int_0^r d\tau \int_\tau^1 \frac{\beta'(s)}{sb(s)}\, ds, \ \forall r \ge 0, \tag{4.1}$$

and define the function $V : D(V) = \{u \in \mathcal{M};\ u \ge 0, \text{ a.e. on } \mathbb{R}^d\} \to \mathbb{R}$ by

$$V(u) = \int_{\mathbb{R}^d} \eta(u(x))dx + \int_{\mathbb{R}^d} \Phi(x)u(x)dx = -S[u] + E[u]. \tag{4.2}$$

Since, by (i), (iv) and (v),

$$\frac{\gamma}{r|b|_\infty} \le \frac{\beta'(r)}{rb(r)} \le \frac{\gamma_1}{rb_0},\ \forall r \ge 0, \tag{4.3}$$

we have

$$\begin{aligned}&\frac{\gamma_1}{b_0}\,\mathbf{1}_{[0,1]}(r)r(\log r-1)+\frac{\gamma}{|b|_\infty}\,\mathbf{1}_{(1,\infty)}(r)r(\log r-1)\le\eta(r)\\&\le\frac{\gamma}{|b|_\infty}\,\mathbf{1}_{[0,1]}(r)r(\log r-1)+\frac{\gamma_1}{b_0}\,\mathbf{1}_{(1,\infty)}(r)r(\log r-1).\end{aligned}\tag{4.4}$$

We also have that $\eta \in C([0,\infty))$, $\eta \in C^2((0,\infty))$, $\eta'' \ge 0$. Since Φ is Lipschitz, hence of at most linear growth, $E[u]$ is well-defined and finite if $u \in \mathcal{M}$. Furthermore, exactly as in [21], p. 16, one proves that $(u\ln u)^- \in L^1$ if $u \in D(V)$. Hence $S[u]$ is well-defined and $-S[u] \in (-\infty,\infty]$ because of (4.4) and thus $V(u) \in (-\infty,\infty]$ for all $u \in D(V)$. We define $V = \infty$ on $L^1 \setminus D(V)$. Then, obviously, $V : L^1 \to (-\infty,\infty]$ is convex and $L^1_{\rm loc}$-lower semicontinuous on balls in $\mathcal{M}$, as easily follows by (4.4) from (4.5) below. If, in addition, $(u\ln u)^+ \in L^1$, then, again by (4.4), we have that $S[u] \in (-\infty,\infty)$ and also V is real-valued. The function (see (1.8))

$$S[u] = -\int_{\mathbb{R}^d} \eta(u(x))dx,\ u \in \mathcal{P},$$

is called in the literature (see, e.g., [18], [26]) the entropy of the system, while $E[u]$ is the mean field energy.

In fact, according to the general theory of thermostatics (see [19]), the functional $S = S[u]$ is a generalized entropy because its kernel $-\eta$ is a strictly concave continuous functions on $(0,\infty)$ and $\lim\limits_{r\downarrow 0}\eta'(r) = +\infty$. In the special case $\beta(s) \equiv s$ and $b(s) \equiv 1$, $\eta(r) \equiv r(\log r - 1)$ and so $S[u] - 1$ reduces to the classical Boltzman-Gibbs entropy.

As in [21] (formula (15), one proves that, for $\alpha \in \left[\frac{m}{m+1}, 1\right)$, where m is as in assumption (iv),

$$\int_{\{\Phi\ge R\}} |\min(u\log u,0)|dx \le C_\alpha \left(\int_{\{\Phi\ge R\}} \Phi^{-m}dx\right)^{1-\alpha} \|u\|^\alpha, \tag{4.5}$$

for all $R > 0$. Indeed, obviously, for every $\alpha \in (0,1)$, there exists $C_\alpha \in (0,\infty)$ such that

$$(r\log r)^- \le C_\alpha r^\alpha \ \text{ for } r \in [0,\infty).$$

Hence, the left hand side of (4.5) by Hölder's inequality is dominated by

$$C_\alpha \left(\int_{\{\Phi \geq R\}} u\Phi dx \right)^\alpha \left(\int_{\{\Phi \geq R\}} \Phi^{-\frac{\alpha}{1-\alpha}} dx \right)^{1-\alpha}.$$

Therefore, for $\alpha \in \left[\frac{m}{m+1}, 1\right)$, we obtain (4.5) since $\Phi \geq 1$. Inequality (4.5) yields

$$V(u) \geq -C(\|u\| + 1)^\alpha, \ \forall u \in D(V). \tag{4.6}$$

We also consider the function $\Psi : D(\Psi) \subset L^1 \to [0, \infty)$ defined by

$$\Psi(u) \ = \ \int_{\mathbb{R}^d} \left| \frac{\beta'(u)\nabla u}{\sqrt{ub(u)}} - D\sqrt{ub(u)} \right|_d^2 dx, \tag{4.7}$$

$$D(\Psi) \ = \ \{u \in L^1 \cap W^{1,1}_{\text{loc}}(\mathbb{R}^d); \ u \geq 0, \ \Psi(u) < \infty\}. \tag{4.8}$$

We extend Ψ to all of L^1 by $\Psi(u) = \infty$ if $u \in L^1 \setminus D(\Psi)$. Since $\nabla u = 0$, a.e. on $\{u = 0\}$, we set here and below

$$\frac{\nabla u}{\sqrt{u}} = 0 \ \text{ on } \{u = 0\}.$$

Theorem 4.1 is the main result and, as mentioned earlier, can be viewed as the H-theorem for NFPE (1.1).

Theorem 4.1 *Assume that hypotheses* (i)–(v) *hold. Then the function* V *defined by* (4.1) *is a Lyapunov function for* $S(t)$, *that is, for* $D_0(V) = D(V) \cap \{V < \infty\}$,

$$\begin{gathered} S(t)u_0 \in D_0(V), \forall t \geq 0, u_0 \in D_0(V) \ \textit{and} \\ V(S(t)u_0) \leq V(S(s)u_0), \forall u_0 \in D_0(V), 0 \leq s \leq t < \infty. \end{gathered} \tag{4.9}$$

Moreover, we have, for all $u_0 \in D_0(V)$,

$$V(S(t)u_0) + \int_s^t \Psi(S(\sigma)u_0) d\sigma \leq V(S(s)u_0) \ \textit{for } 0 \leq s \leq t < \infty. \tag{4.10}$$

In particular, $S(\sigma)u_0 \in D(\Psi)$ *for a.e.* $\sigma \geq 0$. *Furthermore, there exists* $u_\infty \in \omega(u_0)$ *(see* (1.6)*) such that* $u_\infty \in D(\Psi)$, $\Psi(u_\infty) = 0$. *Furthermore, for any such a* u_∞ *we have either* $u_\infty = 0$ *or* $u_\infty > 0$ *a.e., and in the latter case,*

$$u_\infty = g^{-1}(-\Phi + \mu) \ \textit{for some } \mu \in \mathbb{R}, \tag{4.11}$$

where

$$g(r) = \int_1^r \frac{\beta'(s)}{b(s)}\ ds,\ r > 0. \tag{4.12}$$

Moreover, by (4.2), (4.10), we see that the entropy of the semiflow $u(t) = S(t)u_0$ is evolving according to the law

$$S[u(t)] \geq S[u(s)] + \int_{\mathbb{R}^d} \Phi(x)(u(t,x) - u(s,x))ds + \int_s^t \Psi(u(\sigma))d\sigma,$$

for all $0 \leq s \leq t < \infty$.

5 Proof of Theorem 4.1

In the following, we approximate $V : L^1 \to (-\infty, \infty]$ by the functional V_ε defined by

$$V_\varepsilon(u) = \int_{\mathbb{R}^d} (\eta_\varepsilon(u(x)) + \Phi_\varepsilon(x)u(x))dx,\ \forall u \in D(V),$$
$$V_\varepsilon(u) = \infty \ \text{ if } u \in L^1 \setminus D(V),$$

where $\eta_\varepsilon(r) = -\int_0^r d\tau \int_\tau^1 \frac{\beta'(s)}{b^*_\varepsilon(s)+\varepsilon^{2m}}\, ds,\ r \geq 0,\ \varepsilon > 0$. Clearly, $\eta_\varepsilon \to \eta$ as $\varepsilon \to 0$ locally uniformly. We note that V_ε is convex, and $L^1_{\rm loc}$-lower semicontinuous on every ball in $\mathcal{M}$. Furthermore, there exists $C > 0$ such that, for all $\varepsilon \in (0, 1]$, we have $|\eta_\varepsilon(u)| \leq C(1+|u|^2)$. This implies that $V_\varepsilon < \infty$ on L^2 and $V_\varepsilon(u) \to V(u)$ as $\varepsilon \to 0$ for all $u \in D(V) \cap L^2$ and by the generalized Fatou lemma that V_ε is lower semicontinuous on L^2. We set

$$V'_\varepsilon(u) = \eta'_\varepsilon(u) + \Phi_\varepsilon,\ \forall u \in D(V) \cap L^2.$$

It is easy to check that $V'_\varepsilon(u) \in \partial V_\varepsilon(u)$ for all $u \in D(V) \cap L^2$, where ∂V_ε is the subdifferential of V_ε on L^2. As regards the function Ψ defined by (4.7)–(4.8), we have

Lemma 5.1 *We have*

$$D(\Psi) = \{u \in L^1;\ u \geq 0,\ \sqrt{u} \in W^{1,2}(\mathbb{R}^d)\}, \tag{5.1}$$

$$\|\sqrt{u}\|_{W^{1,2}(\mathbb{R}^d)} \leq C(\Psi(u) + 1),\ \forall u \in D(\Psi), \tag{5.2}$$

where $C \in (0, \infty)$ is independent of u. Furthermore, Ψ is $L^1_{\rm loc}$-lower semicontinuous on L^1-balls.

Proof. By (4.7), taking into account (i), (ii), we have

$$\begin{aligned}\gamma|b|_\infty^{-1}\int_{\mathbb{R}^d}\frac{|\nabla u|^2}{u}\,dx &\le \int_{\mathbb{R}^d}\frac{|\beta'(u)|^2\cdot|\nabla u|^2}{ub(u)}\,dx\\ &\le 2\Psi(u)+2\int_{\mathbb{R}^d}|D|^2ub(u)dx<\infty,\ \forall u\in D(\Psi).\end{aligned}\tag{5.3}$$

This yields (5.1) and (5.2) since $\nabla(\sqrt{u})=\frac{1}{2}\,\frac{\nabla u}{\sqrt{u}}$ and (v) holds. To show the lower semicontinuity of Ψ, we rewrite it as

$$\Psi(u)=\int_{\mathbb{R}^d}|\nabla j(u)-D\sqrt{ub(u)}|_d^2dx,\ u\in D(\Psi),\tag{5.4}$$

where

$$j(r)=\int_0^r\frac{\beta'(s)}{\sqrt{sb(s)}}\,ds,\ r\ge 0.\tag{5.5}$$

Clearly,

$$0\le j(r)\le\frac{2\gamma_1}{\sqrt{b_0}}\,\sqrt{r}.\tag{5.6}$$

Let $\{u_n\}\subset L^1$ and $\nu>0$ be such that $\sup\limits_n|u_n|_1<\infty$ and

$$\Psi(u_n)\le\nu<\infty,\ \forall n,\tag{5.7}$$

$$u_n\longrightarrow u\text{ in }L^1_{\rm loc}\text{ as }n\to\infty.\tag{5.8}$$

(5.8) yields

$$\sqrt{u_nb(u_n)}\longrightarrow\sqrt{ub(u)}\text{ in }L^2_{\rm loc}$$

and so, by hypothesis (iii), we have

$$D\sqrt{u_nb(u_n)}\longrightarrow D\sqrt{ub(u)}\text{ in }L^2_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d).\tag{5.9}$$

Hence (5.7) implies that (selecting a subsequence if necessary) for all balls B_N of radius $N\in\mathbb{N}$ around zero we have

$$\sup_n\int_{B_N}|\nabla j(u_n)|^2dx<\infty$$

and

$$j(u_n)\to h(u)\ \text{ in }L^2_{\rm loc}\ \text{ as }n\to\infty.$$

Therefore (again selecting a subsequence, if necessary), for every $N \in \mathbb{N}$,

$$\nabla j(u_n) \to \nabla j(u) \text{ weakly in } L^2(B_N, dx) \text{ as } n \to \infty.$$

Hence, if we define Ψ_N analogously to Ψ, but with the integral over $\mathbb{R}^d$ replaced by an integral over B_N, we conclude that

$$\liminf_{n\to\infty} \Psi_N(u_n) \geq \liminf_{n\to\infty} \int_{B_N} |\nabla j(u_n)|_d^2 dx - 2\int_{B_N} \nabla j(u) \cdot D\sqrt{ub(u)}dx + \int_{B_N} |D|_d^2 ub(u)dx \geq \Psi_N(u).$$

Hence, since $u \in L^1$, we can let $N \to \infty$ to get

$$\liminf_{n\to\infty} \Psi(u_n) \geq \Psi(u).$$

Now, we consider the functional

$$\begin{aligned}\Psi_\varepsilon(u) = &\int_{\mathbb{R}^d} \left| \frac{\beta'(u)\nabla u}{\sqrt{b_\varepsilon^*(u)+\varepsilon^{2m}}} - D_\varepsilon\sqrt{b_\varepsilon^*(u)+\varepsilon^{2m}} \right|^2 dx \\ &+\varepsilon^{2m}\int_{\mathbb{R}^d} D_\varepsilon \cdot \left(\frac{\beta'(u)\nabla u}{b_\varepsilon^*(u)+\varepsilon^{2m}} - D_\varepsilon \right) dx \\ &+\varepsilon \int_{\mathbb{R}^d} \beta(u)(\eta_\varepsilon'(u)+\Phi_\varepsilon)dx,\ \forall u \in D(\Psi_\varepsilon) = D(V)\cap H^1,\end{aligned} \tag{5.10}$$

and

$$\Psi_\varepsilon(u) := \infty \ \text{ if } u \in D(V)\setminus H^1.$$

We have

Lemma 5.2 *For each $\varepsilon > 0$, Ψ_ε is $L^1_{\rm loc}$-lower semicontinuous on every ball in $\mathcal{M}$. Moreover, for any sequence $\{v_\varepsilon\} \subset D(V)\cap H^1$ such that*

$$\sup_{\varepsilon\geq 0} \|v_\varepsilon\| < \infty,\ \lim_{\varepsilon\to 0} v_\varepsilon = v \text{ in } L^1_{\rm loc},$$

we have

$$\liminf_{\varepsilon\to 0} \Psi_\varepsilon(v_\varepsilon) \geq \Psi(v). \tag{5.11}$$

Furthermore, there exists $c \in (0,\infty)$ such that, for all $u \in D(V)$, $\varepsilon \in (0,1]$,

$$\Psi_\varepsilon(u) \geq -c(|u| + \|u\| + 1). \tag{5.12}$$

Proof. We write

$$\Psi_\varepsilon(u) \equiv \Psi^*_\varepsilon(u) + G_\varepsilon(u),$$

where

$$\Psi^*_\varepsilon(u) = \int_{\mathbb{R}^d} \left| \frac{\beta'(u)\nabla u}{\sqrt{b^*_\varepsilon(u) + \varepsilon^{2m}}} - D_\varepsilon \sqrt{b^*_\varepsilon(u) + \varepsilon^{2m}} \right|^2 dx$$
$$+\varepsilon^{2m} \int_{\mathbb{R}^d} D_\varepsilon \cdot \left(\frac{\beta'(u)\nabla u}{b^*_\varepsilon(u) + \varepsilon^{2m}} - D_\varepsilon \right) dx,$$
$$G_\varepsilon(u) = \varepsilon \int_{\mathbb{R}^d} \beta(u)(\eta'_\varepsilon(u) + \Phi_\varepsilon)dx.$$

We have, since $\eta'_\varepsilon(\tau) \geq \frac{\gamma_1}{b_0}(\log\tau - \varepsilon(1-\tau))$ for $\tau \in (0,1]$,

$$\begin{aligned} G_\varepsilon(v_\varepsilon) &\geq \varepsilon\gamma_1 \int_{\{v_\varepsilon \leq 1\}} v_\varepsilon \eta'_\varepsilon(v_\varepsilon)dx \geq \varepsilon\,\frac{\gamma_1^2}{b_0} \int_{\{v_\varepsilon\leq 1\}} (v_\varepsilon \log v_\varepsilon - \varepsilon v_\varepsilon)dx \\ &\geq -\varepsilon\,\frac{\gamma_1^2}{b_0}\left[C_\alpha \left(\int_{\mathbb{R}^d} \Phi^{-m}dx \right)^{1-\alpha} \|v_\varepsilon\|^\alpha + \varepsilon \int_{\mathbb{R}^d} v_\varepsilon \Phi\,dx \right. \\ &\quad \left. + \int_{\{\Phi\leq 1\}} ((v_\varepsilon \log v_\varepsilon)^- + \varepsilon)dx \right], \end{aligned} \tag{5.13}$$

where we used (4.5). Hence

$$\liminf_{\varepsilon\to 0} G_\varepsilon(v_\varepsilon) \geq 0.$$

Now, arguing as in the proof of Lemma 5.1, we represent Ψ^*_ε as (see (5.3))

$$\Psi^*_\varepsilon(u) = \int_{\mathbb{R}^d} |\nabla j^*_\varepsilon(u) - D_\varepsilon \sqrt{b^*_\varepsilon(u) + \varepsilon^{2m}}|^2 dx + \varepsilon^{2m} \int_{\mathbb{R}^d} D_\varepsilon \cdot \left(\frac{\beta'(u)\nabla u}{b^*_\varepsilon(u) + \varepsilon^{2m}} - D_\varepsilon \right) dx,$$

where $u \in D(V) \cap H^1$ and

$$j^*_\varepsilon(r) = \int_0^r \frac{\beta'(s)ds}{\sqrt{b^*_\varepsilon(s) + \varepsilon^{2m}}}.$$

We may assume that $\Psi^*_\varepsilon(v_\varepsilon) \leq \nu < \infty$, $\forall \varepsilon > 0$. Then, as in (5.3), we see that

$$\begin{aligned} \int_{\mathbb{R}^d} \frac{|\beta'(v_\varepsilon)|^2 |\nabla v_\varepsilon|^2}{b^*_\varepsilon(v_\varepsilon) + \varepsilon^{2m}}\,dx &\leq 2\left(\Psi^*_\varepsilon(v_\varepsilon) + \int_{\mathbb{R}^d} |D_\varepsilon|^2 (b^*_\varepsilon(v_\varepsilon) + 2\varepsilon^{2m})dx \right) \\ &\quad + 2\varepsilon^{2m} \int_{\mathbb{R}^d} \frac{|D_\varepsilon|\beta'(v_\varepsilon)|\nabla v_\varepsilon|}{b^*_\varepsilon(v_\varepsilon) + \varepsilon^{2m}}\,dx. \end{aligned} \tag{5.14}$$

Taking into account that

$$\begin{aligned}\varepsilon^{2m}\int_{\mathbb{R}^d}&\frac{|D_\varepsilon|\beta'(v_\varepsilon)|\nabla v_\varepsilon|}{b^*_\varepsilon(v_\varepsilon)+\varepsilon^{2m}}\,dx\\ &\le\frac12\int_{\mathbb{R}^d}\frac{|\beta'(v_\varepsilon)|^2|\nabla v_\varepsilon|^2}{b^*_\varepsilon(v_\varepsilon)+\varepsilon^{2m}}\,dx+\frac{\varepsilon^{4m}}{2}\int_{\mathbb{R}^d}\frac{|D_\varepsilon|^2}{b^*_\varepsilon(v_\varepsilon)+\varepsilon^{2m}}\,dx\\ &\le\frac12\int_{\mathbb{R}^d}\frac{|\beta'(v_\varepsilon)|^2|\nabla v_\varepsilon|^2}{b^*_\varepsilon(v_\varepsilon)+\varepsilon^{2m}}\,dx+\frac{\varepsilon^{2m}}{2}\int_{\mathbb{R}^d}|D_\varepsilon|^2dx,\end{aligned}\tag{5.15}$$

and that $\lim\limits_{\varepsilon\to0}v_\varepsilon=v$ in L^1 by our assumption, it follows by (3.5) and (5.14) that, for some $C>0$ independent of ε,

$$\int_{\mathbb{R}^d}\frac{|\nabla v_\varepsilon|^2}{b^*_\varepsilon(v_\varepsilon)+\varepsilon^{2m}}\ dx\le C,\ \forall\varepsilon>0,$$

and so $\{\nabla j^*_\varepsilon(v_\varepsilon)\}$ is bounded in L^2. Then, arguing as in Lemma 5.1 (see (5.8)–(5.9)), we get for $\varepsilon\to0$

$$D_\varepsilon\sqrt{b^*_\varepsilon(v_\varepsilon)+\varepsilon^{2m}}\longrightarrow D\sqrt{b(u)u}\ \text{ in } L^2(\mathbb{R}^d;\mathbb{R}^d),$$

and, therefore,

$$\liminf_{\varepsilon\to0}\Psi_\varepsilon(v_\varepsilon)\ge\liminf_{\varepsilon\to0}\Psi^*_\varepsilon(v_\varepsilon)\ge\Psi(v),$$

as claimed. By a similar (even easier) argument, one proves that Ψ_ε is $L^1_{\rm loc}$-lower semicontinuous on balls in $\mathcal{M}$. The last part of the assertion is an immediate consequence of (5.13) and (5.15), which holds for all $u\in D(V)\cap H^1$ replacing v_ε. Hence, the lemma is proved.

We denote by $S_\varepsilon(t)$ the continuous semigroup of contractions on L^1 generated by the m-accretive operator A_ε defined by (3.1)–(3.2), that is,

$$S_\varepsilon(t)u_0=\lim_{n\to\infty}\left(I+\frac tn\,A_\varepsilon\right)^{-n}u_0,\ \forall t\ge0,\ u_0\in L^1.\tag{5.16}$$

We note that by (3.7) it follows, by virtue of the Trotter-Kato theorem for nonlinear semigroups of contractions, that (see [11] and [3], p. 169)

$$\lim_{\varepsilon\to0}S_\varepsilon(t)u_0=S(t)u_0,\ \forall u_0\in L^1,\tag{5.17}$$

strongly in L^1 uniformly on compact time intervals.

We shall prove first (4.10) for $S_\varepsilon(t)$. Namely, one has

Lemma 5.3 *For each $u_0 \in L^2 \cap D(V)$, we have $S_\varepsilon(\sigma)u_0 \in D(\Psi_\varepsilon)$ for ds-a.e. $\sigma \geq 0$, and*

$$V_\varepsilon(S_\varepsilon(t)u_0) + \int_s^t \Psi_\varepsilon(S_\varepsilon(\sigma)u_0)d\sigma \leq V_\varepsilon(S_\varepsilon(s)u_0),\ 0 \leq s \leq t < \infty, \tag{5.18}$$

and all three terms are finite.

Proof. First, we shall prove that, for all $\varepsilon > 0$,

$$V_\varepsilon(I + \lambda A_\varepsilon)^{-1}u_0) + \lambda\Psi_\varepsilon((I + \lambda A_\varepsilon)^{-1}u_0) \leq V_\varepsilon(u_0),\ \lambda \in (0, \lambda_0). \tag{5.19}$$

We set $u_\varepsilon^\lambda = (I + \lambda A_\varepsilon)^{-1}u_0$ and note that, by (3.17)–(3.18), we have

$$u_\varepsilon^\lambda \in H^1(\mathbb{R}^d),\ \beta(u_\varepsilon^\lambda) \in H^1(\mathbb{R}^d),\ \forall\lambda \in (0, \lambda_0),\ \varepsilon > 0, \tag{5.20}$$

and

$$V'_\varepsilon(u_\varepsilon^\lambda) = \eta'_\varepsilon(u_\varepsilon^\lambda) + \Phi_\varepsilon \in \partial V_\varepsilon(u_\varepsilon^\lambda), \tag{5.21}$$

where

$$\eta'_\varepsilon(u_\varepsilon^\lambda) \in H^1(\mathbb{R}^d). \tag{5.22}$$

Taking into account that, by Lemma 3.2,

$$\operatorname{div}(\nabla\beta(u_\varepsilon^\lambda) - D_\varepsilon b_\varepsilon^*(u_\varepsilon^\lambda)) = \frac{1}{\lambda}\,(u_\varepsilon^\lambda - u_0) + \varepsilon\beta(u_\varepsilon^\lambda) \in \mathcal{M}, \tag{5.23}$$

it follows, since $\Phi_\varepsilon \in L^2$,

$$\int_{\mathbb{R}^d} (-\Delta\beta(u_\varepsilon^\lambda) + \operatorname{div} D_\varepsilon b_\varepsilon^*(u_\varepsilon^\lambda))\Phi_\varepsilon\, dx = -\int_{\mathbb{R}^d} (\nabla\beta(u_\varepsilon^\lambda) - D_\varepsilon b_\varepsilon^*(u_\varepsilon^\lambda)) \cdot D_\varepsilon\, dx.$$

This yields, by (5.21),

$$\begin{aligned}
&\left\langle A_\varepsilon(u_\varepsilon^\lambda), V'_\varepsilon(u_\varepsilon^\lambda)\right\rangle_2 \\
&\quad = \left\langle -\Delta(\beta(u_\varepsilon^\lambda)) + \varepsilon\beta(u_\varepsilon^\lambda) + \operatorname{div}(D_\varepsilon b_\varepsilon^*(u_\varepsilon^\lambda)), \eta'_\varepsilon(u_\varepsilon^\lambda) + \Phi_\varepsilon\right\rangle_2 \\
&\quad = \int_{\mathbb{R}^d} (\beta'(u_\varepsilon^\lambda)\nabla u_\varepsilon^\lambda - D_\varepsilon b_\varepsilon^*(u_\varepsilon^\lambda)) \cdot \left(\frac{\beta'(u_\varepsilon^\lambda)}{b_\varepsilon^*(u_\varepsilon^\lambda) + \varepsilon^{2m}}\nabla u_\varepsilon^\lambda - D_\varepsilon\right) dx \\
&\qquad\quad + \varepsilon\left\langle \beta(u_\varepsilon^\lambda), \eta'_\varepsilon(u_\varepsilon^\lambda) + \Phi_\varepsilon\right\rangle_2 \\
&\quad = \int_{\mathbb{R}^d} \left| \frac{\beta'(u_\varepsilon^\lambda)\nabla u_\varepsilon^\lambda}{\sqrt{b_\varepsilon^*(u_\varepsilon^\lambda) + \varepsilon^{2m}}} - D_\varepsilon\sqrt{b_\varepsilon^*(u_\varepsilon^\lambda) + \varepsilon^{2m}} \right|^2 dx + \varepsilon\left\langle \beta(u_\varepsilon^\lambda), \eta'_\varepsilon(u_\varepsilon^\lambda) + \Phi_\varepsilon\right\rangle_2 \\
&\qquad\quad + \varepsilon^{2m}\int_{\mathbb{R}^d} \left(D_\varepsilon \cdot \frac{\beta'(u_\varepsilon^\lambda)\nabla u_\varepsilon^\lambda}{b_\varepsilon^* + \varepsilon^{2m}} - D_\varepsilon\right) dx \\
&\quad = \Psi_\varepsilon(u_\varepsilon^\lambda),\ \forall\varepsilon > 0,\ \lambda \in (0, \lambda_0).
\end{aligned}$$

This yields (5.19) because, by the convexity of V_ε, we have by (5.21)

$$V_\varepsilon(u_\varepsilon^\lambda) \le V_\varepsilon(u_0) + \left\langle V_\varepsilon'(u_\varepsilon^\lambda), u_\varepsilon^\lambda - u_0 \right\rangle_2, \ u_\varepsilon^\lambda - u_0 = -\lambda A_\varepsilon(u_\varepsilon^\lambda).$$

To get (5.18), we shall proceed as in the proof of Theorem 3.4 in [25]. Namely, we set

$$\begin{aligned}\lambda\delta(\lambda, v) = V_\varepsilon((I+\lambda A_\varepsilon)^{-1}v) + \lambda\Psi_\varepsilon((I+\lambda A_\varepsilon)^{-1}v) - V_\varepsilon(v),\\ \forall \lambda \in (0,\lambda_0),\ v \in L^2 \cap D(V),\end{aligned}$$

and note that, by (5.19), $\delta(\lambda, u_0) \le 0$, $\lambda \in (0, \lambda_0)$. This yields

$$\begin{aligned}V_\varepsilon((I+\lambda A_\varepsilon)^{-j}u_0) + \lambda\Psi_\varepsilon((I+\lambda A_\varepsilon)^{-j}u_0) - V_\varepsilon((I+\lambda A_\varepsilon)^{-j+1}u_0)\\ = \lambda\delta(\lambda, (I+\lambda A_\varepsilon)^{-j+1}u_0),\ \forall j \in \mathbb{N}.\end{aligned}$$

Then, summing up from $j = 1$ to $j = n$ and taking $\lambda = \frac{t}{n}$, we get

$$\begin{aligned}V_\varepsilon\left(\left(I + \frac{t}{n}A_\varepsilon\right)^{-n} u_0\right) + \sum_{j=1}^n \frac{t}{n}\Psi_\varepsilon\left(\left(I + \frac{t}{n}A_\varepsilon\right)^{-j} u_0\right)\\ = V_\varepsilon(u_0) + \sum_{j=1}^n \frac{t}{n}\ \delta\left(\frac{t}{n}, \left(I + \frac{t}{n}\ A_\varepsilon\right)^{-(j-1)} u_0\right).\end{aligned} \tag{5.24}$$

Note also that, if $n > \frac{t}{\lambda_0}$, then

$$\delta\left(\frac{t}{n}, \left(I + \frac{t}{n}\ A_\varepsilon\right)^{-j} u_0\right) \le 0, \quad 1 \le j \le n. \tag{5.25}$$

We consider the step function

$$f_n(\sigma) = \Psi_\varepsilon\left(\left(I + \frac{t}{n}\ A_\varepsilon\right)^{-j} u_0\right) \text{ for } \frac{(j-1)t}{n} < \sigma \le \frac{jt}{n},$$

and note that, for each $t > 0$,

$$\sum_{j=1}^n \frac{t}{n}\ \Psi_\varepsilon\left(\left(I + \frac{t}{n}A_\varepsilon\right)^{-j} u_0\right) = \int_0^t f_n(\sigma)d\sigma.$$

Then, by (3.40), (5.16) and the L^1_{loc}-lower semicontinuity of Ψ_ε on balls in $\mathcal{M}$, we conclude, by the Fatou lemma, which is applicable because of (5.12), that

$$-\infty < \int_0^t \Psi_\varepsilon(S(\sigma)u_0)d\sigma \le \liminf_{n\to\infty} \int_0^t f_n(\sigma)d\sigma, \tag{5.26}$$

while, by the L^1_{loc}-lower semicontinuity of V_ε on balls in $\mathcal{M}$, we have

$$\liminf_{n\to\infty} V_\varepsilon\left(\left(I+\frac{t}{n}\,A_\varepsilon\right)^{-n} u_0\right) \ge V_\varepsilon(S_\varepsilon(t)u_0).$$

Then, by (5.24)–(5.26), we get

$$V_\varepsilon(S_\varepsilon(t)u_0) + \int_0^t \Psi_\varepsilon(S_\varepsilon(\sigma)u_0)d\sigma \le V_\varepsilon(u_0),\ \forall t \ge 0.$$

In particular, $V_\varepsilon(S_\varepsilon(t)u_0) < \infty$ since $V_\varepsilon(u_0) < \infty$. Taking this into account and that $S_\varepsilon(t+s)u_0 = S_\varepsilon(t)S_\varepsilon(s)u_0$, we get (5.18), as claimed.

Proof of Theorem 4.1 (continued). We shall assume first $u_0 \in L^2 \cap D_0(V)$. We want to let $\varepsilon \to 0$ in (5.18), where $s = 0$.

We note first that we have

$$\liminf_{\varepsilon\to 0} V_\varepsilon(S_\varepsilon(t)u_0) \ge V(S(t)u_0),\ \forall t \ge 0. \tag{5.27}$$

Here is the argument.

First, we note that, if $v_\varepsilon \to v$ in L^1 as $\varepsilon \to 0$ and $\sup\limits_{\varepsilon>0} \|v_\varepsilon\| < \infty$, then $v_\varepsilon(\log v_\varepsilon)^- \to v(\log v)^-$ in L^1_{loc} as $\varepsilon \to 0$. Furthermore, for $\delta > 0$, and $\alpha \in \left[\frac{m+\delta}{m+\delta+1}, 1\right)$, by (4.5),

$$\int_{\{\Phi\ge R\}} v_\varepsilon(\log v_\varepsilon)^- dx \le C_\alpha\ \frac{1}{R^{\varepsilon(1-\alpha)}} \left(\int \Phi^{-m} dx\right)^{1-\alpha} \|v_\varepsilon\|^\alpha,$$

hence

$$\lim_{R\to\infty} \sup_{\varepsilon>0} \int_{\{\Phi\ge R\}} v_\varepsilon(\log v_\varepsilon)^- dx = 0,$$

therefore, $v_\varepsilon(\log v_\varepsilon)^- \to v(\log v)^-$ in L^1. Applying this to $v_\varepsilon = S_\varepsilon(t)u_0$, which by (5.17), (3.40) and (5.16) is justified, and because $\eta_\varepsilon \to \eta$ as $\varepsilon \to 0$ locally uniformly on $[0,\infty)$ and, because for all $\varepsilon \in (0,1]$, $r \in [0,\infty)$,

$$\eta_\varepsilon(r) \ge -\frac{\gamma_1}{b_0}\ (r\wedge 1)(\log(r\wedge 1)^- - 2(r\wedge 1)),$$

we can apply the generalized Fatou lemma to conclude that

$$\liminf_{\varepsilon\to\infty} \int_{\mathbb{R}^d} \eta_\varepsilon(S_\varepsilon(t)u_0)dx \ge \int_{\mathbb{R}^d} \eta(S(t)u_0)dx,$$

and we get (5.27), as claimed.

By Lemma 5.3, (3.40) and (5.16), we have that $v_\varepsilon = S_\varepsilon(t)u_0$, $\varepsilon > 0$, satisfy for dt-a.e. $t > 0$ the assumptions of Lemma 5.2, hence

$$\lim_{\varepsilon\to 0} \Psi_\varepsilon(S_\varepsilon(t)u_0) \ge \Psi(S(t)u_0), \text{ a.e. } t > 0.$$

Moreover, by Fatou's lemma, which is applicable by (5.12), it follows that

$$\liminf_{\varepsilon\to 0} \int_0^t \Psi_\varepsilon(S_\varepsilon(s)u_0)ds \ge \int_0^t \Psi(S(s)u_0)ds, \ \forall t \ge 0. \tag{5.28}$$

Because, as mentioned earlier, $V_\varepsilon(u) \to V(u)$ as $\varepsilon \to 0$, if $u \in D(V) \cap L^2$, (5.27), (5.28) and (5.18) with $s = 0$ imply

$$V(S(t)u_0) + \int_0^t \Psi(S(\sigma)u_0)d\sigma \le V(u_0), \ \forall u_0 \in D(V) \cap L^2, \ t \ge 0. \tag{5.29}$$

We note that, by (2.26) and (4.6), we have

$$\begin{array}{rcl} V(S(t)u_0) & \ge & -C(\|S(t)u_0\| + 1)^\alpha \\ & \ge & -C(\|u_0\| + t|u_0|_1)^\alpha, \ \alpha \in \left[\frac{m}{m+1}, 1\right). \end{array} \tag{5.30}$$

Hence

$$0 \le \int_0^t \Psi(S(\sigma)u_0)d\sigma < \infty, \ \forall t \ge 0,$$

which implies that

$$S(\sigma)u_0 \in D(\Psi) \ \text{ a.e. } \sigma > 0. \tag{5.31}$$

Now, to extend (5.29) to all $u_0 \in D_0(V)$, take $u_0^n \in D(V) \cap L^2 (\subset D_0(V))$ with $u_0^n \le u_0$ and $u_0^n \to u_0$ as $n \to \infty$ in L^1. Then, because for all $r \ge 0$

$$\eta(r) \ge -\frac{\gamma_0}{b_0} \left[(r \wedge 1)(\log(r \wedge 1)^- + (r \wedge 1))\right],$$

arguing as above (using again (4.5)), we conclude the monotone convergence applies to get

$$\lim_{n\to\infty} V(u_0^n) = V(u_0)$$

and the generalized Fatou lemma applies to get eventually (5.29) and (5.31) for all $u_0 \in D_0(V)$. Since $S(t)u_0 \in D_0(V)$, if $u_0 \in D_0(V)$, the first part including (4.10) follows.

To prove (4.11), we note that since $\alpha < 1$, by (4.10) and (5.30), we have

$$\begin{aligned} 0 &= \lim_{t\to\infty} \frac{1}{t}\int_0^t \Psi(S(\sigma)u_0)d\sigma \\ &\geq \lim_{t\to\infty} \frac{1}{t}\int_n^t \inf_{r\geq n} \Psi(S(r)u_0)d\sigma \\ &= \inf_{r\geq n} \Psi(S(r)) \quad \text{for all } n \in \mathbb{N}. \end{aligned}$$

Hence, there exists $t_n \to \infty$ such that

$$\lim_{n\to\infty} \Psi(S(t_n)u_0) = 0. \tag{5.32}$$

Furthermore, we obtain by Lemma 5.1 that

$$\sup_{t\geq 0} |S(t)u_0|_1 + \limsup_{t\to\infty} \frac{1}{t}\int_0^t |\nabla(\sqrt{S(s)u_0})|_2^2 ds < \infty.$$

Hence, there exist $t_n \to \infty$ such that

$$\sup_n \|\sqrt{S(t_n)u_0}\|_{W^{1,2}(\mathbb{R}^d)} < \infty. \tag{5.33}$$

So, by the Rellich-Kondrachov theorem (see, e.g., [11], p. 284), the set

$$\{S(t_n)u_0 \mid n \in \mathbb{N}\}$$

is relatively compact in $L^1_{\rm loc}$. Hence, along a subsequence $\{t_{n'}\} \to \infty$, we have

$$\lim S(t_{n'})u_0 = u_\infty \text{ in } L^1_{\rm loc} \tag{5.34}$$

for some $u_\infty \in L^1$. Since Ψ is $L^1_{\rm loc}$-lower semicontinuous on L^1-balls by Lemma 5.1, this together with (5.32) implies that $u_\infty \in D(\Psi)$ and $\Psi(u_\infty) = 0$.

If $u_\infty \in D(\Psi)$, such that $\Psi(u_\infty) = 0$, then

$$\frac{\beta'(u_\infty)\nabla u_\infty}{\sqrt{u_\infty b(u_\infty)}} = D\sqrt{u_\infty b(u_\infty)}, \text{ a.e. in } \mathbb{R}^d. \tag{5.35}$$

Let us prove now that either $u_\infty \equiv 0$ or $u = u_\infty > 0$, a.e. in $\mathbb{R}^d$. To this end, we consider the solution $y = y(t,x)$ to the system

$$\begin{aligned}&y_i'(t) = \widetilde{D}_i(y_i(t)),\ t \geq 0,\ i = 1,...,d,\\&y_i(0) = x_i,\end{aligned}$$

where $\widetilde{D}_i \in C^1(\mathbb{R})$, $i = 1,....,d$, is an arbitrary vector field on $\mathbb{R}$, and $y(t) = \{y_i(t)\}_{i=1}^d$, $x = \{x_i\}_{i=1}^d$. If j is defined by (5.5), we have

$$\begin{aligned}\frac{d}{dt}\, j(u(y(t,x))) &= j_u(u(y(t,x))\nabla u(y(t,x)) \cdot y'(t))\\&= \frac{\beta'(u(y(t,x)))}{\sqrt{b(u(y(t,x)))u(y(t,x))}}\, \nabla u(y(t,x))\cdot\mathcal{D}(y(t,x)), \forall t \geq 0,\end{aligned}$$

where $\mathcal{D}(y) = (\mathcal{D}_{ij}(y))_{ij}$ with $\mathcal{D}_{ij}(y) = \delta_{ij}D_j(y)$. Then, by (5.35), it follows that

$$\frac{d}{dt}\, j\left(u(y(t,x))\right) = \sum_{i=1}^d \widetilde{D}_i(y_i(t))D_i(u(y(t,x)))y_i(t)\left(u(y(t,x))b(u(y(t,x)))\right)^{\frac{1}{2}}.$$

We note that

$$C_2 j(r) \leq \sqrt{rb(r)} \leq C_1 j(r), \forall r \geq 0,$$

where $C_1, C_2 > 0$. This yields

$$\frac{d}{dt}\, j(u(y(t,x))) \geq -\sum_{i=1}^d \widetilde{D}_i(y_i(t))D_i(u(y(t,x)))j(u(y(t,x))),\ \forall t \geq 0.$$

Hence

$$j(u(y(t,x))) \geq Cj(u(x)), \forall t \geq 0,\ x \in \mathbb{R}^d,$$

and, therefore,

$$j(u(x)) \geq Cj(u(e^{-\mathcal{D}t}x)),\ \forall t \geq 0,\ x \in \mathbb{R}^d,$$

where $e^{\mathcal{D}t}$ is the flow generated by $\mathcal{D}$. Since $\mathcal{D}$ is an arbitrary vector field on $\mathbb{R}^d$, it follows that, for fixed x and t, $\{e^{-\mathcal{D}t}x\}$ covers all $\mathbb{R}^d$. We infer that, if $u \not\equiv 0$, then $j(u(x)) > 0$, $\forall x \in \mathbb{R}^d$, and this implies that $u = u_\infty > 0$, a.e. on $\mathbb{R}^d$. For such a u_∞, this yields, because $\Psi(u_\infty) = 0$,

$$\nabla(g(u_\infty) + \Phi) = 0,\ \text{a.e. in } \mathbb{R}^d, \tag{5.36}$$

where

$$g(r)=\int_1^r \frac{\beta'(s)}{sb(s)}\,ds,\ \forall r>0.$$

By (5.36), we see that $g(u_\infty)+\Phi=\mu$ for some $\mu\in\mathbb{R}$, in $\mathbb{R}^d$ and, since g is strictly monotone, we have

$$u_\infty(x)=g^{-1}(-\Phi(x)+\mu),\quad x\in\mathbb{R}^d. \tag{5.37}$$

6 The asymptotic behaviour in L^1

We assume here that, besides (i)–(v), condition (vi) also holds.

Theorem 6.1 *Assume that hypotheses* (i)–(vi) *hold and let* $u_0\in D_0(V)\backslash\{0\}$. *Set*

$$\widetilde{\omega}(u_0)=\left\{\lim_{n\to\infty} S(t_n)u_0 \quad \textit{in } L^1,\ \{t_n\}\to\infty\right\}.$$

Then

$$\omega(u_0)=\widetilde{\omega}(u_0)=\{u_\infty\}, \tag{6.1}$$

and $u_\infty>0$, *a.e. Furthermore,* $u_\infty\in D_0(V)\cap D(\Psi), \Psi(u_\infty)=0, S(t)u_\infty=u_\infty$ *for* $t\geq 0$, $|u_\infty|_1=|u_0|_1$, *and it is given by*

$$u_\infty(x)=g^{-1}(-\Phi(x)+\mu),\ \forall x\in\mathbb{R}^d, \tag{6.2}$$

where μ *is the unique number in* $\mathbb{R}$ *such that*

$$\int_{\mathbb{R}^d} g^{-1}(-\Phi(x)+\mu)dx=\int_{\mathbb{R}^d} u_0\,dx, \tag{6.3}$$

and

$$g(r)=\int_1^r \frac{\beta'(s)}{sb(s)}\,ds,\ r>0.$$

In particular, for all $u_0\in D_0(V)$ *with the same* L^1*-norm, the sets in* (6.1) *coincide, and thus* u_∞ *is the only element in* $D_0(V)$ *such that* $S(t)u_\infty=u_\infty$ *for all* $t\geq 0$.

Proof. Let us first prove the following version of Proposition 2.3.

Lemma 6.2 *Under hypotheses* (i)–(vi), *we have, for all* $u_0\in\mathcal{M}$, $u_0\geq 0$, *a.e. in* $\mathbb{R}^d$,

$$\|S(t)u_0\|\leq\|u_0\|,\ \forall t\geq 0. \tag{6.4}$$

Proof. We note first that we have

$$\|(I+\lambda A)^{-1}u_0\| \le \|u_0\|,\ \forall \lambda \in (0,\lambda_0). \tag{6.5}$$

Indeed, arguing as in the proof of Lemma 3.2 and taking into account that $u_\varepsilon \ge 0$, we get by (3.41)–(3.44),

$$\begin{aligned}\int_{\mathbb{R}^d} u_\varepsilon \varphi_\nu dx \le \lambda \int_{\mathbb{R}^d} &(b^*_\varepsilon(u_\varepsilon)\nabla\Phi_\varepsilon \cdot (\nu\Phi_\varepsilon\nabla\Phi_\varepsilon - \nabla\Phi_\varepsilon) \\ &+\beta(u_\varepsilon)(\Delta\Phi_\varepsilon - \nu\Phi_\varepsilon\Delta\Phi_\varepsilon + \nu^2\Phi_\varepsilon|\nabla\Phi_\varepsilon|^2)\exp(-\nu\Phi_\varepsilon))dx \\ &+\int_{\mathbb{R}^d} u_0\varphi_\nu\, dx,\end{aligned} \tag{6.6}$$

which, for $\nu \to 0$, because both $\Delta\Phi_\varepsilon$ and $\nabla\Phi_\varepsilon$ are bounded, yields

$$\|u_\varepsilon\| \le \|u_0\| + \lambda \int_{\mathbb{R}^d} (\beta(u_\varepsilon)\Delta\Phi_\varepsilon - b^*_\varepsilon(u_\varepsilon)|\nabla\Phi_\varepsilon|^2)dx.$$

Since $0 \le \beta(u_\varepsilon) \le \gamma_1 u_\varepsilon$, $0 \le b^*_\varepsilon(u_\varepsilon) \le |b|_\infty u_\varepsilon$ and $u_\varepsilon \to u$ in L^1 as $\varepsilon \to 0$, we obtain by Fatou's lemma and (3.35), (3.44)

$$\begin{aligned}\|u\| &\le \|u_0\| + \lambda \int_{\mathbb{R}^d} (\beta(u)\Delta\Phi - b(u)u|\nabla\Phi|^2)dx \\ &\le \|u_0\| + \lambda \int_{\mathbb{R}^d} \beta(u)\left(\Delta\Phi - \frac{b_0}{\gamma_1}|\nabla\Phi|^2\right)dx \\ &\le \|u_0\|,\end{aligned}$$

where we used assumptions (v) and (vi). Hence, (6.5) is proved. Then, by (2.15) and (6.5), one gets (6.4), as claimed, and the lemma is proved.

Now, by (4.6) and (6.4), we have, for all $t \ge 0$,

$$V(S(t)u_0) \ge -C(\|S(t)u_0\|+1)^\alpha \ge -C(\|u_0\|+1)^\alpha,$$

hence, by (4.10),

$$\int_0^\infty \Psi(S(\sigma)u_0)d\sigma < \infty. \tag{6.7}$$

This implies that

$$\omega(u_0) \subset \{u \in D(\Psi);\ \Psi(u) = 0\}. \tag{6.8}$$

To prove this, we shall use a modification of the argument from the proof of Theorem 4.1 in [25].

Let $u_\infty \in \omega(u_0)$ and $\{t_n\} \to \infty$ such that

$$S(t_n)u_0 \to u_\infty \text{ in } L^1_{\rm loc}.$$

Assume that $\Psi(u_\infty) > \delta > 0$ and argue from this to a contradiction. This implies that there is a bounded open subset $\mathcal{O}$ of $\mathbb{R}^d$ such that

$$\Psi_{\mathcal{O}}(u_\infty) > \frac{\delta}{2} > 0, \tag{6.9}$$

where $\Psi_{\mathcal{O}}$ is the integral for (4.7) restricted to $\Psi_{\mathcal{O}}$. Since $\Psi_{\mathcal{O}}$ is lower semicontinuous in L^1, it follows by (6.9) that there is a $\mu = \mu(\delta) > 0$ such that

$$\Psi_{\mathcal{O}}(u) \ge \frac{\delta}{4} \ \text{ if } |u_\infty - u|_1 \le \mu. \tag{6.10}$$

Since $S(t),\ t > 0$, is a semigroup of contractions, we have

$$|S(t)u_0 - S(s)u_0|_1 \le \nu(|t-s|),\ \forall s,t \ge 0, \tag{6.11}$$

where $\nu(r) := \sup\{|S(s)u_0 - u_0|_1 : 0 \le s \le r\},\ r > 0$. Clearly, $\nu(r) \to 0$ as $r \to 0$. By (6.11), we have

$$|S(t)u_0 - u_\infty|_1 \le |S(t)u_0 - S(t_n)u_0|_1 + |S(t_n)u_0 - u_\infty|_1 \le \mu,$$

for $|t - t_n| \le \nu^{-1}\left(\frac{\mu}{2}\right),\ n \ge N(\mu)$, where ν^{-1} is the inverse function of ν. By (6.10), this yields

$$\Psi_{\mathcal{O}}(S(t)u_0) \ge \frac{\delta}{4} \text{ for } |t - t_n| \le \nu^{-1}\left(\frac{\mu}{2}\right),$$

and $n \ge N(\mu)$. But this contradicts (6.7).

(6.8) and Theorem 4.1 imply (6.2). By (6.4), we also have

$$\lim_{R\to\infty} \sup_{t\ge 0} \int_{\{\Phi \ge R\}} S(t)u_0\, dx = 0,$$

which implies that $\omega(u_0) = \widetilde{\omega}(u_0)$ and that $|u_\infty|_1 = |u_0|_1$ by (2.16) and (2.18).

Hence (6.3) follows and thus (6.1) also holds. By Fatou's lemma, it follows that $u_\infty \in D(V)$ and, by (5.37), (4.9) and the $L^1_{\rm loc}$-lower semicontinuity of

V on balls in $\mathcal{M}$, we conclude that $u_\infty \in D_0(V)$. Now, let us check that, for $t > 0$,

$$S(t)u_\infty = u_\infty.$$

So, let $t_n \to \infty$, such that

$$\lim_{n\to\infty} S(t_n)u_0 = u_\infty.$$

Then, for all $t > 0$, by the semigroup property and the L^1-continuity of $S(t)$,

$$S(t)u_\infty = \lim_{n\to\infty} S(t + t_n)u_0 \in \widetilde{\omega}(u_0) = \{u_\infty\}.$$

The last part of the assertion is obvious by (6.3).

Corollary 6.3 *Let u_∞ be as in Theorem 6.1. Then*

$$|u_\infty|_\infty \le \max\left(1, e^{\frac{|b|_\infty}{\gamma}(\mu-1)}\right),$$

where $\mu \in \mathbb{R}$ is as in (6.2).

Proof. For g as above, we have that g is strictly increasing and $g : (0,\infty) \to \mathbb{R}$ is bijective. Furthermore, by (4.3), we have, for $(0,\infty)$,

$$\frac{\gamma_1}{b_0}\, \mathbf{1}_{(0,1]}(r) \log r + \frac{\gamma}{|b|_\infty}\, \mathbf{1}_{(1,\infty)}(r) \log r \le g(r).$$

Hence, replacing r by $e^{\frac{b_0}{\gamma_1} r}$, $r \le 0$, we get

$$g^{-1}(r) \le e^{\frac{b_0}{\gamma_1} r},\ \ r \in (-\infty, 0],$$

and, replacing r by $e^{\frac{|b|_\infty}{\gamma} r}$, $r \in (0,\infty)$, we obtain

$$g^{-1}(r) \le e^{\frac{|b|_\infty}{\gamma} r},\ \ r \in (0,\infty).$$

This implies, by (6.2), for all $x \in \mathbb{R}^d$,

$$\begin{aligned}(0 <)u_\infty(x) &= g^{-1}(\mu - \Phi(x)) \\ &\le \mathbf{1}_{\{\mu\le\Phi\}}(x)e^{\frac{b_0}{\gamma_1}(\mu-\Phi(x))} + \mathbf{1}_{\{\mu>\Phi\}}(x)e^{\frac{|b|_\infty}{\gamma}(\mu-\Phi(x))} \\ &\le \max\left(1, e^{\frac{|b|_\infty}{\gamma}(\mu-1)}\right),\end{aligned}$$

since $\Phi \ge 1$.

We show now that Theorem 6.1 implies the uniqueness of solutions $u^* \in \mathcal{M}\cap\mathcal{P}\cap\{V<\infty\}$ of the stationary version of (1.1), that is, to the equation

$$-\Delta\beta(u^*) + \mathrm{div}(Db(u^*)u^*) = 0 \text{ in } \mathcal{D}'(\mathbb{R}^d). \tag{6.12}$$

We note that the set of all $u^* \in L^1(\mathbb{R}^d)$ satisfying (6.12) is just $A^{-1}(0)$.

Theorem 6.4 *Under hypotheses* (i)–(vi)*, there is a unique solution* $u^* \in \mathcal{M}\cap\mathcal{P}\cap\{V<\infty\}$ *to equation* (6.12)*.*

Proof. By Theorem 6.1, it follows that u_∞ is a solution to (6.12), so it only remains to prove the uniqueness. So, let $u^* \in \mathcal{M}^+\cap\mathcal{P}\cap A^{-1}(0)$. Then, by construction, $S(t)u^* = u^*$, $\forall t \ge 0$, in particular,

$$\lim_{t\to\infty} S(t)u^* = u^*.$$

So, if, in addition, $u^* \in \{V<\infty\}$, it follows by the above (taking $u_0 = u^*$) that $u^* = u_\infty$ with u_∞ being uniquely determined by $\int_{\mathbb{R}^d} u^* dx = 1$.

Theorem 6.5 *Let* $X^i(t),\ t \ge 0,\ i = 1,2,$ *be two stationary nonlinear distorted Brownian motions, i.e., both satisfy* (1.5) *with* $(\mathcal{F}^i_t)$*-Wiener processes* $W^i(t),\ t \ge 0,$ *on probability spaces* $(\Omega^i, \mathcal{F}^i, \mathbb{P}^i)$ *equipped with normal filtrations* $\mathcal{F}^i_t,\ t \ge 0,$ *with*

$$\mathbb{P}^i \circ (X^i(t))^{-1} = u^i_\infty\, dx,$$

and $u(t,x)$ *in* (1.5) *replaced by* $u^i_\infty(x)$ *for* $i = 1,2$*, respectively. Assume that* $u^i_\infty \in \mathcal{M}\cap\{V<\infty\},\ i = 1,2.$ *Then*

$$\mathbb{P}^i \circ (X^1)^{-1} = \mathbb{P}^i \circ (X^2)^{-1},$$

i.e., we have uniqueness in law of stationary nonlinear distorted Brownian motions with stationary measures in $\mathcal{M}\cap\{V<\infty\}$*.*

Proof. By Itô's formula, both u^1_∞ and u^2_∞ satisfy (6.12). Hence, by Theorem 6.4, we have $u^1_\infty = u^2_\infty = u_\infty$. Fix $T > 0$ and let

$$\Phi(r) := \frac{\beta(r)}{r},\ r \in \mathbb{R}.$$

Then Theorem 3.1 in [7] implies that, for each $s \in [0,T]$ and each $v_0 \in L^1\cap L^\infty$, there is at most one solution $v = v(t,x)$, $t \in [s,T]$, to

$$v_t - \Delta(\Phi(u_\infty)v) + \mathrm{div}(Db(u_\infty)v) = 0 \text{ in } \mathcal{D}'((0,T)\times\mathbb{R}^d,$$
$$v(0,\cdot) = v_0,$$

such that $v \in L^\infty((s,T)\times\mathbb{R}^d)$ and $t \mapsto v(t,x)dx$, $t \in [s,T]$ is narrowly continuous. But u_∞, the time marginal law of X^i under $\mathbb{P}^i$, $i = 1, 2$, is such a solution with $v_0 = u_\infty$, since $u_\infty \in L^\infty$ by Corollary 6.3. Hence, Lemma 2.12 in [27] implies the assertion, since by Itô's formula $\mathbb{P}^i \circ (X^i)^{-1}$, $i = 1, 2$, both satisfy the martingale problem for the Kolmogorov operator

$$L_{u_\infty} = \Phi(u_\infty)\Delta + b(u_\infty)D\cdot\nabla.$$

Remark 6.6 By [6], a stationary nonlinear distorted Brownian motion as above always exists under the assumptions in this section.

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

References

[1] Arnold, A., Markowich, P., Toscani, G., Untesreiter, A., On convex inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations, *Comm. Partial Differential Equations*, vol. 26 (2001), 43-100.

[2] Bakry, D., Gentil, I., Ledoux, M., *Analysis and geometry of Markov diffusion operators*, Springer, 2014, xx+552 pp. ISBN: 378-3-319-00226-2.

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

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

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

[6] Barbu, V., Röckner, M., From nonlinear Fokker-Planck equations to solutions of distribution dependent SDE. arXiv:1808.10706v2, to appear in *Annals of Probability.*

[7] Barbu, V., Röckner, M., Uniqueness for nonlinear Fokker-Planck equations and weak uniqueness for McKean–Vlasov SDEs. arXiv: 1909.04464v1.

[8] Bogachev, V.I., Krylov, N.V., Röckner, M., Shaposhnikov, S.V., *Fokker-Planck-Kolmogorov equations,* Mathematical Surveys and Monographs, 207, American Mathematical Sociedy, Providence, R.I., 2015, xii+479 pp. ISBN: 978-1-4704-2558-6.

[9] Bogachev, V.I., Röckner, M., Shaposhnikov, S.V., Convergence in variation of solutions of nonliner Fokker-Planck-Kolmogorov equations to stationary measures, *J. Funct. Anal.*, 276 (12) (2019), 3681-3713.

[10] Brezis, H., *Functional Analysis Sobolev Spaces and Partial Differential Equations*, Springer 2010, New York, Dordrecht, Heidelberg, London.

[11] Brezis, H., Pazy, A., Convergence and apporiximation of semigroups of nonlinear operators in Banach spaces, *J. Funct. Anal.*, 9 (1972), 63-74.

[12] Carillo, J.A., Toscani, G., Asymptotic L^1-decay6 of solutions of the porous media equation to self-similarity, *Indiana Univ. Math. J.,* 49 (1) (2000), 113-142.

[13] Chavanis, P.H., Generalized stochastic Fokker-Planck equations, *Entropy*, 2015, 3205-3252.

[14] Chen, G.Q., Perthame, B., Well poosedness for nonisotropic degenerate parabolic hyperbolic equations, *Ann. Institute H. Poincaré*, 4 (2003), 645-668.

[15] Crandall, M.G., The semigroup approach to first order quasilinear equations in several space variables, *Israel J. Math.,* 10 (1972), 108-132.

[16] Dafermos, C., Slemrod, M., Asymptotic behavior of nonlinear contractions semigroups, *J. Funct. Anal.*, 13 (1973), 97-100.

[17] Eberle, E., Guillin, A., Zimmer, R., Quantitative Harris-type theorems for diffusions and McKean-Vlasov processes, *Trans. Amer. Math. Soc.*, 371 (10) (2019), 7135-7173.

[18] Franck, T.D., Generalized Fokker-Planck equations derived from generalized linear nonequilibrium thermodynamics, *Physica* A, 310 (2002), 397-412.

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

[20] Franck, T.D., Daffertshofer, A., H-theorem for nonlinear Fokker-Planck equations related to generalized thermostatics,*Physica A. Statistical Mechanics and its Applications*, 295(2001), 455-474.

[21] Jordan, R., Kinderlehrer, D., Otto, F., The variational formulation of the Fokker-Planck equation, *SIAM J. Math. Anal.*, 29 (1998), 1-17.

[22] Manita, O.A., Romanov, M.S., Shaposhnikov, S.V., On uniqueness of solutions to nonlinear Fokker-Planck-Kolmogorov equations, *Nonlin. Anal.*, 128 (2015), 199-226.

[23] Manita, O.A., Shaposhnikov, S.V., Nonlinear parabolic equations for measures, *St. Petersburg Math. J.*, 25 (1) (2014), 43-62.

[24] Markowich, P.A., Villani, C., On the trend to equilibrium for the Fokker-Planck equations: an interplay between physics and functional analysis, *Mathematics Contemporary*, 2000.

[25] Pazy, A., The Lyapunov method for semigroups of nonlinear contractions in Banach spaces, *Journal d'Analyse Math.*, 40 (1981), 239-262.

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

[27] Trevisan, D., Well–posedness of multidimensional diffusion processes with weakly differentiable coefficients, *Electron. J. of Probab.*, 21 (22) (2016), 1-41.

[28] Wang, F.-Y., *Functional inequalities, Markov semigroups and spectral theory*, Science Press, 2005, xx+379 pp. ISBN:7-03-014415-5.