



HAL
open science

Nonlinear Fokker-Planck equation with reflecting boundary conditions

Ioana Ciotir, Rim Fayad

► **To cite this version:**

Ioana Ciotir, Rim Fayad. Nonlinear Fokker-Planck equation with reflecting boundary conditions. 2021. hal-03151286

HAL Id: hal-03151286

<https://hal.science/hal-03151286v1>

Preprint submitted on 24 Feb 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Nonlinear Fokker-Planck equation with reflecting boundary conditions

Ioana CIOTIR^a, Rim FAYAD^a

^a Normandie University, INSA de Rouen Normandie,
LMI (EA 3226 - FR CNRS 3335), 76000 Rouen, France
685 Avenue de l'Université, 76801 St Etienne du Rouvray cedex

February 24, 2021

Abstract

We study the existence and uniqueness of a mild solution to a nonlinear Fokker-Planck equation with reflecting boundary conditions, by using a monotonicity approach. Then we prove that the mild solution is also a distributional one.

Keywords: Fokker-Planck equation, m-accretive operator, mild solution, McKean-Vlasov stochastic differential equation

AMS [2020] Primary 60H30, 60H10, 60G46; secondary 35Q84, 35D99

1 Introduction

In the present work we are interested in the study of a nonlinear Fokker-Planck equation with Neumann boundary conditions of the following type

$$\begin{cases} \frac{\partial y}{\partial t} - \frac{1}{2} \Delta \beta(y) - \operatorname{div}(b(\xi)y) = 0, & [0, T] \times \mathcal{O}, \\ \frac{1}{2} \frac{\partial}{\partial n} \beta(y) + (b(\xi) \cdot n(\xi))y = 0, & [0, T] \times \partial \mathcal{O}, \\ y(0, \xi) = y_0(\xi), & \xi \in \mathcal{O}. \end{cases} \quad (1)$$

Here $\mathcal{O} \subset \mathbb{R}^d$, $d \geq 1$, is assumed to be an open bounded and convex domain with smooth boundary $\partial \mathcal{O}$. We denoted by n the outward normal to $\partial \mathcal{O}$ and by $\frac{\partial}{\partial n}$ the outward normal derivative. The closer of \mathcal{O} is denoted by $\overline{\mathcal{O}} = K$.

Assumptions

We shall assume the following hypotheses on the functions $\beta : \mathbb{R} \rightarrow \mathbb{R}$ and $b : \overline{\mathcal{O}} \rightarrow \mathbb{R}^d$:

- i) $\beta \in C(\mathbb{R})$, $\beta(0) = 0$, and β is increasing maximal monotone such that, for $m \geq 1$ we have

$$|\beta(r)| \leq \alpha_1 |r|^m + \alpha_2, \text{ for all } r \in \mathbb{R},$$

for some constants $\alpha_1 > 0$ and $\alpha_2 \geq 0$.

- ii) j is the potential of β , i.e. $\partial j = \beta$, and we assume that for the same $m \geq 1$ we have

$$j(r) \geq \alpha_3 r^{m+1}, \text{ for all } r \in \mathbb{R},$$

for some positive constant α_3 .

- iii) $b \in C^1(\overline{\mathcal{O}})$ is a given function such that $\operatorname{div} b \leq 0$ in $\overline{\mathcal{O}}$, $b \cdot n \geq 0$ on $\partial\mathcal{O}$ and $|b|_\infty < 1$.

Equation (1), where y is a probability density, is known in the literature as the nonlinear Fokker-Planck equation with reflection. For reader's convenience we shall give some details concerning the connection with a stochastic differential equation.

If we choose the solution $y : [0, T] \times \mathcal{O} \rightarrow \mathbb{R}$ in the class of probability density functions on \mathcal{O} , i.e. such that $y \in L^1(\mathcal{O})$, $y \geq 0$ a.e. on \mathcal{O} and $\int_{\mathcal{O}} y(\xi) d\xi = 1$, then y can be seen as the probability density of the law \mathcal{L}_X , where X is the weak solution (in the stochastic sense) to the following stochastic differential equation with reflection

$$\begin{cases} dX(t) + b(X(t)) dt + N_K(X(t)) dt \ni \sqrt{\frac{\beta(y(t, X(t)))}{y(t, X(t))}} dW_t \\ X(0) = X_0 \end{cases} \quad (2)$$

in a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$ where $(\mathcal{F}_t)_t$ is the normal filtration and W_t , $t \geq 0$, is a Wiener process.

Keeping in mind that $K = \overline{\mathcal{O}}$, we recall that $N_K : K \rightarrow 2^{\mathbb{R}^d}$ is the normal cone to K , i.e.

$$N_K(x) = \{z \in \mathbb{R}^d; (z, x - y)_{\mathbb{R}^d} \geq 0, \quad \forall y \in K\}.$$

Actually, the reflecting boundary condition from equation (1) gives the second drift term $N_K(X)$ from the stochastic equation (2).

Concerning the stochastic differential equation, we recall the definition of the solution.

Definition 1 *We call a strong solution to the stochastic differential equation (2) a pair of continuous \mathbb{R}^d -valued and \mathcal{F}_t -adapted processes $(X(t), \eta(t))$, $t \in [0, T]$ such that*

- 1) $X(t) \in K$, \mathbb{P} -a.s., for all $t \in [0, T]$.

2) $\exists \eta \in BV([0, T]; \mathbb{R}^d)$, \mathbb{P} -a.s., such that

$$X(t) + \int_0^t b(X(s)) ds + \eta(t) = X_0 + \int_0^t \sqrt{\frac{\beta(y(s, X(s)))}{y(s, X(s))}} dW_s$$

$$\int_0^T (X(s) - Z(s)) \cdot d\eta(s) \geq 0, \quad \forall Z \in C([0, T]; K).$$

As usual, $BV([0, T]; \mathbb{R}^d)$ is the space of functions with bounded variation and $\int_0^T (X(s) - Z(s)) \cdot d\eta(s)$ is the Riemann-Stieltjes integral.

Equations as (2) are usually referred as McKean-Vlasov SDEs. For classical results concerning this equation see [12], [16], [17], [19], [21]. Recently there was a large interest concerning this subject in different situations (see e.g. [5], [13], [14], [18]).

One can prove that, if y is a solution to (1), then there is a martingale solution X to (2) which has y as probability density, by using the superposition principle by Trevisan [22] (which is a generalization of the work of Figalli [11]) and by a classical result of Stroock-Varadhan (see [20]). For the general McKean-Vlasov equation see e.g. [6], [7]. For an immediate extension to a stochastic reflection problem as in the present case see [3].

In the present work we are interested to treat the well-posedness of the nonlinear Fokker-Planck equation (1) in $L^1(\mathcal{O})$. This work follows the direction opened by Barbu and Röckner in [5], [6], [7], [8] for the Fokker-Planck equation with Dirichlet boundary conditions under different sets of assumptions. This approach is based on the Crandall and Liggett existence theorem for a nonlinear Cauchy problem in a Banach space.

Concerning the Fokker-Planck equation with Neumann boundary condition we refer to [3] for a controllability problem treating an equation with a particular form of the function β which is Lipschitz.

In the present work we study a general Fokker-Planck equation with Neumann boundary conditions, under general growth conditions on β . The main objective of this paper is to prove existence and uniqueness of the mild solution to equation (1) and then to show that it is also a variational solution.

For reader's convenience we shall first recall the following definitions.

Definition 2 *Let $y_0 \in L^2(\mathcal{O})$. The function y is called a mild solution to equation (1) if $y \in C([0, T]; L^1(\mathcal{O}))$ and*

$$y(t) = \lim_{\delta \rightarrow \infty} y_\delta(t) \text{ in } L^1(\mathcal{O}), \text{ uniformly on } [0, T],$$

where $y_\delta : [0, T] \rightarrow \mathbb{R}$ is a step function such that

$$y_\delta(t) = y_\delta^i, \quad \forall t \in [i\delta, (i+1)\delta), \quad i = 0, 1, \dots, N_\delta - 1,$$

for $N_\delta = \lceil \frac{T}{\delta} \rceil$ which satisfies

$$\frac{1}{\delta} \int_{\mathcal{O}} (y_\delta^{i+1}(\xi) - y_\delta^i(\xi)) \psi(\xi) d\xi + \int_{\mathcal{O}} \left(\frac{1}{2} \nabla \beta(y_\delta^{i+1}) + b(\xi) y_\delta^{i+1} \right) \nabla \psi(\xi) d\xi = 0,$$

for $\forall \psi \in H^1(\mathcal{O})$ such that $y_\delta^0 = y_0$ and $\beta(y_\delta^{i+1}) \in H^1(\mathcal{O})$.

Definition 3 Let $y_0 \in L^2(\mathcal{O})$. We say that y is a distributional solution to (1) if

$$y \in C([0, 1]; L^1(\mathcal{O})), \quad \beta(y) \in L^2(0, T; H_0^1(\mathcal{O}))$$

such that

$$\begin{aligned} \int_0^T \int_{\mathcal{O}} \left(y \frac{\partial \varphi}{\partial t} - \frac{1}{2} \nabla \beta(y) \cdot \nabla \varphi - (b \cdot \nabla \varphi) y \right) d\xi dt \\ + \int_{\mathcal{O}} y_0(\xi) \varphi(0, \xi) d\xi = 0 \end{aligned} \quad (3)$$

for all $\varphi \in C^1([0, T] \times \overline{\mathcal{O}})$, $\varphi(T, \cdot) = 0$.

Notations

For an open subset $\mathcal{O} \subset \mathbb{R}^d$ we denote by $L^p(\mathcal{O})$, $1 \leq p \leq \infty$, the space of standard Lebesgue p -integrable functions on \mathcal{O} , and by $W^{1,p}(\mathcal{O})$ the Sobolev space $\{u \in L^p(\mathcal{O}), D_i u \in L^p(\mathcal{O}) \text{ for } i = 1, 2, \dots, d\}$ where $D_i = \frac{\partial}{\partial x_i}$ is taken in the sense of Schwartz distributions. In particular, we denote by $H^1(\mathcal{O}) = W^{1,2}(\mathcal{O})$.

2 The main results

We can formulate now the main result of this work.

Theorem 4 Assuming the hypotheses above, for each $y \in L^1(\mathcal{O})$, there is a unique mild solution y to equation (1).

Moreover, if $y_0 \geq 0$ a.e. on \mathcal{O} and $\int_{\mathbb{R}} y_0(\xi) d\xi = 1$, then $y(t)$ is a probability density function on \mathcal{O} for $\forall t \in [0, T]$.

Furthermore, y is also a distributional solution to (1).

The main idea is to rewrite equation (1) as a Cauchy problem of the form

$$\begin{cases} dy + A(y) = 0, & (0, T), \\ y(0) = y_0, \end{cases}$$

in the space $L^1(\mathcal{O})$, where $A : D(A) \subset L^1(\mathcal{O}) \rightarrow L^1(\mathcal{O})$ is an appropriate m -accretive operator and to apply the Crandall-Liggett theorem. To this purpose we shall first study this operator in the $L^2(\mathcal{O})$ space framework.

The $L^2(\mathcal{O})$ approach

We define the operator

$$\begin{aligned} A_0 & : D(A_0) \subset L^2(\mathcal{O}) \rightarrow L^2(\mathcal{O}) \\ D(A_0) & = \left\{ u \in L^2(\mathcal{O}); \Delta\beta(u) + \operatorname{div}(bu) \in L^2(\mathcal{O}), \right. \\ & \quad \left. \frac{\partial}{\partial n}\beta(u) + (b \cdot n)u = 0 \text{ a.e. on } \partial\mathcal{O} \right\} \\ A_0(u) & = -\Delta\beta(u) - \operatorname{div}(bu). \end{aligned}$$

For some $f \in L^2(\mathcal{O})$ arbitrarily fixed and $u \in H^1(\mathcal{O})$, if we consider the equation

$$A_0 u = f$$

i.e.

$$\begin{cases} -\frac{1}{2}\Delta\beta(u) - \operatorname{div}(bu) = f, & \text{in } \mathcal{O} \\ \frac{1}{2}\frac{\partial}{\partial n}\beta(u) + (b \cdot n)u = 0, & \text{on } \partial\mathcal{O} \end{cases}$$

we can write its variational formulation as follows

$$\int_{\mathcal{O}} \left(\frac{1}{2}\nabla\beta(u(\xi)) + b(\xi)u(\xi) \right) \nabla\psi(\xi) d\xi = \int_{\mathcal{O}} f(\xi)\psi(\xi) d\xi, \quad \forall \psi \in H^1(\mathcal{O}).$$

Proposition 5 *Assuming the hypotheses above, there is $\lambda_0 > 0$ such that, for all $f \in L^2(\mathcal{O})$ and $\lambda \in (0, \lambda_0)$ arbitrarily fixed, the equation*

$$u + \lambda A_0 u = f \tag{4}$$

has a unique variational solution u .

Furthermore, if we denote by $(I + \lambda A_0)^{-1}(f)$ the solution to equation (4) for each λ and f , we have the following propertie

$$\left| (I + \lambda A_0)^{-1}(f_1) - (I + \lambda A_0)^{-1}(f_2) \right|_{L^1(\mathcal{O})} \leq |f_1 - f_2|_{L^1(\mathcal{O})}, \quad \forall f_1, f_2 \in L^2(\mathcal{O}). \tag{5}$$

In addition, we have that $(I + \lambda A_0)^{-1}(f)$ is a probability density function on \mathcal{O} for all $f \in L^2(\mathcal{O})$ which is also a probability density function.

Proof. We shall first rewrite (4) as

$$\begin{cases} u - \frac{1}{2}\lambda\Delta\beta(u) - \lambda\operatorname{div}(bu) = f, & \text{in } \mathcal{O}, \\ \frac{1}{2}\frac{\partial}{\partial n}\beta(u) + (b \cdot n)u = 0, & \text{on } \partial\mathcal{O}, \end{cases} \tag{6}$$

with the variational formulation

$$\begin{aligned} & \int_{\mathcal{O}} u\psi d\xi + \frac{\lambda}{2} \int_{\mathcal{O}} \nabla\beta(u(\xi)) \nabla\psi d\xi + \lambda \int_{\mathcal{O}} (b(\xi) \cdot \nabla\psi(\xi)) u(\xi) d\xi \quad (7) \\ &= \int_{\mathcal{O}} f(\xi) \psi(\xi) d\xi, \quad \psi \in H^1(\mathcal{O}). \end{aligned}$$

We approximate the operator β by $\tilde{\beta}_\varepsilon(r) = \beta_\varepsilon(r) + \varepsilon r$ where β_ε is the Yosida approximation of β , i.e.,

$$\beta_\varepsilon(r) = \frac{1}{\varepsilon} (r - J_\varepsilon(r)) = \beta(J_\varepsilon(r))$$

for

$$J_\varepsilon(r) = (Id + \varepsilon\beta)^{-1}(r), \quad \forall r \in \mathbb{R}.$$

Since β_ε is Lipschitz and increasing, we can easily see that $\tilde{\beta}_\varepsilon$ is also Lipschitz and strongly monotone, that is

$$\left(\tilde{\beta}_\varepsilon(r) - \tilde{\beta}_\varepsilon(\bar{r})\right) (r - \bar{r}) \geq \varepsilon |r - \bar{r}|^2, \quad \forall r, \bar{r} \in \mathbb{R}.$$

Note that this property implies that $\left(\tilde{\beta}_\varepsilon\right)^{-1}$ is also Lipschitz, where $\left(\tilde{\beta}_\varepsilon\right)^{-1}$ is the inverse of $\tilde{\beta}_\varepsilon$.

We approximate equation (6) by

$$\begin{cases} u^\varepsilon - \frac{1}{2} \lambda \Delta \tilde{\beta}_\varepsilon(u^\varepsilon) - \lambda \operatorname{div}(bu^\varepsilon) = f, & \text{in } \mathcal{O}, \\ \frac{1}{2} \frac{\partial}{\partial n} \tilde{\beta}_\varepsilon(u^\varepsilon) + (b \cdot n) u^\varepsilon = 0, & \text{on } \partial\mathcal{O}, \end{cases} \quad (8)$$

with the variational formulation

$$\begin{aligned} & \int_{\mathcal{O}} u^\varepsilon \psi d\xi + \frac{\lambda}{2} \int_{\mathcal{O}} \nabla \tilde{\beta}_\varepsilon(u^\varepsilon(\xi)) \nabla \psi d\xi + \lambda \int_{\mathcal{O}} (b(\xi) \cdot \nabla \psi(\xi)) u^\varepsilon(\xi) d\xi \quad (9) \\ &= \int_{\mathcal{O}} f(\xi) \psi(\xi) d\xi, \quad \psi \in H^1(\mathcal{O}). \end{aligned}$$

If we denote by $\left(\tilde{\beta}_\varepsilon\right)^{-1}$ the inverse of $\tilde{\beta}_\varepsilon$, we can rewrite (8) as follows

$$\begin{cases} \left(\tilde{\beta}_\varepsilon\right)^{-1}(v^\varepsilon) - \frac{1}{2} \lambda \Delta v^\varepsilon - \lambda \operatorname{div}\left(b \left(\tilde{\beta}_\varepsilon\right)^{-1}(v^\varepsilon)\right) = f, & \text{in } \mathcal{O}, \\ \frac{1}{2} \frac{\partial}{\partial n} v^\varepsilon + (b \cdot n) \left(\tilde{\beta}_\varepsilon\right)^{-1}(v^\varepsilon) = 0, & \text{on } \partial\mathcal{O}. \end{cases} \quad (10)$$

Note that since $\tilde{\beta}_\varepsilon$ is strictly monotone, we can easily see that $\left(\tilde{\beta}_\varepsilon\right)^{-1}$ is Lipschitz on \mathbb{R} .

Let us define the operator

$$B^\varepsilon : H^1(\mathcal{O}) \rightarrow (H^1)^*(\mathcal{O})$$

such that, for each $v \in H^1(\mathcal{O})$ we have

$$\begin{aligned} (H^1)^* \langle B^\varepsilon(v), \varphi \rangle_{H^1} &= \int_{\mathcal{O}} \left(\tilde{\beta}_\varepsilon \right)^{-1}(v) \varphi d\xi \\ &+ \frac{\lambda}{2} \int_{\mathcal{O}} \nabla v \nabla \varphi d\xi \\ &+ \lambda \int_{\mathcal{O}} b(\xi) \cdot \nabla \varphi \left(\tilde{\beta}_\varepsilon \right)^{-1}(v) d\xi, \text{ for } \forall \varphi \in H^1(\mathcal{O}). \end{aligned}$$

We can now rewrite equation (10) as

$$B^\varepsilon(v^\varepsilon) = f. \quad (11)$$

We shall first show that the operator B^ε is strictly monotone. We have

$$\begin{aligned} (H^1)^* \langle B^\varepsilon(v) - B^\varepsilon(\bar{v}), v - \bar{v} \rangle_{H^1} &= \int_{\mathcal{O}} \left(\left(\tilde{\beta}_\varepsilon \right)^{-1}(v) - \left(\tilde{\beta}_\varepsilon \right)^{-1}(\bar{v}) \right) (v - \bar{v}) d\xi \\ &+ \frac{\lambda}{2} \int_{\mathcal{O}} \nabla(v - \bar{v}) d\xi \\ &+ \lambda \int_{\mathcal{O}} b(\xi) \cdot \nabla(v - \bar{v}) \left(\left(\tilde{\beta}_\varepsilon \right)^{-1}(v) - \left(\tilde{\beta}_\varepsilon \right)^{-1}(\bar{v}) \right) d\xi. \end{aligned}$$

Since by assumption i) we have that

$$\left(\left(\tilde{\beta}_\varepsilon \right)^{-1}(v) - \left(\tilde{\beta}_\varepsilon \right)^{-1}(\bar{v}) \right) (v - \bar{v}) \geq \varepsilon \left| \left(\tilde{\beta}_\varepsilon \right)^{-1}(v) - \left(\tilde{\beta}_\varepsilon \right)^{-1}(\bar{v}) \right|^2$$

we obtain

$$\begin{aligned} (H^1)^* \langle B^\varepsilon(v) - B^\varepsilon(\bar{v}), v - \bar{v} \rangle_{H^1} &\geq \frac{\lambda}{4} \int_{\mathcal{O}} |\nabla(v - \bar{v})|^2 d\xi \\ &+ (\varepsilon - |b|_\infty \lambda) \int_{\mathcal{O}} \left| \left(\tilde{\beta}_\varepsilon \right)^{-1}(v) - \left(\tilde{\beta}_\varepsilon \right)^{-1}(\bar{v}) \right|^2 \\ &\geq 0, \text{ for } \lambda < \frac{\varepsilon}{2|b|_\infty} = \lambda_0. \end{aligned}$$

Keeping in mind that $\beta(0) = 0$ we have also that

$$(H^1)^* \langle B^\varepsilon(v), v \rangle_{H^1} \geq \frac{\lambda}{4} \int_{\mathcal{O}} |\nabla(v)|^2 d\xi + (\varepsilon - |b|_\infty \lambda) \int_{\mathcal{O}} \left| \left(\tilde{\beta}_\varepsilon \right)^{-1}(v) \right|^2,$$

for $\lambda < \frac{\varepsilon}{2|b|_\infty} = \lambda_0$. Since $\tilde{\beta}_\varepsilon$ is Lipschitz we have that $\left| \left(\tilde{\beta}_\varepsilon \right)^{-1}(v) \right| \geq C|v|$, for some positive constant $C(\varepsilon)$ and for each ε fixed, and therefore B^ε is coercive from H^1 to $(H^1)^*$.

Keeping in mind that B^ε is also continuous and monotone, we can now use the Minty-Browder theorem and get that B^ε is surjective. Therefore equation (11) has a unique solution $v^\varepsilon \in H^1(\mathcal{O})$ for each ε fixed.

We obtain that $u^\varepsilon = \tilde{\beta}_\varepsilon(v^\varepsilon) \in L^2(\mathcal{O})$ is a solution to (6).

We shall now prove that

$$|u^\varepsilon(\lambda, f_1) - u^\varepsilon(\lambda, f_2)|_{L^1(\mathcal{O})} \leq |f_1 - f_2|_{L^1(\mathcal{O})}, \quad \forall f_1, f_2 \in L^2(\mathcal{O}),$$

by using a similar approach with the one suggested in [4] [5] and [6]. To this purpose we define the function

$$\chi_\alpha(r) = \begin{cases} 1, & r > \alpha, \\ \frac{r}{\alpha}, & |r| \leq \alpha, \\ -1, & r < -\alpha, \end{cases}$$

for $\alpha > 0$.

On the other hand, from the variational form of the solution, we have for $u_i^\varepsilon = u^\varepsilon(\lambda, f_i)$ where $f_i \in L^2(\mathcal{O})$, $i = \overline{1, 2}$, that

$$\begin{aligned} & \int_{\mathcal{O}} (u_1^\varepsilon - u_2^\varepsilon) \psi d\xi + \frac{\lambda}{2} \int_{\mathcal{O}} \left(\nabla \tilde{\beta}_\varepsilon(u_1^\varepsilon) - \nabla \tilde{\beta}_\varepsilon(u_2^\varepsilon) \right) \nabla \psi d\xi \\ & + \lambda \int_{\mathcal{O}} b(\xi) \cdot \nabla \psi (u_1^\varepsilon - u_2^\varepsilon) d\xi \\ & = \int_{\mathcal{O}} (f_1 - f_2) \psi d\xi, \quad \psi \in H^1(\mathcal{O}). \end{aligned}$$

If we take $\psi = \chi_\alpha \left(\tilde{\beta}_\varepsilon(u_1^\varepsilon) - \tilde{\beta}_\varepsilon(u_2^\varepsilon) \right)$, in the previous relation, we get

$$\begin{aligned} & \int_{\mathcal{O}} (u_1^\varepsilon - u_2^\varepsilon) \chi_\alpha \left(\tilde{\beta}_\varepsilon(u_1^\varepsilon) - \tilde{\beta}_\varepsilon(u_2^\varepsilon) \right) d\xi \\ & + \frac{\lambda}{2} \int_{\mathcal{O}} \left| \nabla \tilde{\beta}_\varepsilon(u_1^\varepsilon) - \nabla \tilde{\beta}_\varepsilon(u_2^\varepsilon) \right|^2 \chi'_\alpha d\xi \\ & + \lambda \int_{\mathcal{O}} b(\xi) \cdot \nabla \left(\tilde{\beta}_\varepsilon(u_1^\varepsilon) - \tilde{\beta}_\varepsilon(u_2^\varepsilon) \right) \chi'_\alpha (u_1^\varepsilon - u_2^\varepsilon) d\xi \\ & \leq \int_{\mathcal{O}} |f_1 - f_2| d\xi, \end{aligned}$$

where we denoted by $\chi'_\alpha = \chi'_\alpha \left(\tilde{\beta}_\varepsilon(u_1^\varepsilon) - \tilde{\beta}_\varepsilon(u_2^\varepsilon) \right)$.

Since $\chi_\alpha < 1$, $\chi'_\alpha > 0$ and keeping in mind the explicit form of χ_α we have

$$\begin{aligned} & \int_{\mathcal{O}} (u_1^\varepsilon - u_2^\varepsilon) \chi_\alpha \left(\tilde{\beta}_\varepsilon(u_1^\varepsilon) - \tilde{\beta}_\varepsilon(u_2^\varepsilon) \right) d\xi \\ & \leq \frac{\lambda}{\alpha} |b|_\infty \int_{[\tilde{\beta}_\varepsilon(u_1) - \tilde{\beta}_\varepsilon(u_2) < \alpha]} \left| \nabla \left(\tilde{\beta}_\varepsilon(u_1^\varepsilon) - \tilde{\beta}_\varepsilon(u_2^\varepsilon) \right) \right| |u_1^\varepsilon - u_2^\varepsilon| d\xi \\ & \quad + \int_{\mathcal{O}} |f_1 - f_2| d\xi. \end{aligned} \tag{12}$$

We calculate the first term from the right hand side of the previous relation.

Since $\tilde{\beta}_\varepsilon$ is strongly increasing and therefore $\left(\tilde{\beta}_\varepsilon \right)^{-1}$ is Lipschitz, we have that

$$|u_1^\varepsilon - u_2^\varepsilon| \leq \frac{1}{\varepsilon} \left| \tilde{\beta}_\varepsilon(u_1^\varepsilon) - \tilde{\beta}_\varepsilon(u_2^\varepsilon) \right|,$$

and therefore

$$\begin{aligned} & \lim_{\alpha \rightarrow 0} \frac{\lambda}{\alpha} |b|_\infty \int_{[\tilde{\beta}_\varepsilon(u_1) - \tilde{\beta}_\varepsilon(u_2) < \alpha]} \left| \nabla \left(\tilde{\beta}_\varepsilon(u_1^\varepsilon) - \tilde{\beta}_\varepsilon(u_2^\varepsilon) \right) \right| |u_1^\varepsilon - u_2^\varepsilon| d\xi \\ & \leq \lim_{\alpha \rightarrow 0} \frac{\lambda}{\varepsilon} |b|_\infty \int_{[\tilde{\beta}_\varepsilon(u_1) - \tilde{\beta}_\varepsilon(u_2) < \alpha]} \left| \nabla \left(\tilde{\beta}_\varepsilon(u_1^\varepsilon) - \tilde{\beta}_\varepsilon(u_2^\varepsilon) \right) \right| d\xi = 0. \end{aligned}$$

Going back to (12) and considering that $\lim_{\alpha \rightarrow 0} \chi_\alpha = \text{sgn}$ and

$$\text{sgn} \left(\tilde{\beta}_\varepsilon(u_1^\varepsilon) - \tilde{\beta}_\varepsilon(u_2^\varepsilon) \right) = \text{sgn}(u_1^\varepsilon - u_2^\varepsilon)$$

we obtain that

$$\lim_{\alpha \rightarrow 0} \int_{\mathcal{O}} (u_1^\varepsilon - u_2^\varepsilon) \chi_\alpha \left(\tilde{\beta}_\varepsilon(u_1^\varepsilon) - \tilde{\beta}_\varepsilon(u_2^\varepsilon) \right) d\xi = |u_1^\varepsilon - u_2^\varepsilon|_{L^1(\mathcal{O})}.$$

This leads to

$$|u_1^\varepsilon - u_2^\varepsilon|_{L^1(\mathcal{O})} \leq |f_1 - f_2|_{L^1(\mathcal{O})}, \quad \forall f_1, f_2 \in L^2(\mathcal{O}). \tag{13}$$

We shall now pass to the limit for $\varepsilon \rightarrow 0$ in the approximating equation (8) with the variational formulation (9). To this purpose we need the following estimates.

First we take $\psi = u^\varepsilon$ and we get

$$\int_{\mathcal{O}} |u^\varepsilon|^2 d\xi + \frac{1}{2} \lambda \int_{\mathcal{O}} \nabla \tilde{\beta}_\varepsilon(u^\varepsilon) \nabla u^\varepsilon d\xi + \lambda \int_{\mathcal{O}} (b \cdot \nabla u^\varepsilon) u^\varepsilon d\xi = \int_{\mathcal{O}} f u^\varepsilon d\xi.$$

Since, by assumption *iii*) we have that

$$\frac{1}{2} \int_{\mathcal{O}} b \cdot \nabla (u^\varepsilon)^2 d\xi = -\frac{1}{2} \int_{\mathcal{O}} \text{div} b (u^\varepsilon)^2 d\xi + \frac{1}{2} \int_{\partial \mathcal{O}} (b \cdot n) (u^\varepsilon)^2 d\xi \geq 0$$

we get $\{u^\varepsilon\}$ is bounded in $L^2(\mathcal{O})$.

Then we take $\psi = \tilde{\beta}_\varepsilon(u^\varepsilon)$ and we get

$$\begin{aligned} & \int_{\mathcal{O}} u^\varepsilon \tilde{\beta}_\varepsilon(u^\varepsilon) d\xi + \frac{1}{2} \lambda \int_{\mathcal{O}} \left| \nabla \tilde{\beta}_\varepsilon(u^\varepsilon) \right|^2 d\xi + \lambda \int_{\mathcal{O}} \left(b \cdot \nabla \tilde{\beta}_\varepsilon(u^\varepsilon) \right) u^\varepsilon d\xi \quad (14) \\ &= \int_{\mathcal{O}} f \tilde{\beta}_\varepsilon(u^\varepsilon) d\xi. \end{aligned}$$

We calculate

$$\begin{aligned} \int_{\mathcal{O}} u^\varepsilon \tilde{\beta}_\varepsilon(u^\varepsilon) d\xi &= \int_{\mathcal{O}} u^\varepsilon \beta_\varepsilon(u^\varepsilon) d\xi + \int_{\mathcal{O}} |u^\varepsilon|^2 d\xi \\ &\geq \int_{\mathcal{O}} (u^\varepsilon - J_\varepsilon(u^\varepsilon)) \beta(J_\varepsilon(u^\varepsilon)) d\xi + \int_{\mathcal{O}} J_\varepsilon(u^\varepsilon) \beta(J_\varepsilon(u^\varepsilon)) d\xi \\ &\geq \alpha_3 \int_{\mathcal{O}} |J_\varepsilon(u^\varepsilon)|^{m+1} d\xi \end{aligned}$$

and

$$\lambda \int_{\mathcal{O}} \left(b \cdot \nabla \tilde{\beta}_\varepsilon(u^\varepsilon) \right) u^\varepsilon d\xi \leq \frac{1}{4} \lambda \int_{\mathcal{O}} \left| \nabla \tilde{\beta}_\varepsilon(u^\varepsilon) \right|^2 d\xi + C \int_{\mathcal{O}} |u^\varepsilon|^2 d\xi.$$

By going back to (14) we get

$$\begin{aligned} & \alpha_3 \int_{\mathcal{O}} |J_\varepsilon(u^\varepsilon)|^{m+1} d\xi + \frac{1}{4} \lambda \int_{\mathcal{O}} \left| \nabla \tilde{\beta}_\varepsilon(u^\varepsilon) \right|^2 d\xi \quad (15) \\ & \leq C + \frac{\rho}{2} \int_{\mathcal{O}} |f|^{m+1} d\xi + \frac{1}{2\rho} \int_{\mathcal{O}} \left| \tilde{\beta}_\varepsilon(u^\varepsilon) \right|^{\frac{m+1}{m}} d\xi. \end{aligned}$$

Since

$$\tilde{\beta}_\varepsilon(u^\varepsilon) = \beta_\varepsilon(u^\varepsilon) + \varepsilon u^\varepsilon$$

we have

$$\int_{\mathcal{O}} \left| \tilde{\beta}_\varepsilon(u^\varepsilon) \right|^{\frac{m+1}{m}} d\xi \leq 2^{\frac{1}{m}} |\beta(J_\varepsilon(u^\varepsilon))|^{\frac{m+1}{m}} + 2^{\frac{1}{m}} \varepsilon^{\frac{m+1}{m}} |u^\varepsilon|^{\frac{m+1}{m}}.$$

On the other hand, since $|\beta(r)| \leq \alpha_1 |r|^m + \alpha_2, \forall r \in \mathbb{R}$, we have that

$$|\beta(J_\varepsilon(u^\varepsilon))|^{\frac{m+1}{m}} = \int_{\mathcal{O}} |\beta(J_\varepsilon(u^\varepsilon))|^{\frac{m+1}{m}} d\xi \leq C \left(\int_{\mathcal{O}} |J_\varepsilon(u^\varepsilon)|^{m+1} d\xi + 1 \right).$$

By replacing in (15) we get that, for a constant ρ sufficiently large, that

$$\frac{\alpha_3}{2} \int_{\mathcal{O}} |J_\varepsilon(u^\varepsilon)|^{m+1} d\xi + \frac{1}{4} \lambda \int_{\mathcal{O}} \left| \nabla \tilde{\beta}_\varepsilon(u^\varepsilon) \right|^2 d\xi \leq C.$$

From the previous estimates, it follows that

$$\{u_\varepsilon\} \text{ bounded in } L^2(\mathcal{O})$$

$$\begin{aligned} \{J_\varepsilon(u^\varepsilon)\} &\text{ bounded in } L^{m+1}(\mathcal{O}) \\ \{\beta_\varepsilon(u^\varepsilon)\} &\text{ bounded in } L^{\frac{m+1}{m}}(\mathcal{O}) \\ \{\nabla\beta_\varepsilon(u^\varepsilon)\} &\text{ bounded in } L^2(\mathcal{O}) \subset L^{\frac{m+1}{m}}(\mathcal{O}). \end{aligned}$$

Since the Sobolev space $W^{1, \frac{m+1}{m}}(\mathcal{O})$ is compactly embedded in $L^{\frac{m+1}{m}}(\mathcal{O})$ we have that

$$\beta_\varepsilon(u^\varepsilon) \rightarrow \eta \text{ strongly in } L^{\frac{m+1}{m}}(\mathcal{O})$$

and therefore $\beta(u) = \eta$.

We can get (5) by passing to the limit in (13).

In order to conclude the proof of the proposition, it only remain to show that u is a probability density function on \mathcal{O} for each $f \in L^2(\mathcal{O})$ which is a probability density function.

We assume that f is a probability density, i.e. $f \geq 0$ a.e. in \mathcal{O} and $\int_{\mathcal{O}} f(\xi) d\xi = 1$.

Let us first take $\psi = u^-$ in (7). We obtain

$$\begin{aligned} & - \int_{[u<0]} |u|^2 d\xi - \frac{\lambda}{2} \int_{[u<0]} \nabla\beta(u) \nabla u d\xi + \lambda \int_{\mathcal{O}} b \cdot \nabla u^- u d\xi \quad (16) \\ & = - \int_{[u<0]} f u d\xi. \end{aligned}$$

We calculate

$$\begin{aligned} & \lambda \int_{\mathcal{O}} b \cdot \nabla u^- u d\xi \\ & = -\lambda \int_{\mathcal{O}} b \cdot \nabla u^- u^- d\xi \\ & = -\frac{\lambda}{2} \int_{\mathcal{O}} b \cdot \nabla (u^-)^2 d\xi \\ & = \frac{\lambda}{2} \int_{\mathcal{O}} \operatorname{div} b (u^-)^2 d\xi - \frac{\lambda}{2} \int_{\partial\mathcal{O}} b \cdot n (u^-)^2 d\xi \\ & = \frac{\lambda}{2} \int_{[u<0]} \operatorname{div} b (u^-)^2 d\xi - \frac{\lambda}{2} \int_{\partial\mathcal{O}} (b \cdot n) (u^-)^2 d\xi < 0. \end{aligned}$$

Keeping in mind the assumptions on b , we obtain that $u^- = 0$ a.e. on \mathcal{O} since the left hand side of (16) is negative and $-\int_{[u<0]} f u d\xi > 0$. It is then immediate that

$$\int_{\mathcal{O}} u(\xi) d\xi = \int_{\mathcal{O}} f(\xi) d\xi = 1$$

and we can conclude that u is a probability density function. The proof of the proposition is now complete. ■

The $L^1(\mathcal{O})$ approach

Let us first consider the operator $\bar{A}: D(\bar{A}) \subset L^1(\mathcal{O}) \rightarrow L^1(\mathcal{O})$ defined by

$$\bar{A}(u) = -\Delta\beta(u) + \operatorname{div}(bu)$$

where

$$D(\bar{A}) = \left\{ u \in L^1(\mathcal{O}); \Delta\beta(u) + \operatorname{div}(bu) \in L^1(\mathcal{O}), \right. \\ \left. \frac{\partial}{\partial n}\beta(u) + (b \cdot n)u = 0 \text{ a.e. on } \partial\mathcal{O} \right\},$$

where Δ is taken in the sense of distributions on \mathcal{O} .

We can prove now the following result.

Proposition 6 *For all $f \in L^1(\mathcal{O})$ and $\forall \lambda \in (0, \lambda_0)$, the equation*

$$u + \lambda\bar{A}u = f$$

has a solution $u = (I + \lambda\bar{A})^{-1}f$.

Moreover

$$\left| (I + \lambda\bar{A})^{-1}f_1 - (I + \lambda\bar{A})^{-1}f_2 \right|_{L^1(\mathcal{O})} \leq |f_1 - f_2|_{L^1(\mathcal{O})}, \quad (17)$$

$\forall f_1, f_2 \in L^1(\mathcal{O})$.

Proof. For $f \in L^1(\mathcal{O})$ arbitrarily fixed we have a sequence $\{f_n\} \subset L^2(\mathcal{O})$ such that $f_n \rightarrow f$ strongly in $L^1(\mathcal{O})$ for $n \rightarrow \infty$. Since we have existence of a solution for the following equation

$$u_n + \lambda\bar{A}u_n = f_n,$$

we can rewrite it as

$$u_n = (I + \lambda\bar{A})^{-1}f_n.$$

Using Proposition 5 we have that

$$\left| (I + \lambda A_0)^{-1}f_1 - (I + \lambda A_0)^{-1}f_2 \right|_{L^1(\mathcal{O})} \leq |f_1 - f_2|_{L^1(\mathcal{O})},$$

for all $f_1, f_2 \in L^2(\mathcal{O})$ which implies directly that

$$|u_n - u_m|_{L^1(\mathcal{O})} \leq |f_n - f_m|_{L^1(\mathcal{O})}, \quad \forall n, m \in \mathbb{N},$$

and then $(u_n)_n$ converges strongly in $L^1(\mathcal{O})$ to some $u \in L^1(\mathcal{O})$. Since $(I + \lambda A_0)^{-1}(f) = (I + \lambda\bar{A})^{-1}(f)$, for $f \in L^2(\mathcal{O})$ we get that $\bar{A}u_n \rightarrow \bar{A}u$ in $L^1(\mathcal{O})$ and

$$u + \lambda\bar{A}u = f.$$

Hence the operator \bar{A} is closed in $L^1 \times L^1$. The estimate follows by the previous result. ■

We can finally define the operator

$$A : D(A) \subset L^1(\mathcal{O}) \rightarrow L^1(\mathcal{O})$$

$$Ay = \bar{A}y, \quad \forall y \in D(A) = (I + \lambda \bar{A})^{-1}(L^1(\mathcal{O}))$$

for same $\lambda > 0$ (equivalently, for each $\lambda > 0$).

Concerning the operator above we have the following result.

Proposition 7 *The operator A is m -accretive in $L^1(\mathcal{O})$, that is, $R(I + \lambda A) = L^1(\mathcal{O})$, for $\forall \lambda > 0$ and*

$$\left| (I + \lambda A)^{-1} f_1 - (I + \lambda A)^{-1} f_2 \right|_{L^1(\mathcal{O})} \leq |f_1 - f_2|_{L^1(\mathcal{O})},$$

for $\forall f_1, f_2 \in L^1(\mathcal{O})$, $\lambda > 0$.

The proof of this result follows directly from the previous one by the definition of the operator A .

Proof of the Theorem 4. In order to use the Crandall-Liggett theorem for the Cauchy problem

$$\begin{cases} \frac{dy}{dt} + A(y) = 0, & t \in (0, T), \\ y(0) = y_0, \end{cases}$$

where $y_0 \in L^1(\mathcal{O})$, we need to have $\overline{D(A)} = L^1(\mathcal{O})$, but this follows by a similar argument as the one used in Theorem 2 from [4].

For reader's convenience we shall sketch it here. We first consider for each $\varepsilon > 0$ the approximating operator

$$A_\varepsilon : D(A_\varepsilon) \subset L^2(\mathcal{O}) \rightarrow L^2(\mathcal{O})$$

defined by

$$A_\varepsilon(u) = -\Delta \tilde{\beta}_\varepsilon(u) - \operatorname{div}(b \cdot u), \quad \text{for } \forall u \in D(A_\varepsilon)$$

where

$$D(A_\varepsilon) = \left\{ u \in H^1(\mathcal{O}); \tilde{\beta}_\varepsilon(u) \in H^2(\mathcal{O}), \right. \\ \left. \frac{\partial}{\partial n} \tilde{\beta}_\varepsilon(u) + (b \cdot n)u = 0, \text{ a.e. on } \partial\mathcal{O} \right\}.$$

By Proposition 5 we have that

$$|u_\varepsilon^1 - u_\varepsilon^2|_{L^1(\mathcal{O})} \leq |f_1 - f_2|_{L^1(\mathcal{O})}, \quad \text{for } \forall f_1, f_2 \in L^2(\mathcal{O}),$$

where u_ε^i is the solution of $u_\varepsilon^i + \lambda A_\varepsilon (u_\varepsilon^i) = f_i$, $i = \overline{1, 2}$ and also that

$$\lim_{\varepsilon \rightarrow 0} u_\varepsilon = u = J_\lambda (f) \text{ in } L^1(\mathcal{O}), \quad \forall f \in L^2(\mathcal{O}), \quad \lambda \in (0, \lambda_0).$$

Since $C_0^\infty(\mathbb{R}^d) \subset D(A_\varepsilon)$ it follows that $D(A_\varepsilon)$ is dense in $L^1(\mathcal{O})$ and therefore so is $D(A)$.

We get that, for each $y_0 \in L^1(\mathcal{O}) = \overline{D(A)}$, the problem above has a unique mild solution $y \in C([0, T]; L^1(\mathcal{O}))$. Moreover,

$$y(t) = \lim_{n \rightarrow \infty} \left(I + \frac{t}{n} A \right)^{-n} y_0, \quad \text{in } L^1(\mathcal{O}),$$

uniformly in t on compact intervals (see e.g. [1]).

Keeping in mind the definition of a mild solution, this means that

$$y(t) = \lim_{\delta \rightarrow 0} y_\delta(t), \quad \text{in } L^1(\mathcal{O})$$

uniformly on $[0, T]$, where

$$y_\delta(t) = y_\delta^{i+1}, \quad \forall t \in [i\delta, (i+1)\delta),$$

where $i = 0, 1, \dots, N_\delta - 1$ for $N_\delta = \lceil \frac{T}{\delta} \rceil$, and

$$\begin{cases} y_\delta^{i+1} + \delta A(y_\delta^{i+1}) = y_\delta^i, & i = 0, 1, \dots, N_\delta - 1 \\ y_\delta^0 = y_0. \end{cases} \quad (18)$$

Moreover $y = y(t, y_0)$ generates a semigroup of contraction in $L^1(\mathcal{O})$, i.e.

$$|y(t, y_0^1) - y(t, y_0^2)|_{L^1(\mathcal{O})} \leq |y_0^1 - y_0^2|_{L^1(\mathcal{O})}, \quad \forall y_0^1, y_0^2 \in L^1(\mathcal{O}).$$

If we take $y_0 \in L^2(\mathcal{O})$ and keep in mind that by Proposition 6 we have

$$(I + \lambda \bar{A})^{-1} f = (I + \lambda A)^{-1} f, \quad \forall f \in L^2(\mathcal{O}),$$

it follows that $y_\delta^{i+1} \in D(A)$ and

$$\frac{1}{\delta} (y_\delta^{i+1} - y_\delta^i) = -A y_\delta^{i+1}, \quad \forall i = 0, 1, \dots, N_\delta - 1.$$

This means that y is the mild solution to equation (1) in the sense of Definition 2. By the accretivity of the operator A in $L^1(\mathcal{O})$ it is easily seen that we have also uniqueness of the solution.

In order to conclude the proof of the theorem we only need to show that y is also a distributional solution to equation (1).

From the definition of the mild solution we have that

$$\frac{1}{\delta} \int_{\mathcal{O}} (y_\delta^{i+1} - y_\delta^i) \beta(y_\delta^{i+1}) d\xi + \frac{1}{2} \int_{\mathcal{O}} |\nabla \beta(y_\delta^{i+1})|^2 d\xi$$

$$+ \int_{\mathcal{O}} b(\xi) \cdot \nabla \beta (y_{\delta}^{i+1}) y_{\delta}^{i+1} d\xi = 0,$$

and keeping in mind that $\partial j = \beta$ we can compute

$$\begin{aligned} & \int_{\mathcal{O}} j (y_{\delta}^{i+1}) d\xi - \int_{\mathcal{O}} j (y_{\delta}^i) d\xi + \frac{\delta}{2} \int_{\mathcal{O}} |\nabla \beta (y_{\delta}^{i+1})|^2 d\xi \\ & + \delta \int_{\mathcal{O}} b(\xi) \cdot \nabla \beta (y_{\delta}^{i+1}) y_{\delta}^{i+1} d\xi = 0. \end{aligned}$$

If we take the sum for $i = \overline{1, k}$ we get

$$\begin{aligned} & \int_{\mathcal{O}} j (y_{\delta}^{k+1}) d\xi - \int_{\mathcal{O}} j (y_{\delta}^0) d\xi + \frac{\delta}{2} \sum_{i=0}^k |\nabla \beta (y_{\delta}^{i+1})|_2^2 \\ & + \delta \sum_{i=0}^k \int_{\mathcal{O}} b(\xi) \cdot \nabla \beta (y_{\delta}^{i+1}) y_{\delta}^{i+1} d\xi = 0, \end{aligned}$$

which leads to

$$\begin{aligned} & \int_{\mathcal{O}} j (y_{\delta}^{k+1}) d\xi + \frac{\delta}{2} \sum_{i=0}^k |\nabla \beta (y_{\delta}^{i+1})|_2^2 \\ & \leq \int_{\mathcal{O}} j (y_{\delta}^0) d\xi + \delta \sum_{i=0}^k |b|_{\infty} |y_{\delta}^{i+1}|_2 |\nabla \beta (y_{\delta}^{i+1})|_2, \end{aligned}$$

for $\forall k = 0, 1, \dots, N_{\delta} - 1$.

We obtain that

$$\int_{\mathcal{O}} j (y_{\delta}^{k+1}) d\xi + \frac{\delta}{4} \sum_{i=0}^k |\nabla \beta (y_{\delta}^{i+1})|_2^2 \leq \int_{\mathcal{O}} j (y_{\delta}^0) d\xi + \delta |b|_{\infty} \sum_{i=0}^k |y_{\delta}^{i+1}|_2^2.$$

Since $y_{\delta}(t) = y_{\delta}^i$ for $t \in [i\delta, (i+1)\delta)$ we get by the previous estimate that

$$\begin{aligned} & \int_{\mathcal{O}} j (y_{\delta}(t)) d\xi + \frac{1}{4} \int_0^t |\nabla \beta (y_{\delta}(s))|_2^2 ds \\ & \leq \int_{\mathcal{O}} j (y_0) d\xi + |b|_{\infty} \int_0^t |y_{\delta}(s)|_2^2 ds, \quad \forall t \in [0, T]. \end{aligned}$$

From the assumptions we have on j we get that

$$\begin{aligned} & \sup_{t \in [0, T]} |y_{\delta}(t)|_2^2 + \frac{1}{4} \int_0^T |\nabla \beta (y_{\delta}(s))|_2^2 ds \\ & \leq \int_{\mathcal{O}} j (y_0) d\xi + |b|_{\infty}^2 T \sup_{t \in [0, T]} |y_{\delta}(s)|_2^2 ds, \end{aligned}$$

and therefore

$$\left(1 - |b|_\infty^2\right) \sup_{t \in [0, T]} \|y_\delta(t)\|_2^2 + \frac{1}{4} \int_0^T \|\nabla \beta(y_\delta(s))\|_2^2 ds \leq \int_{\mathcal{O}} j(y_0) d\xi$$

We know that

$$\begin{aligned} y_\delta &\rightarrow y \text{ in } C([0, T]; L^1(\mathcal{O})) \\ y_\delta &\rightharpoonup y \text{ weakly in } L^\infty(0, T; L^2(\mathcal{O})) \\ \beta(y_\delta) &\rightharpoonup \eta \text{ weakly in } L^2(0, T; H^1(\mathcal{O})) \\ \beta &\text{ maximal monotone in } L^2(\mathcal{O}) \times L^2(\mathcal{O}). \end{aligned}$$

Since β is continuous and

$$y_\delta(t, \xi) \rightarrow y(t, \xi) \text{ a.e. on } (0, T) \times \mathcal{O}$$

it follows that

$$\eta(t, \xi) = \lim_{\delta \rightarrow 0} \beta(y_\delta(t, \xi)) \text{ a.e. on } (0, T) \times \mathcal{O}.$$

We can now conclude the proof that y is also a distributional solution. We shall use the previous estimate to pass to the limit in the mild solution in order to obtain that it is also a distributional solution. More precisely, if we take

$$\frac{1}{\delta} \int_{\mathcal{O}} (y_\delta^{i+1} - y_\delta^i) \psi d\xi + \int_{\mathcal{O}} \left(\frac{1}{2} \nabla \beta(y_\delta^{i+1}) + b(\xi) y_\delta^{i+1} \right) \cdot \nabla \psi d\xi = 0$$

and keep in mind that

$$y_\delta(t) = y_\delta^i, \quad t \in [i\delta, (i+1)\delta)$$

we obtain, for all $\psi \in C^1([0, T] \times \overline{\mathcal{O}})$ such that $\psi(T, \cdot) \equiv 0$, that

$$\begin{aligned} &\frac{1}{\delta} \int_{\mathcal{O}} \int_0^{T-\delta} (y_\delta(t+\delta) - y_\delta(t)) \psi(t, \xi) d\xi \\ &+ \int_{\mathcal{O}} \int_0^{T-\delta} \left(\frac{1}{2} \nabla \beta(y_\delta(t+\delta)) + b(\xi) y_\delta(t+\delta) \right) \cdot \nabla \psi(t, \xi) d\xi = 0. \end{aligned}$$

After some elementary calculus we get for the first term that

$$\begin{aligned} &\frac{1}{\delta} \int_{\mathcal{O}} \int_0^{T-\delta} y_\delta(t+\delta) \psi(t, \xi) d\xi - \frac{1}{\delta} \int_{\mathcal{O}} \int_0^{T-\delta} y_\delta(t) \psi(t, \xi) d\xi \\ &= -\frac{1}{\delta} \int_{\mathcal{O}} \int_0^{T-\delta} y_\delta(t) (\psi(t, \xi) - \psi(t-\delta, \xi)) d\xi \\ &\quad - \frac{1}{\delta} \int_{\mathcal{O}} \int_0^\delta y_\delta(t) \psi(t, \xi) d\xi \\ &\quad + \frac{1}{\delta} \int_{\mathcal{O}} \int_{T-\delta}^T y_\delta(t) \psi(t, \xi) d\xi. \end{aligned}$$

By going back to the previous relation and replacing this expression we get that

$$\begin{aligned} & -\frac{1}{\delta} \int_{\mathcal{O}} \int_0^{T-\delta} y_\delta(t) (\psi(t, \xi) - \psi(t - \delta, \xi)) d\xi \\ & -\frac{1}{\delta} \int_{\mathcal{O}} \int_0^\delta y_\delta(t) \psi(t, \xi) d\xi + \frac{1}{\delta} \int_{\mathcal{O}} \int_{T-\delta}^T y_\delta(t) \psi(t, \xi) d\xi \\ & + \int_{\mathcal{O}} \int_0^{T-\delta} \left(\frac{1}{2} \nabla \beta(y_\delta(t + \delta)) + b(\xi) y_\delta(t + \delta) \right) \cdot \nabla \psi(t, \xi) d\xi = 0 \end{aligned}$$

and finally, by letting $\delta \rightarrow 0$, it follows that y is a variational solution in the sense of Definition 3. ■

Acknowledgement 8 *The first author would like to thank professor Viorel Barbu for helpful discussions and constructive suggestions.*

This work was partially supported by the EU with the European regional development fund (ERDF, HN0002137 and 18P03390/18E01750/18P02733), by the Normandie Regional Council (via the M2NUM and M2SiNum projects) and by the French ANR grant ANR-18-CE46-0013 QUTE-HPC.

References

- [1] V. Barbu, *Nonlinear Differential Equations of Monotone Types in Banach Spaces*, Springer, New York, 2010.
- [2] V. Barbu, *Analysis and Control of Nonlinear Infinite Dimensional Systems*, Volume 190 1st Edition, Academic Press, 1992.
- [3] V. Barbu, *The controllability of Fokker–Planck equations with reflecting boundary conditions and controllers in diffusion term* working document
- [4] V. Barbu, I. Ciotir, I. Danaïla, *Existence and uniqueness of solution to the two-phase Stefan problem with convection* submitted preprint
- [5] V. Barbu, M. Röckner, *Nonlinear Fokker-Planck equations driven by Gaussian linear multiplicative noise*, J. Diff. Equations 265 (2018), no. 10, 4993-5030.
- [6] V. Barbu, M. Röckner, *The evolution to equilibrium of solutions to nonlinear Fokker-Planck equation*, arXiv:1904.08291, submitted
- [7] V. Barbu, M. Röckner, *From nonlinear Fokker-Planck equations to solutions of distribution dependent SDE*, arXiv:1808.10706V2, to appear in Ann. Prob.

- [8] V. Barbu, M. Röckner, *Solutions for nonlinear Fokker-Planck equations with measures as initial data and McKean-Vlasov equations*, arXiv:2005.02311, submitted
- [9] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, 2011
- [10] M.G. Crandall, *Generation of semigroups of nonlinear transformations in general*, Banach spaces, Amer. J. Math., 93 (1971), 265–298.
- [11] A. Figalli, *Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients*, J. Funct. Anal., 254 (1) (2008), 109–153.
- [12] T. Funaki, *A certain class of diffusion processes associated with nonlinear parabolic equations*, Z. Wahrsch. Verw. Gebiete, 67 (3) (1984), 331-348.
- [13] W. Hammersley, D. Siska, L. Szpruch, *McKean-Vlasov SDEs under measure independent Lyapunov conditions*, arXiv:1802.03974, 2018.
- [14] X. Huang, M. Röckner, F.-Y. Wang, *Nonlinear Fokker-Planck equations for probability measures on path space and path-distribution dependent SDEs*, arXiv:1709.00556, 2017.
- [15] J.-L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod, 2002
- [16] H.P. McKean, Jr., *A class of Markov processes associated with nonlinear parabolic equations*, Proc. Nat. Acad. Sci. U.S.A., 56 (1966), 1907-1911.
- [17] H.P. McKean, Jr., *Propagation of chaos for a class of non-linear parabolic equations*. In Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967), 41-57.
- [18] Yu.S. Mishura, A.Yu. Veretennikov, *Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations*, arXiv:1603.02212, 2016.
- [19] M. Scheutzow, *Uniqueness and nonuniqueness of solutions of Vlasov McKean equations*, J. Austral. Math. Soc. Ser. A, 43 (2) (1987), 246-256.
- [20] D.W. Stroock, S.R. Srinivasa Varadhan, *Multidimensional diffusion processes*, Springer, Berlin. New York, 1997.
- [21] A.-S. Sznitman, *Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated*, J. Funct. Anal., 56 (3), (1984), 311-336.
- [22] D. Trevisan, *Well-posedness of multidimensional diffusion processes*, Electronic J. Probab., 21 (2016), 1-41.