Existence and uniqueness of nonnegative solutions to the stochastic porous media equation

Viorel Barbu [*],
University Al. I. Cuza
and
Institute of Mathematics “Octav Mayer”, Iasi, Romania ,
Giuseppe Da Prato [†],
Scuola Normale Superiore di Pisa, Italy
and
Michael Röckner [‡]
Faculty of Mathematics, University of Bielefeld, Germany
and
Department of Mathematics and Statistics, Purdue University, U. S. A.

May 30, 2007

Abstract. It is proved that the stochastic porous media equation in a bounded domain of $\mathbb{R}^3$, with multiplicative noise, with a monotone nonlinearity of polynomial growth has a unique nonnegative solution in H^{-1} (in particular is nonnegative measure-valued), provided the initial data is in H^{-1} and nonnegative.

AMS subject Classification 2000: 76S05, 60H15.

Key words: Porous media equation, Stochastic PDEs, Yosida approximation.

[*]Supported by the CEEX Project 05 of Romanian Minister of Research.

[†]Supported by the research program “Equazioni di Kolmogorov” from the Italian “Ministero della Ricerca Scientifica e Tecnologica”

[‡]Supported by the DFG-Research Group 399, the SFB-701, the BIBOS-Research Center, the INTAS project 99-559, the RFBR project 04–01–00748, the Russian–Japanese Grant 05-01-02941-JF, the DFG Grant 436 RUS 113/343/0(R).

1 Introduction

Let $\mathscr{O}$ be an open bounded domain of $\mathbb{R}^n$ with smooth boundary $\partial\mathscr{O}$. We consider the linear operator Δ in $L^2(\mathscr{O})$ defined on $H^2(\mathscr{O}) \cap H_0^1(\mathscr{O})$. It is well known that $-\Delta$ is self-adjoint positive and anti-compact. So, there exists a complete orthonormal system $\{e_k\}$ in $L^2(\mathscr{O})$ of eigenfunctions of $-\Delta$. In fact we have $e_k \in \cap_{p\geq 1} L^p(\mathscr{O})$ for all $k \in \mathbb{N}$. We denote by $\{\lambda_k\}$ the corresponding sequence of eigenvalues,

$$\Delta e_k = -\lambda_k e_k, \quad k \in \mathbb{N}.$$

We shall consider a cylindrical Wiener process in $L^2(\mathscr{O})$ of the following form

$$W(t) = \sum_{k=1}^{\infty} \gamma_k(t) e_k, \quad t \geq 0,$$

where $\{\gamma_k\}$ is a sequence of mutually independent standard Brownian motions on a filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$. To be more specific, we shall assume that $1 \leq n \leq 3$.

In this work we consider the stochastic partial differential equation,

$$\begin{cases} dX(t) - \Delta\beta(X(t))dt = \sigma(X)dW(t), \quad t \geq 0, \\ \beta(X(t)) = 0, \quad \text{on } \partial\mathscr{O}, \quad t \geq 0, \\ X(0) = x, \end{cases} \tag{1.1}$$

where

$$\sigma(x)h := \sum_{k=1}^{\infty} \mu_k (h, e_k) x e_k, \quad x \in H^{-1}(\mathscr{O}),\ h \in L^2(\mathscr{O}), \tag{1.2}$$

and $\{\mu_k\}$ is a suitable sequence of nonnegative numbers.

The solution X to (1.1) is a function of $\omega \in \Omega$, $t \geq 0$ and $\xi \in \mathscr{O}$, but in the following it will be simply written $X(t)$ omitting ω and ξ. Here β is a continuous, differentiable, monotonically increasing function on $\mathbb{R}$ which satisfies the following conditions,

$$\begin{cases} \beta'(r) \leq \alpha_1 |r|^{m-1} + \alpha_2, \quad \forall\, r \in \mathbb{R},\ \beta(0) = 0, \\ B(r) := \int_0^r \beta(s)ds \geq \alpha_3 |r|^{m+1} + \alpha_4 r^2, \quad \forall\, r \in \mathbb{R}, \end{cases} \tag{1.3}$$

where $\alpha_1, \alpha_3 > 0,\ \alpha_2, \alpha_4 \geq 0$ (with $\alpha_2 > 0$ if $\alpha_4 > 0$) and $1 \leq m$. We note that since β is increasing, we have

$$r\beta(r) \geq B(r), \quad r \geq 0. \tag{1.4}$$

A standard example is $\beta(r) = a|r|^{m-1}r + br$ where $m \in \mathbb{R}$ and $a > 0, b \geq 0$.

Equation (1.1) with additive noise was recently studied in [4],[5],[7], [8],[9], see also [3]. In particular, [7] contains an existence result under similar conditions on β, but with additive noise.

It should be mentioned that existence and uniqueness of solutions to equation (1.1) follow also by the general results in [7] and [12] (see also [13] for generalizations). In this paper we present, however, an alternative proof, based on the Yosida approximations of $-\Delta\beta$, which allows us to prove the positivity of solutions for nonnegative initial data x. Taking into account the physical meaning of the equation (in mathematical modeling of flows through porous media, X is the density of the driven flow), the positivity of solutions is an essential property which should be addressed and much of the substance of this work is devoted to the rigorous proof of positivity via Itô's formula in appropriate functional spaces. The main difficulty in proving positivity is due to the fact that the Sobolev space $H^{-1}(\mathscr{O})$, which as in the deterministic case is natural for studying equation (1.1), is not suitable for truncation techniques usually used for proving positivity, and so a more involved approach is necessary.

Equation (1.1) can be written in the abstract form

$$\begin{cases} dX(t) + AX(t)dt = \sigma(X(t))dW(t), \quad t \geq 0, \\ \\ X(0) = x, \end{cases} \tag{1.5}$$

where the operator $A\colon D(A) \subset H^{-1}(\mathscr{O}) \to H^{-1}(\mathscr{O})$ is defined by

$$\begin{cases} Ax = -\Delta\beta(x), \quad x \in D(A), \\ \\ D(A) = \{x \in H^{-1}(\mathscr{O}) \cap L^1(\mathscr{O}) :\ \beta(x) \in H_0^1(\mathscr{O})\}, \end{cases} \tag{1.6}$$

and

$$\sigma(X(t))dW(t) = \sum_{k=1}^{\infty} X(t)e_k\, \mu_k d\gamma_k(t). \tag{1.7}$$

Troughout this paper we shall assume that

$$\sum_{k=1}^{\infty} \mu_k^2 \lambda_k^2 =: C < \infty. \tag{1.8}$$

To give a rigorous sense to this noise term we first note that since $n \leq 3$, by Sobolev embedding it follows that $\sup_{k\in\mathbb{N}} \frac{1}{\lambda_k} |e_k|_\infty < \infty$, since

$$|e_k|_\infty \leq C|e_k|_{H^2(\mathscr{O})} \leq C|\Delta e_k|_{L^2(\mathscr{O})} \leq C\lambda_k.$$

Thus, by equation (1.8) and for some constant $c_1 > 0$

$$\sum_{k=1}^{\infty} \mu_k^2 |xe_k|_{-1}^2 \leq c_1 \sum_{k=1}^{\infty} \mu_k^2 \ \lambda_k^2 \ |x|_{-1}^2 \leq c_1 C|x|_{-1}^2, \quad \forall\, x \in H^{-1}(\mathscr{O}), \quad (1.9)$$

because $|xe_k|_{-1}^2 \leq c_1\lambda_k^2|x|_{-1}^2$ by an elementary calculation.

We obtain by (1.9) that the series in (1.2) converges in $H^{-1}(\mathscr{O})$ and that $\sigma(x) \in L_2(L^2(\mathscr{O}), H^{-1}(\mathscr{O}))$. It follows that $W_1(t) := \int_0^t \sigma(X(s))dW(s)$ is well defined as a process on $H = H^{-1}(\mathscr{O})$. Note that since σ is linear we also have that $x \to \sigma(x)$ is Lipschitz from $H^{-1}(\mathscr{O})$ to $L_2(L^2(\mathscr{O}), H^{-1}(\mathscr{O}))$.

The plan of the paper is the following: main results are stated in §2 and proofs are given in §3.

The following notations will be used throughout in the following.

(i) $H_0^1(\mathscr{O}), H^2(\mathscr{O})$ are standard Sobolev spaces on $\mathscr{O}$ endowed with their usual norms denoted by $|\cdot|_{H_0^1(\mathscr{O})}$ and $|\cdot|_{H^2(\mathscr{O})}$ respectively.

(ii) H is the space $H^{-1}(\mathscr{O})$ (the dual of $H_0^1(\mathscr{O})$) endowed with the norm

$$|x|_H = |x|_{-1} = |-\Delta^{-1}x|_{H_0^1(\mathscr{O})}.$$

(Here $(-\Delta)^{-1}x = y$ is the solution to Dirichlet problem $-\Delta y = x$ in $\mathscr{O}$, $y \in H_0^1(\mathscr{O})$). The scalar product in H is

$$\langle x, z\rangle_{-1} = \int_{\mathscr{O}} (-\Delta)^{-1}\, x\, z\, d\xi, \quad \forall\, x, z \in H_0^1(\mathscr{O}).$$

(iii) The scalar product and the norm in $L^2(\mathscr{O})$ will be denoted by $(\cdot,\cdot)$ and $|\cdot|_2$, respectively and the norm in $L^p(\mathscr{O})$, $1 \leq p \leq \infty$ by $|\cdot|_p$.

(iv) For two Hilbert spaces H_1, H_2 the space of Hilbert-Schmidt operators from H_1 to H_2 is denoted by $L_2(H_1, H_2)$.

2 The main result

To begin with let us define the solution concept we shall work with. Formally, a solution to (1.1) (equivalently (1.5)) might be an H–valued continuous adapted process such that $X, AX \in C_W([0,T]; L^2(\Omega; H))$ and

$$X(t) = x - \int_0^t AX(s)ds + \int_0^t \sigma(X(s))dW(s), \quad t \in [0,T]. \tag{2.1}$$

By $C_W([0,T]; L^2(\Omega; H))$ we mean the Banach space of all the processes X in $(\Omega, \mathscr{F}, \mathbb{P})$ with values in H which are adapted and mean square continuous, endowed with the norm

$$\|X\|^2_{C_W([0,T];L^2(\Omega;H))} := \sup_{t\in[0,T]} \mathbb{E}|X(t)|^2_H.$$

Moreover, $L^p_W([0,T]; L^2(\Omega; H))$, $p \in [1,\infty]$, is the space of all processes $X \in L^p([0,T]; L^2(\Omega; H))$ which are adapted to W.

However, such a concept of solution might fail to exist for equation (1.1) and so we shall confine to a weaker one inspired by [7] and [11].

Definition 2.1 *An H-valued continuous $\mathscr{F}_t$-adapted process X is called a solution to* (1.1) *on* $[0,T]$ *if* $X \in L^{m+1}(\Omega \times (0,T) \times \mathscr{O})$ *and*

$$\begin{aligned}(X(t), e_j) &= (x, e_j) + \int_0^t \int_{\mathscr{O}} \beta(X(s))\Delta e_j d\xi ds \\ &\quad + \sum_{k=1}^\infty \mu_k \int_0^t (X(s)e_k, e_j)d\gamma_k(s), \quad \forall\, j \in \mathbb{N},\ t \in [0,T].\end{aligned} \tag{2.2}$$

Here m is the exponent arising in (1.3) and $\{e_k\}$ is the above orthonormal basis. Taking into account that $-\Delta e_j = \lambda_j e_j$ in $\mathscr{O}$ we may equivalently write (2.2) as follows

$$\begin{aligned}\langle X(t), e_j\rangle_{-1} &= \langle x, e_j\rangle_{-1} - \int_0^t \int_{\mathscr{O}} \beta(X(s))e_j d\xi ds \\ &\quad + \sum_{k=1}^\infty \mu_k \int_0^t \langle X(s)e_k, e_j\rangle_{-1} d\gamma_k(s), \quad \forall\, j \in \mathbb{N},\end{aligned}$$

i.e.

$$d\langle X(t), e_j\rangle_{-1} + (\beta(X(t)), e_j)dt = \sum_{k=1}^\infty \mu_k \langle X(s)e_k, e_j\rangle_{-1} d\gamma_k(s).$$

Recalling (1.7) we see that

$$\sum_{k=1}^{\infty} \mu_k(X(t)e_k, e_j)d\gamma_k(t) = (\sigma(X(t))dW(t), e_j), \quad j \in \mathbb{N}.$$

We also note that since by assumption (1.3), $\beta(X) \in L^{\frac{m+1}{m}}((0,T) \times \Omega \times \mathscr{O})$, the integral arising in the right hand side of (2.2) makes sense because $e_j \in C^\infty(\overline{\mathscr{O}})$ for all $j \in \mathbb{N}$. Of course, one might derive a vector valued version of Definition 2.1. As a matter of fact, it turns out (see [7]) that (2.2) can be rewritten equivalently as an equation in B' (the dual of $L^{m+1}(\mathscr{O}) \subset H = H' \subset B'$)

$$X(t) = x + \int_0^t LX(s)ds + \sqrt{Q}W(t),$$

where H and H' are identified by the Riesz isomorphism and L is a suitable extension of $x \to -\Delta\beta(x)$ to B'. In particular, this shows that Definition 2.1 is independent of the basis $\{e_k\}$.

Now we are ready to formulate the main results.

Theorem 2.2 *Assume that* (1.3) *and* (1.8) *hold. Then for each* $x \in H^{-1}(\mathscr{O})$ *there is a unique solution* X *to* (1.1). *Moreover, if* $x \in L^p(\mathscr{O})$ *is nonnegative a.e. on* $\mathscr{O}$ *where* $p \geq \max\{m+1, 4\}$ *is a natural number then* $X \in L^\infty_W(0,T; L^p(\Omega; L^p(\mathscr{O})))$ *and* $X \geq 0$ *a.e. on* $\Omega \times (0,\infty) \times \mathscr{O}$. *If* $x \in H^{-1}(\mathscr{O})$ *is such that* $x \geq 0$, *i.e.* x *is a positive measure, then* $\mathbb{P}$ *a.s.*

$$X(t) \geq 0 \quad \textit{for all}\, t \geq 0.$$

The positivity of the solution X to (1.1) will be proven below by choosing an appropriate Lyapunov function.

3 Proof of Theorem 2.2

We mention that in our estimates in the sequel constants may change from line to line though we do not express this in our notation.

We recall that the operator A, defined by (1.6), is maximal monotone in H (see e.g. [6]). Then we consider the Yosida approximation

$$A_\varepsilon(x) = \frac{1}{\varepsilon}\,(x - J_\varepsilon(x)) = A(1+\varepsilon A)^{-1}(x), \quad \varepsilon > 0,\ x \in H,$$

where $J_\varepsilon(x) = (1+\varepsilon A)^{-1}(x)$. The operator A_ε is monotone and Lipschitzian on H. Then, by (1.9) it follows by standard existence theory for stochastic equations in the Hilbert spaces (see e.g. [10]) that the approximating equation

$$\begin{cases} dX_\varepsilon(t) + A_\varepsilon X_\varepsilon(t)dt = \sigma(X_\varepsilon(t))dW(t), \quad t \geq 0, \\ \\ X_\varepsilon(0) = x, \end{cases} \tag{3.1}$$

has a unique solution $X_\varepsilon \in C_W([0,T];L^2(\Omega;H))$ such that $X_\varepsilon \in C([0,T];H)$, $\mathbb{P}$-a.s. with $A_\varepsilon X_\varepsilon \in C_W([0,T];L^2(\Omega;H))$.

By Itô's formula we have

$$\begin{aligned} &\frac{1}{2}\, d|X_\varepsilon(t)|_{-1}^2 + \langle A_\varepsilon X_\varepsilon(t), X_\varepsilon(t)\rangle_{-1} dt \\ &= \langle \sigma(X_\varepsilon(t))dW(t), X_\varepsilon(t)\rangle_{-1} + \frac{1}{2}\sum_{k=1}^{\infty} \mu_k^2 |X_\varepsilon(t)e_k|_{-1}^2 dt. \end{aligned} \tag{3.2}$$

This yields (see (1.9))

$$\begin{aligned} &\frac{1}{2}\,\mathbb{E}|X_\varepsilon(t)|_{-1}^2 + \mathbb{E}\int_0^t \langle A_\varepsilon X_\varepsilon(s), X_\varepsilon(s)\rangle_{-1} ds \\ &\leq \frac{1}{2}\,|x|_{-1}^2 + C\mathbb{E}\int_0^t |X_\varepsilon(s)|_{-1}^2 ds \end{aligned}$$

and therefore

$$\frac{1}{2}\,\mathbb{E}|X_\varepsilon(t)|_{-1}^2 + \mathbb{E}\int_0^t \langle A_\varepsilon X_\varepsilon(s), X_\varepsilon(s)\rangle_{-1} ds \leq C|x|_{-1}^2, \quad \forall\, \varepsilon > 0. \tag{3.3}$$

We set $Y_\varepsilon(t) = J_\varepsilon(X_\varepsilon(t))$ (see (3.1)). Then

$$\begin{aligned} &\frac{1}{2}\,\mathbb{E}|X_\varepsilon(t)|_{-1}^2 + \mathbb{E}\int_0^t \int_{\mathscr{O}} B(Y_\varepsilon(s))d\xi ds \\ &+\frac{1}{\varepsilon}\,\mathbb{E}\int_0^t |X_\varepsilon(s) - Y_\varepsilon(s)|_{-1}^2 ds \leq C|x|_{-1}^2, \quad \forall\, \varepsilon > 0. \end{aligned} \tag{3.4}$$

(Here we have used the equality

$$\langle A_\varepsilon x, x\rangle_{-1} = \langle AJ_\varepsilon x, J_\varepsilon x\rangle_{-1} + \frac{1}{\varepsilon}\,|x - J_\varepsilon(x)|_{-1}^2,$$

and (1.4).)

Now we fix $X \in C_W([0,T];L^2(\Omega,H))$ and consider the equation

$$\begin{cases} d\widetilde{X}_\varepsilon(t) + A_\varepsilon \widetilde{X}_\varepsilon(t)dt = \sigma(X(t))dW(t), \quad t \geq 0, \\ \\ \widetilde{X}_\varepsilon(0) = x. \end{cases} \tag{3.5}$$

Equivalently,

$$\begin{cases} d\widetilde{X}_\varepsilon(t) - \Delta\beta(\widetilde{Y}_\varepsilon(t))dt = \sigma(X(t))dW(t), \quad t \geq 0, \\ \\ \widetilde{X}_\varepsilon(0) = x, \end{cases} \tag{3.6}$$

where

$$\widetilde{Y}_\varepsilon = (1+\varepsilon A)^{-1}\widetilde{X}_\varepsilon = J_\varepsilon(\widetilde{X}_\varepsilon)).$$

We are going to prove that for $\varepsilon \to 0$, $\widetilde{X}_\varepsilon$ is convergent to a solution X^* to equation

$$dX^* + AX^*dt = \sigma(X)dW(t), \quad X^*(0) = x.$$

For equation (3.5) we have the same estimates as for (3.1). In fact by Itô's formula we get (see (3.4))

$$\begin{aligned} \mathbb{E}|\widetilde{X}_\varepsilon(t)|_{-1}^2 + \mathbb{E}\int_0^t \int_{\mathscr{O}} B(\widetilde{Y}_\varepsilon(s))d\xi ds \ &+\ \frac{1}{\varepsilon}\,\mathbb{E}\int_0^t |\widetilde{X}_\varepsilon(s) - \widetilde{Y}_\varepsilon(s)|_{-1}^2 ds \\ &\leq\ C|x|_{-1}^2 + C\mathbb{E}\int_0^t |X(s)|_{-1}^2, \end{aligned} \tag{3.7}$$

(where we have used (1.9) to estimate $\mathbb{E}\int_0^t \|\sigma(X(s))\|^2_{L_2(L^2(\mathscr{O});H)}ds$). By virtue of assumption (1.3) this implies that

$$\mathbb{E}\int_0^T \int_{\mathscr{O}} |\beta(\widetilde{Y}_\varepsilon(s))|^{\frac{m+1}{m}} d\xi ds \leq C(|x|_{-1}^2 + 1), \quad \varepsilon > 0,$$

(because $|\beta(r)| \leq \tilde{\alpha}_1|r|^m + \tilde{\alpha}_2,\ \tilde{\alpha}_1 \geq 0$), and so along a subsequence, we have

$$\beta(\widetilde{Y}_\varepsilon) \to \eta \quad \text{weakly in } L^{\frac{m+1}{m}}((0,T)\times\Omega\times\mathscr{O}). \tag{3.8}$$

On the other hand, we have by (3.6) that for $t \in [0,T]$

$$\langle \widetilde{X}_\varepsilon(t), e\rangle_{-1} + \int_0^t \int_{\mathscr{O}} \beta(\widetilde{Y}_\varepsilon(s))e\, d\xi ds = \langle x, e\rangle_{-1} + \int_0^t \langle \sigma(X(s))dW(s), e\rangle_{-1} ds,$$

for all $e \in L^{m+1}(\mathscr{O})$. We note that by (3.7) there exists $X^* \in L^2_W([0,T];L^2(\Omega;H))$ such that

$$\widetilde{X}_\varepsilon \to X^* \quad \text{weakly in } \ L^2_W([0,T];L^2(\Omega;H)) \tag{3.9}$$

and by (3.7) and (1.3) we obtain that also

$$\widetilde{Y}_\varepsilon \to X^* \quad \text{weakly in } L^2_W([0,T];L^2(\Omega;H)) \cap L^{m+1}(\Omega\times(0,T)\times\mathscr{O}). \quad (3.10)$$

Hence along a subsequence $\varepsilon \to 0$

$$\mathbb{E}\langle \widetilde{X}_\varepsilon(t), e\rangle_{-1} \to \mathbb{E}\langle X^*(t), e\rangle_{-1} \quad \text{weakly in } L^2(0,T).$$

Then letting ε tend to 0 we get for a.e. $t \in [0,T]$

$$\langle X^*(t), e\rangle_{-1} = \langle x, e\rangle_{-1} - \int_0^t \int_{\mathscr{O}} \eta(s) e d\xi ds + \int_0^t \langle \sigma(X(s))dW(s), e\rangle_{-1} ds. \quad (3.11)$$

Taking into account (3.9)-(3.10), to conclude the proof of existence it suffices to show that

$$\eta(t,\xi,\omega) = \beta(X^*(t,\xi,\omega)) \quad \text{a.e. } (\omega,t,\xi) \in \Omega\times(0,T)\times\mathscr{O}. \quad (3.12)$$

Indeed, in such a case we may take in (3.11) $e = \Delta e_j$ for $j \in \mathbb{N}$.

To this end we consider the operator

$$F\colon L^m(\Omega\times(0,T)\times\mathscr{O}) \to L^{\frac{m}{m+1}}(\Omega\times(0,T)\times\mathscr{O}) = (L^m(\Omega\times(0,T)\times\mathscr{O}))',$$

defined by

$$(Fx)(t,\xi,\omega) = \beta(x(t,\xi,\omega)) \quad \text{a.e. } (\omega,t,\xi) \in \Omega\times(0,T)\times\mathscr{O}.$$

This operator is maximal monotone and more precisely, it is the subgradient of the convex function $\Phi : L^{m+1}(\Omega\times(0,T)\times\mathscr{O}) \to \mathbb{R}$ defined as,

$$\Phi(x) = \mathbb{E}\int_0^T \int_{\mathscr{O}} B(x(t,\xi)) d\xi dt.$$

For each $Z \in L^{m+1}(\Omega\times(0,T)\times\mathscr{O})$ we have

$$\Phi(\widetilde{Y}_\varepsilon) - \Phi(Z) \le \mathbb{E}\int_0^T \int_{\mathscr{O}} \beta(\widetilde{Y}_\varepsilon(t,\xi))(\widetilde{Y}_\varepsilon(t,\xi) - Z(t,\xi)) d\xi dt.$$

Letting ε tend to 0 we have by (3.8), (3.9), (3.10) and by the weak lower semicontinuity of Φ

$$\Phi(X^*) - \Phi(Z) \le \liminf_{\varepsilon\to 0} \mathbb{E}\int_0^T \int_{\mathscr{O}} \beta(\widetilde{Y}_\varepsilon(t,\xi))\widetilde{Y}_\varepsilon(t,\xi) d\xi dt - \mathbb{E}\int_0^T \int_{\mathscr{O}} \eta Z d\xi dt.$$

To prove (3.12) by the uniqueness of the subgradient it suffices to show that

$$\liminf_{\varepsilon\to 0} \mathbb{E}\int_0^T \int_{\mathscr{O}} \beta(\widetilde{Y}_\varepsilon(t,\xi))\widetilde{Y}_\varepsilon(t,\xi)d\xi dt \le \mathbb{E}\int_0^T \int_{\mathscr{O}} \eta X^* d\xi dt. \tag{3.13}$$

To this end we come back to equation (3.6) and note that by Itô's formula we have

$$\frac{1}{2}\,\mathbb{E}|\widetilde{X}_\varepsilon(t)|_{-1}^2 + \mathbb{E}\int_0^t \int_{\mathscr{O}} \beta(\widetilde{Y}_\varepsilon(s))\widetilde{X}_\varepsilon(s)d\xi ds$$

$$= \frac{1}{2}\,|x|_{-1}^2 + \frac{1}{2}\,\sum_{k=1}^\infty \mathbb{E}\int_0^t \mu_k^2 |X(s)e_k|_{-1}^2 ds.$$

Equivalently,

$$\begin{aligned}
&\frac{1}{2}\,\mathbb{E}|\widetilde{X}_\varepsilon(t)|_{-1}^2 + \mathbb{E}\int_0^t \int_{\mathscr{O}} \beta(\widetilde{Y}_\varepsilon(s))\widetilde{Y}_\varepsilon(s)d\xi ds \\
&+\mathbb{E}\int_0^t \int_{\mathscr{O}} \beta(\widetilde{Y}_\varepsilon(s))(\widetilde{X}_\varepsilon(s)-\widetilde{Y}_\varepsilon(s))d\xi ds \\
&= \frac{1}{2}\,|x|_{-1}^2 + \frac{1}{2}\,\sum_{k=1}^\infty \mathbb{E}\int_0^t \mu_k^2 |X(s)e_k|_{-1}^2 ds.
\end{aligned} \tag{3.14}$$

By (3.9)-(3.10) we have

$$\int_{\mathscr{O}} \beta(\widetilde{Y}_\varepsilon(s))(\widetilde{X}_\varepsilon(s)-\widetilde{Y}_\varepsilon(s))d\xi = \langle A_\varepsilon \widetilde{X}_\varepsilon(s), \widetilde{X}_\varepsilon(s)-J_\varepsilon(\widetilde{X}_\varepsilon(s))\rangle_{-1} = \varepsilon|A_\varepsilon \widetilde{X}_\varepsilon(s)|_{-1}^2.$$

Fix $\varphi \in L^\infty(0,T)$, $\varphi \ge 0$. Then $\varphi X^* \in L^2_W(0,T;L^2(\Omega;H))$. Thus by (3.9)-(3.10)

$$\mathbb{E}\int_0^T \varphi(t)|X^*(t)|_{-1}^2\,dt = \lim_{\varepsilon\to 0}\mathbb{E}\int_0^T \langle X^*(t), X_\varepsilon(t)\rangle_{-1}\varphi(t)dt$$

$$\le \left(\mathbb{E}\int_0^T \varphi(t)|X^*(t)|_{-1}^2\,dt\right)^{1/2} \liminf_{\varepsilon\to 0}\left(\mathbb{E}\int_0^T \varphi(t)|X_\varepsilon(t)|_{-1}^2\,dt\right)^{1/2}.$$

Hence simplifying we obtain

$$\mathbb{E}\int_0^T \varphi(t)|X^*(t)|_{-1}^2\,dt \le \liminf_{\varepsilon\to 0}\mathbb{E}\int_0^T \varphi(t)|X_\varepsilon(t)|_{-1}^2\,dt.$$

Hence (3.14), Fatou's Lemma (see also (1.4)) and the arbitrariness of φ implies that for a.e. $t \in [0,T]$ we obtain that

$$\liminf_{\varepsilon\to 0} \mathbb{E}\int_0^t \int_{\mathscr{O}} \beta(\widetilde{Y}_\varepsilon(s)\widetilde{Y}_\varepsilon(s)d\xi ds + \frac{1}{2}\,\mathbb{E}|X^*(t)|_{-1}^2$$
$$\leq \frac{1}{2}\,|x|_{-1}^2 + \frac{1}{2}\,\sum_{k=1}^{\infty}\mathbb{E}\int_0^t \mu_k^2 |X(s)e_k|_{-1}^2 ds. \tag{3.15}$$

On the other hand, by (3.11) we see via Itô's formula (applied to the right hand side of (3.11), since the left hand side might not be continuous in t) that for all $j \in \mathbb{N}$ and a.e. $t \in [0,T]$,

$$\frac{1}{2}\,\mathbb{E}|\langle X^*(t), e_j\rangle_{-1}|^2 + \mathbb{E}\int_0^t \langle \eta_s, e_j\rangle\langle X^*(s), e_j\rangle_{-1} ds$$
$$= \frac{1}{2}\,\langle x, e_j\rangle_{-1}^2 + \frac{1}{2}\,\mathbb{E}\sum_{k=1}^{\infty}\mu_k^2 \int_0^t \langle X(s)e_k, e_j\rangle^2 ds$$

and dividing by $|e_j|_{-1}^2$ and summing over j we obtain

$$\frac{1}{2}\,\mathbb{E}|X^*(t)|_{-1}^2 + \mathbb{E}\int_0^t \int_{\mathscr{O}} \eta(s)X^*(s)d\xi ds$$
$$= \frac{1}{2}\,|x|_{-1}^2 + \frac{1}{2}\,\sum_{k=1}^{\infty}\mu_k^2\mathbb{E}\int_0^t |X(s)e_k|_{-1}^2 ds. \tag{3.16}$$

We note that the integral in the left hand side makes sense since by (3.4), $X^* \in L^{m+1}((0,T)\times\Omega\times\mathscr{O})$ while $\eta \in L^{\frac{m+1}{m}}((0,T)\times\Omega\times\mathscr{O})$.

Comparing (3.15) and (3.16) we infer that

$$\liminf_{\varepsilon\to 0} \mathbb{E}\int_0^T \int_{\mathscr{O}} \beta(\widetilde{Y}_\varepsilon(t))\widetilde{Y}_\varepsilon(t)d\xi dt \leq \mathbb{E}\int_0^T \int_{\mathscr{O}} \eta(t)X^*(t)d\xi dt,$$

as claimed. A formal problem arises, however, because $X^*(t)$ as constructed before might not be H-continuous. However, arguing as in [11], [12] we may replace it by an H-continuous version defined by

$$\widetilde{X}^*(t) = x + \int_0^t \Delta\eta(s)ds + \int_0^t \sigma(X(s))dW(s).$$

It follows that $X^* = \widetilde{X}^*$ a.e. and that $\widetilde{X}^*$ is also an $\mathscr{F}_t$-adapted process. Moreover, the Itô formula from ([11, Theorem I-3-2]) holds. Hence $\widetilde{X}^* \in$

$C_W([0,T];L^2(\Omega;H))\cap L^{m+1}((0,T)\times\Omega\times\mathscr{O})$ is a solution (in the sense of Definition 2.1) to

$$\left\{\begin{array}{l} dX^*+AX^*dt=\sigma(X)dW \\ \\ X^*(0)=x, \end{array}\right. \tag{3.17}$$

as claimed.

Uniqueness for equation (3.17). Let X_1^*,X_2^* be two solutions to equation (3.17) for $X=X_i,\ i=1,2$. We have (see (2.2))

$$d\langle X_1^*-X_2^*,e_j\rangle_{-1}+\int_{\mathscr{O}}(\beta(X_1^*)-\beta(X_2^*))e_jd\xi dt=\sum_{k=1}^{\infty}\mu_k\langle (X_1-X_2)e_k,e_j\rangle_{-1}d\gamma_k.$$

By Itô's formula we obtain

$$\begin{array}{l}\frac12\,\mathbb{E}|\langle X_1^*(t)-X_2^*(t),e_j\rangle_{-1}|^2 \\ \\ +\mathbb{E}\displaystyle\int_0^t(\beta(X_1^*(s))-\beta(X_2^*(s)),e_j)\langle X_1^*(s)-X_2^*(s),e_j\rangle_{-1}ds \\ \\ =\displaystyle\frac12\,\mathbb{E}\int_0^t\sum_{k=1}^{\infty}\mu_k^2\langle (X_1(s)-X_2(s))e_k,e_j\rangle_{-1}^2ds\end{array}$$

Dividing by $|e_j|_{-1}^2$ and summing over j we see that

$$\begin{array}{l}\displaystyle\frac12\,\mathbb{E}|X_1^*(t)-X_2^*(t)|_{-1}^2+\mathbb{E}\int_0^t(\beta(X_1^*)-\beta(X_2^*),X_1^*(s)-X_2^*(s)\;ds \\ \\ =\displaystyle\frac12\,\mathbb{E}\int_0^t\sum_{j,k=1}^{\infty}\mu_k^2\langle (X_1(s)-X_2(s))e_k,|e_j|_{-1}^{-1}\,e_j\rangle_{-1}^2ds.\end{array}$$

Hence (see (1.9))

$$\mathbb{E}|X_1^*(t)-X_2^*(t)|_{-1}^2\le CE\int_0^t|X_1(s)-X_2(s)|_{-1}^2ds,\quad \forall\, t\in[0,T] \tag{3.18}$$

Now we shall use the latter inequality to prove existence of a unique solution

$$X\in C_W([0,T];L^2(\Omega;H))\cap L^{m+1}((0,T)\times\Omega\times\mathscr{O})$$

to equation (1.1). Indeed the operator $X \to X^*$ is a contraction on the space $C_W([0,T]; L^2(\Omega; H))$ if T is sufficiently small and so, we have existence (and uniqueness) for $T > 0$ small. By a standard unique continuation argument it follows existence and uniqueness on an arbitrary interval $[0,T]$.

Positivity. We shall assume now that $x \in L^p(\mathscr{O})$, where $p \geq \max\{m+1, 4\}$, and $x(\xi) \geq 0$ a.e. in $\mathscr{O}$. We shall prove that

$$X \geq 0 \quad \text{a.e. in } (0,T) \times \mathscr{O} \times \Omega. \tag{3.19}$$

We shall first assume in addition that β is strictly monotone, i.e.

$$(\beta(r) - \beta(\bar{r}))(r - \bar{r}) \geq \alpha(r - \bar{r})^2, \quad \forall\, r, \bar{r} \in \mathbb{R}, \tag{3.20}$$

where $\alpha > 0$. Below we shall use the following lemma.

Lemma 3.1 *Let $y \in D(A)$ and $g : \mathbb{R} \to \mathbb{R}$ Lipschitz and increasing. Then*

$$\langle \nabla \beta(y), \nabla g(y) \rangle_{\mathbb{R}^n} \geq 0, \quad \textit{a.e. on } \mathscr{O}.$$

Proof. First note that by definition of $D(A)$ we have that $y, \beta(y) \in H_0^1(\mathscr{O})$. Using a Dirac sequence we can find mollifiers $g_k \in C^1(\mathbb{R})$, $g_k' \geq 0$, $k \in \mathbb{N}$, such that

$$\nabla g(y) = \lim_{k \to \infty} g_k'(y) \nabla y \quad \text{in } L^2(\mathscr{O}).$$

So, it suffices to prove that

$$\langle \nabla \beta(y), \nabla g(y) \rangle_{\mathbb{R}^n} \geq 0, \quad \text{a.e. on } \mathscr{O}.$$

But

$$\langle \nabla \beta(y), \nabla y \rangle_{\mathbb{R}^n} = \langle \nabla \beta(y), \nabla \beta^{-1}\beta(y) \rangle_{\mathbb{R}^n}.$$

Since β is strictly monotone, β^{-1} is Lipschitz, so applying the above mollifier argument with β^{-1} replacing g, we prove the assertion. $\square$

We shall use the approximating equation (3.1) whose solution X_ε is weakly convergent to X in $L^2_W(\Omega; L^2(0,T;H))$. Namely, we have for $Y_\varepsilon(t) := J_\varepsilon(X_\varepsilon(t))$, $t \geq 0$,

$$dX_\varepsilon(t) - \Delta\beta(Y_\varepsilon(t))dt = \sigma(X_\varepsilon(t))dW(t), \quad t \geq 0. \tag{3.21}$$

We note that equation (3.1) can be equivalently written as

$$\begin{cases} dX_\varepsilon(t) + \dfrac{1}{\varepsilon}\, X_\varepsilon(t)dt = \dfrac{1}{\varepsilon}\, J_\varepsilon(X_\varepsilon(t))dt + \sigma(X_\varepsilon(t))dW(t), \quad t \geq 0, \\[2ex] X_\varepsilon(0) = x, \end{cases} \tag{3.22}$$

Fix $x \in H$ and set

$$y = J_\varepsilon(x) = (1 - \varepsilon\Delta\beta)^{-1}x,$$

i.e.

$$y - \varepsilon\Delta\beta(y) = x \tag{3.23}$$

Then $y \in D(A)$. Since β is strictly monotone, β^{-1} is Lipschitz. Therefore, since $\beta(y) \in H_0^1(\mathscr{O})$, also $y \in H_0^1(\mathscr{O})$. Now assume $x \in L^p(\mathscr{O})$. By multiplying both sides of (3.23) by $\frac{y^{p-1}}{1+\lambda y^{p-2}}$ and integrating over $\mathscr{O}$ we get by Lemma 3.1

$$\int_{\mathscr{O}} \frac{y^p}{1+\lambda|y|^{p-2}}\, d\xi \le \int_{\mathscr{O}} \frac{y^{p-1}x}{1+\lambda|y|^{p-2}}\, d\xi.$$

Then, letting $\lambda \to 0$ we find the estimate

$$|y|_p^p \le \int_{\mathscr{O}} y^{p-1}x d\xi \le |y|_p^{p-1}\ |x|_p. \tag{3.24}$$

Hence

$$|J_\varepsilon(x)|_p \le |x|_p, \quad \forall\, x \in L^p(\mathscr{O}), \tag{3.25}$$

and therefore,

$$|A_\varepsilon(x)|_p = \frac{1}{\varepsilon}\ |x - J_\varepsilon(x)|_p \le \frac{2}{\varepsilon}\ |x|_p, \quad \forall\, x \in L^p(\mathscr{O}).$$

(3.23) and (3.25) imply that J_ε is continuous from $L^p(\mathscr{O})$ into itself.

Lemma 3.2 *For each $x \in L^2(\mathscr{O})$ equation* (3.22) *has a unique solution $X_\varepsilon \in C_W([0,T]; L^2(\Omega; L^2(\mathscr{O})))$.*

Proof. Let us first prove that $J_\varepsilon = (1 - \varepsilon\Delta\beta)^{-1}$ is Lipschitz continuous in $L^2(\mathscr{O})$. Indeed, by the equation

$$J_\varepsilon(x) - \varepsilon\Delta\beta(J_\varepsilon(x)) = x, \quad \text{in } \mathscr{O},$$

(taking into account that $\beta(J_\varepsilon(x)) \in H_0^1(\mathscr{O})$) we have for $x, \bar{x} \in L^2(\mathscr{O})$

$$\int_{\mathscr{O}} (J_\varepsilon(x) - J_\varepsilon(\bar{x}))(\beta(J_\varepsilon(x)) - \beta(J_\varepsilon(\bar{x})))d\xi$$

$$+\varepsilon \int_{\mathscr{O}} |\nabla(\beta(J_\varepsilon(x)) - \beta(J_\varepsilon(\bar{x}))|^2 d\xi \le \int_{\mathscr{O}} (x - \bar{x})(\beta(J_\varepsilon(x)) - \beta(J_\varepsilon(\bar{x})))d\xi.$$

This yields, recalling (3.20)

$$\alpha|J_\varepsilon(x)-J_\varepsilon(\bar{x})|_2^2+\varepsilon|\beta(J_\varepsilon(x))-\beta(J_\varepsilon(\bar{x}))|^2_{H^1_0(\mathscr{O})} \le |x-\bar{x}|_2\, |\beta(J_\varepsilon(x))-\beta(J_\varepsilon(\bar{x}))|_2.$$

On the other hand, by the Poincaré inequality there exists $C>0$ such that

$$|\beta(J_\varepsilon(x))-\beta(J_\varepsilon(\bar{x}))|_2^2 \le C|\beta(J_\varepsilon(x))-\beta(J_\varepsilon(\bar{x}))|^2_{H^1_0(\mathscr{O})}.$$

Therefore

$$\alpha|J_\varepsilon(x)-J_\varepsilon(\bar{x}))|_2^2+\frac{\varepsilon}{2}\,|\beta(J_\varepsilon(x))-\beta(J_\varepsilon(\bar{x}))|^2_{H^1_0(\mathscr{O})}+\frac{\varepsilon}{2C}\,|\beta(J_\varepsilon(x))-\beta(J_\varepsilon(\bar{x}))|_2^2$$

$$\le \frac{C}{2\varepsilon}\,|x-\bar{x}|_2^2\;+\frac{\varepsilon}{2C}\,|\beta(J_\varepsilon(x))-\beta(J_\varepsilon(\bar{x}))|_2^2,$$

and consequently

$$\alpha|J_\varepsilon(x)-J_\varepsilon(\bar{x}))|_2^2+\frac{\varepsilon}{2}\,|\beta(J_\varepsilon(x))-\beta(J_\varepsilon(\bar{x}))|^2_{H^1_0(\mathscr{O})} \le \frac{C}{2\varepsilon}\,|x-\bar{x}|_2.$$

So, J_ε is Lipschitz continuous in $L^2(\mathscr{O})$ as claimed. Consequently $A_\varepsilon = \frac{1}{\varepsilon}\,(1-J_\varepsilon)$ is Lipschitz continuous in $L^2(\mathscr{O})$ as well. Moreover, since

$$\|\sigma(x)\|_{L_2(L^2(\mathscr{O}),L^2(\mathscr{O}))} \le \sum_{k=1}^\infty \mu_k^2|xe_k|_2^2 \le \sum_{k=1}^\infty \mu_k^2|e_k|^2_{L^\infty(\mathscr{O})}\,|x|_2^2 \le C_1\sum_{k=1}^\infty \mu_k^2\lambda_k^2\,|x|_2^2$$

we infer by standard existence theory for stochastic PDEs that for each $x\in L^2(\mathscr{O})$ equation (3.22) has a unique solution in $X_\varepsilon\in C_W([0,T];L^2(\Omega;L^2(\mathscr{O})))$ (see e.g. [10]). $\square$

For $R>0$ define

$$K_R := \{X\in L^\infty_W(0,T;L^p(\Omega\times\mathscr{O})) :\; e^{-4\alpha t}\mathbb{E}|X(t)|_p^p \le R^p \quad \text{for a.e. } t\in[0,T]\}$$

Lemma 3.3 *Let $T>0$ and $x\in L^p(\mathscr{O})$. Then for the solution X_ε of* (3.1) *(or equivalently* (3.22)*) we have $X_\varepsilon\in L^\infty_W(0,T;L^p(\Omega\times\mathscr{O}))$ and X_ε is bounded in $L^\infty_W(0,T;L^p(\Omega\times\mathscr{O}))$*

Proof. Obviously, K_R is a closed subset of $L^\infty_W(0,T;L^p(\Omega\times\mathscr{O}))$. Since by (3.22) X_ε is a fixed point of the map

$$X \overset{F}{\mapsto} e^{-\frac{t}{\varepsilon}}x+\frac{1}{\varepsilon}\int_0^t e^{-\frac{(t-s)}{\varepsilon}}J_\varepsilon(X(s))ds+\int_0^t e^{-\frac{(t-s)}{\varepsilon}}\sigma(X(s))dW(s), \quad t\in[0,T],$$

obtained by iteration in $C_W(0,T;L^2(\Omega\times\mathscr{O}))$, it suffices to prove that this map leaves K_R invariant for R large enough. But for $X\in K_R$ we have by (3.25) for $t\ge 0$

$$\left(e^{-p\alpha t}\mathbb{E}\left|e^{-\frac{t}{\varepsilon}}x+\frac{1}{\varepsilon}\int_0^t e^{-\frac{(t-s)}{\varepsilon}}J_\varepsilon(X(s))ds\right|_p^p\right)^{1/p}$$

$$\le e^{-\alpha t}e^{-\frac{t}{\varepsilon}}|x|_p+e^{-\alpha t}\left(\mathbb{E}\left(\int_0^t\frac{1}{\varepsilon}\,e^{-\frac{(t-s)}{\varepsilon}}|J_\varepsilon(X(s))|_p ds\right)^p\right)^{1/p}$$

$$\le e^{-(\frac{1}{\varepsilon}+\alpha)t}|x|_p+e^{-\alpha t}\Bigg(\int_0^t\cdots\int_0^t\frac{1}{\varepsilon}\,e^{-\frac{(t-s_1)}{\varepsilon}}e^{\alpha s_1}\cdots\frac{1}{\varepsilon}\,e^{-\frac{(t-s_p)}{\varepsilon}}e^{\alpha s_p}$$

$$\times e^{-\alpha s_1}(\mathbb{E}|(|X(s_1)|_p^p)^{1/p}\cdots e^{-\alpha s_p}(\mathbb{E}|(|X(s_p)|_p^p)^{1/p}\,ds_1\cdots ds_p\Bigg)^{1/p}$$

$$\le e^{-(\frac{1}{\varepsilon}+\alpha)t}|x|_p+e^{-\alpha t}R\int_0^t\frac{1}{\varepsilon}\,e^{-\frac{(t-s)}{\varepsilon}}e^{\alpha s}ds$$

$$\le e^{-(\frac{1}{\varepsilon}+\alpha)t}|x|_p+\tfrac{R}{1+\alpha\varepsilon}.$$

Now we set

$$Y(t)=\int_0^t e^{-\frac{(t-s)}{\varepsilon}}X(s))dW(s),\quad t\ge 0.$$

Then

$$\begin{cases} dY(t)+\dfrac{1}{\varepsilon}\,Y(t)dt=\sigma(X(t))dW(t), & t\ge 0,\\[2ex] Y(0)=0.\end{cases}$$

Let $\lambda>0$. Applying Itô's formula to the function

$$\Psi_\lambda(y):=\frac{1}{p}\,|(1+\lambda A_0)^{-1}y|_p^p,\quad y\in L^p(\mathscr{O}),$$

(see the beginning of the proof of the next lemma for a detailed justification) we obtain via Hölder's inequality that

$$\mathbb{E}[\Psi_\lambda(Y(t))] + \frac{1}{\varepsilon}\,\mathbb{E}\int_0^t \int_{\mathscr{O}} |(1+\lambda A_0)^{-1}Y(s)|^p d\xi\, ds$$

$$= \frac{p-1}{2}\ \sum_{k=1}^{\infty} \mu_k^2\, \mathbb{E}\int_0^t \int_{\mathscr{O}} |(1+\lambda A_0)^{-1}Y(s)|^{p-2}$$

$$\times |(1+\lambda A_0)^{-1}(X(s)e_k)|^2 d\xi\, ds$$

$$\leq C\mathbb{E}\int_0^t |(1+\lambda A_0)^{-1}Y(s)|_p^2\, |X(s)|_p^2 ds$$

$$\leq \frac{1}{2\varepsilon}\,\mathbb{E}\int_0^t |(1+\lambda A_0)^{-1}Y(s)|_p^p ds + \frac{9C^2\varepsilon}{8}\,\mathbb{E}\int_0^t |X(s)|_p^p ds$$

$$\leq \frac{1}{2\varepsilon}\,\mathbb{E}\int_0^t |(1+\lambda A_0)^{-1}Y(s)|_p^p ds + \frac{9C^2\varepsilon(e^{4\alpha t}-1)}{32\alpha}\, R^p.$$

Then letting $\lambda \to \infty$, we see by Fatou's lemma that for a.e. $t \in [0,T]$ we have for C_1 independent of ε

$$e^{-4\alpha t}\mathbb{E}|Y(t)|_p^p \leq \frac{C_1\varepsilon}{\alpha} R^p, \quad \forall\, t \in [0,T].$$

This means that for α large enough and $R > 2|x|_p$ the map leaves K_R invariant as claimed.

Lemma 3.4 *For $x \in L^p(\mathscr{O}))$ we have*

$$X_\varepsilon \to X \quad \textit{strongly in } L_W^\infty(0,T;L^2(\Omega;H)),$$

$$X_\varepsilon \to X \quad \textit{weakly in } L_W^\infty(0,T;L^p(\Omega;L^p(\mathscr{O}))),$$

where X is the solution to (1.1).

Proof. By (3.4) and Lemma 3.3 we know that $\{X_\varepsilon\}$ is bounded in

$$L_W^2(0,T;L^2(\Omega;H)) \cap L_W^\infty(0,T;L^p(\Omega;L^p(\mathscr{O})))$$

Subtracting equations (1.1) and (3.1) we get via Itô's formula and because β is increasing that

$$\frac{1}{2}\,\mathbb{E}|X_\varepsilon(t)-X(t)|_{-1}^2 + \mathbb{E}\int_0^t \int_{\mathscr{O}} (\beta((1+\varepsilon A)^{-1}X) - \beta(X))(X_\varepsilon - X) d\xi ds$$

$$\leq c\mathbb{E}\int_0^t |X_\varepsilon(s)-X(s)|_{-1}^2 ds,$$

and by Gronwall's lemma we obtain

$$\mathbb{E}|X_\varepsilon(t)-X(t)|_{-1}^2 \le C\mathbb{E}\int_0^1\int_{\mathscr{O}}(\beta((1+\varepsilon A)^{-1}X)-\beta(X))(X_\varepsilon-X)d\xi ds. \quad (3.26)$$

On the other hand, it follows by (3.25) that

$$\int_{\Omega\times[0,T]\times\mathscr{O}}|(1+\varepsilon A)^{-1}X|^p\mathbb{P}(d\omega)\,dt\,d\xi \le \int_{\Omega\times[0,T]\times\mathscr{O}}|X|^p\mathbb{P}(d\omega)\,dt\,d\xi,$$

while for $\varepsilon\to 0$

$$(1+\varepsilon A)^{-1}X\to X \quad \text{in} \;\; L^1(\mathscr{O})$$

for $(\omega,t)\in\Omega\times[0,T]$ (which is a consequence of the fact that the operator A is m-accretive in $L^1(\mathscr{O})$, cfr. [2]). Hence (at least along a subsequence)

$$(1+\varepsilon A)^{-1}X\to X \quad \text{a.e. on} \;\; \Omega\times[0,T]\times\mathscr{O}.$$

Hence

$$(1+\varepsilon A)^{-1}X\to X \quad \text{weakly in} \;\; L^p(\Omega\times[0,T]\times\mathscr{O})$$

as $\varepsilon\to 0$ and according to the above inequality this implies that for $\varepsilon\to 0$, $|(1+\varepsilon A)^{-1}X|_{L^p}\to|X|_{L^p}$. Hence since $L^p(\Omega\times[0,T]\times\mathscr{O})$ is uniformly convex,

$$(1+\varepsilon A)^{-1}X\to X \quad \text{strongly in} \;\; L^p(\Omega\times[0,T]\times\mathscr{O}),$$

see [2]. Next by assumption (1.3) we have

$$\begin{aligned}
&|\beta((1+\varepsilon A)^{-1}X)-\beta(X)|\\
&\le\int_0^1\beta'(\lambda(1+\varepsilon A)^{-1}X)+(1-\lambda)X)|(1+\varepsilon A)^{-1}X-X|d\lambda\\
&\le C\left(|(1+\varepsilon A)^{-1}X|^{m-1}+|X|^{m-1}+1\right)|(1+\varepsilon A)^{-1}X-X|.
\end{aligned}$$

This yields, via Hölder's inequality

$$\begin{aligned}
&\left|\mathbb{E}\int_0^t\int_{\mathscr{O}}(\beta((1+\varepsilon A)^{-1}X)-\beta(X))(X_\varepsilon-X)d\xi ds\right|\\
&\le C|X_\varepsilon-X|_{L^p(\Omega\times[0,T]\times\mathscr{O})}|(1+\varepsilon A)^{-1}X-X|_{L^p(\Omega\times[0,T]\times\mathscr{O})}\\
&\times\left(|(1+\varepsilon A)^{-1}X|^{m-1}_{L^p(\Omega\times[0,T]\times\mathscr{O})}+|X|^{m-1}_{L^p(\Omega\times[0,T]\times\mathscr{O})}+1\right)\\
&\le C_1|(1+\varepsilon A)^{-1}X-X|_{L^p(\Omega\times[0,T]\times\mathscr{O})}\to 0,
\end{aligned}$$

because $\{X_\varepsilon\}$ is bounded in $L^p(\Omega\times[0,T]\times\mathscr{O})$ and $(m-1)\frac{p}{p-2}\le p$. Now the assertion follows by (3.26).

Consider now the function

$$\varphi(x)=\frac{1}{p}\,|x^-|_p^p.$$

For any $x\in L^p(\mathscr{O})$, φ is Gâteaux differentiable and its differential $D\varphi\colon L^p(\mathscr{O})\to L^{p/(p-1)}(\mathscr{O})$ is given by

$$D\varphi(x)=-(x^-)^{p-1},$$

while the second Gâteaux derivative $D^2\varphi(x)\in L(L^p(\mathscr{O});L^{p/(p-1)}(\mathscr{O}))$ is given by

$$(D^2\varphi(x)h,g)=(p-1)\int_{\mathscr{O}} h\,g\,|x^-|^{p-2}d\xi,\quad \forall\, h,g,x\in L^p(\mathscr{O}).$$

Lemma 3.5 *Let $n\le 3$. For each $x\in L^p(\mathscr{O})$ we have*

$$\begin{aligned}&\mathbb{E}[\varphi(X_\varepsilon(t))]+\mathbb{E}\int_0^t(A_\varepsilon X_\varepsilon(s),D\varphi(X_\varepsilon(s))ds\\&=\varphi(x)+\frac{p-1}{2}\sum_{k=1}^\infty\mu_k^2\,\mathbb{E}\int_0^t\int_{\mathscr{O}}|X_\varepsilon^-(s)e_k|^2|X_\varepsilon^-(s)|^{p-2}d\xi ds.\end{aligned}\tag{3.27}$$

Proof. We note first that since $X_\varepsilon\in L^\infty_W(0,T;L^p(\Omega;L^p(\mathscr{O})))$ the above formula makes sense. Next we approximate φ by

$$\varphi_\lambda(x)=\varphi((1+\lambda A_0)^{-1}x),\quad A_0=-\Delta,\ \ D(A_0)=H^2(\mathscr{O})\cap H^1_0(\mathscr{O}),\ \lambda>0.$$

Since $\varphi\in C^2(C(\overline{\mathscr{O}}))$ and $(1+\lambda A_0)^{-1}$ is linear continuous from $L^2(\mathscr{O})$ to $C(\overline{\mathscr{O}})$ (due to our assumption $n\le 3$) we infer that $\phi_\lambda\in C^2(L^2(\mathscr{O}))$ and its first order and second order differentials are given, respectively, by

$$D\varphi_\lambda(x)=D\varphi((1+\lambda A_0)^{-1}x))(1+\lambda A_0)^{-1},$$

$$(D^2\varphi_\lambda(x)h,k)=(D^2\varphi((1+\lambda A_0)^{-1}x))((1+\lambda A_0)^{-1}h,(1+\lambda A_0)^{-1}k)$$

for $h,k\in L^2(\mathscr{O}),x\in L^2(\mathscr{O})$. Note that if $x\in L^p(\mathscr{O})$, then

$$D\varphi_\lambda(x)=-(1+\lambda A_0)^{-1}(((1+\lambda A_0)^{-1}x)^-)^{p-1}.$$

So, for $\lambda\to 0$ we have $\varphi_\lambda(x)\to\varphi(x)$ and $D\varphi_\lambda(x)\to D\varphi(x)$ in $L^{p/(p-1)}(\mathscr{O})$. Next we write Itô's formula for φ_λ in the space $L^2(\mathscr{O})$ which makes sense by Lemma 3.2.

We get

$$\mathbb{E}[\varphi_\lambda(X_\varepsilon(t))]+\mathbb{E}\int_0^t(A_\varepsilon(X_\varepsilon(s)),D\varphi_\lambda(X_\varepsilon(s)))ds=\varphi_\lambda(x)$$

$$+\frac{p-1}{2}\sum_{k=1}^\infty\mu_k^2\,\mathbb{E}\int_0^t\int_{\mathscr{O}}|((1+\lambda A_0)^{-1}(X_\varepsilon(s)e_k)|^2\;|((1+\lambda A_0)^{-1}X_\varepsilon(s))^-|^{p-2}d\xi\,ds.$$

This yields

$$\mathbb{E}[\varphi_\lambda(X_\varepsilon(t))]-\mathbb{E}\int_0^t\int_{\mathscr{O}}(1+\lambda A_0)^{-1}(A_\varepsilon(X_\varepsilon(s)))(((1+\lambda A_0)^{-1}X_\varepsilon(s))^-)^{p-1}d\xi ds$$

$$=\varphi_\lambda(x)$$

$$+\frac{p-1}{2}\sum_{k=1}^\infty\mu_k^2\,\mathbb{E}\int_0^t\int_{\mathscr{O}}|((1+\lambda A_0)^{-1}X_\varepsilon(s))^-|^{p-2}|(1+\lambda A_0)^{-1}(X_\varepsilon(s)e_k)|^2d\xi\,ds. \tag{3.28}$$

We know that for $\lambda\to 0$, $(1+\lambda A_0)^{-1}X_\varepsilon(s)\to X_\varepsilon(s)$ strongly in $L^p(\mathscr{O})$ a.e. in $\Omega\times(0,T)$ and

$$|(1+\lambda A_0)^{-1}X_\varepsilon|_p\le|X_\varepsilon|_p,\quad\text{a.e. in }\Omega\times(0,T).$$

Then by the Lebesgue dominated convergence theorem we have

$$\lim_{\lambda\to 0}(1+\lambda A_0)^{-1}X_\varepsilon=X_\varepsilon\quad\text{strongly in }L^p(\Omega\times(0,T)\times\mathscr{O}). \tag{3.29}$$

Similarly, since $A_\varepsilon(X_\varepsilon)\in L^p(\Omega\times(0,T)\times\mathscr{O})$ we have for $\lambda\to 0$

$$(1+\lambda A_0)^{-1}(A_\varepsilon(X_\varepsilon))\to A_\varepsilon(X_\varepsilon),\quad\text{strongly in }L^p(\Omega\times(0,T)\times\mathscr{O}).$$

and

$$((1+\lambda A_0)^{-1}X_\varepsilon)^-\to X_\varepsilon^-,\quad\text{strongly in }L^p(\Omega\times(0,T)\times\mathscr{O}).$$

This yields

$$\lim_{\lambda\to 0}\mathbb{E}\int_0^t\int_{\mathscr{O}}(1+\lambda A_0)^{-1}(A_\varepsilon(X_\varepsilon(s)))(((1+\lambda A_0)^{-1}X_\varepsilon(s))^-)^{p-1}d\xi ds$$
$$=\int_0^t\int_{\mathscr{O}}A_\varepsilon(X_\varepsilon(s))(X_\varepsilon^-(s))^{p-1}d\xi ds. \tag{3.30}$$

Then, if $x \in L^p(\mathscr{O})$ letting $\lambda \to 0$ in (3.28) we get (since by Fatou's lemma $\mathbb{E}\varphi(X_\varepsilon(t)) \le \liminf_{\lambda\to 0} \mathbb{E}\varphi_\lambda(X_\varepsilon(t)),\ \forall\, t \ge 0$)

$$\mathbb{E}[\varphi(X_\varepsilon(t))] - \mathbb{E}\int_0^t \int_{\mathscr{O}} A_\varepsilon(X_\varepsilon(s))(X_\varepsilon^-(s))^{p-1} d\xi ds$$

$$= \varphi(x) + \frac{p-1}{2} \sum_{k=1}^{\infty} \mu_k^2\, \mathbb{E}\int_0^t \int_{\mathscr{O}} |X_\varepsilon(s)e_k|^2\, |X_\varepsilon^-(s)|^{p-2} d\xi ds,$$

and so (3.27) follows. □

We have by (3.27) and the definition of Y_ε that for $x \in L^p(\mathscr{O}), x \ge 0$,

$$\mathbb{E}[\varphi(X_\varepsilon(t))] + \mathbb{E}\int_0^t \int_{\mathscr{O}} \Delta\beta(Y_\varepsilon(s))(X_\varepsilon^-(s))^{p-1} d\xi ds$$

$$= \frac{p-1}{2} \sum_{k=1}^{\infty} \mu_k^2\, \mathbb{E}\int_0^t \int_{\mathscr{O}} |X_\varepsilon^-(s)e_k|^2\, |X_\varepsilon^-(s)|^{p-2} d\xi ds$$

$$\le C\mathbb{E}\int_0^t |X_\varepsilon^-(s)|_p^p ds.$$

(Recall that $A_\varepsilon(X_\varepsilon) = -\Delta\beta(Y_\varepsilon)$.)

We therefore have, taking into account that $\Delta\beta(Y_\varepsilon) = \frac{1}{\varepsilon}(Y_\varepsilon - X_\varepsilon)$,

$$\begin{aligned}&\frac{1}{p}\, \mathbb{E}|X_\varepsilon^-(t)|_p^p + \frac{1}{\varepsilon}\, \mathbb{E}\int_0^t \int_{\mathscr{O}} (Y_\varepsilon(s) - X_\varepsilon(s))(X_\varepsilon^-(s))^{p-1} d\xi ds\\ &\le C\mathbb{E}\int_0^t |X_\varepsilon^-(s)|_p^p ds.\end{aligned} \tag{3.31}$$

We have

$$|Y_\varepsilon^-(t)|_p^p \le -\int_{\mathscr{O}} X_\varepsilon(t))(Y_\varepsilon^-(t))^{p-1} d\xi, \quad \mathbb{P}\text{-a.s.}, \tag{3.32}$$

analogously to deriving (3.24), for $x \in L^p(\mathscr{O})$. To see this multiply (3.23) by $g(y)$ where

$$g(y){:}\ = \frac{-(y^-)^{p-1}}{1 + \lambda(y^-)^{p-2}},$$

to get (after integration by parts) that

$$\int_{\mathscr{O}} \frac{(y^-)^p}{1 + \lambda(y^-)^{p-2}}\, d\xi + \varepsilon \int_{\mathscr{O}} \langle \nabla\beta(y), \nabla g(y)\rangle_{\mathbb{R}^n}\, d\xi = \int_{\mathscr{O}} \frac{-x^-(-y^-)^{p-1}}{1 + \lambda(y^-)^{p-2}}\, d\xi.$$

Note that g is Lipschitz and increasing. So, we can apply Lemma 3.1 to obtain

$$\int_{\mathscr{O}} \frac{(y^-)^p}{1+\lambda(y^-)^{p-2}}\, d\xi \le \int_{\mathscr{O}} \frac{-x^-(-y^-)^{p-1}}{1+\lambda(y^-)^{p-2}}\, d\xi$$

and (3.32) follows by taking $\lambda \to 0$. By (3.32) we have

$$-|Y_\varepsilon^-(t)|_p^p \ge \int_{\mathscr{O}} (X_\varepsilon^+(t) - X_\varepsilon^-(t))(Y_\varepsilon^-(t))^{p-1} d\xi \ge -\int_{\mathscr{O}} X_\varepsilon^-(t)(Y_\varepsilon^-(t))^{p-1} d\xi$$

and therefore $|Y_\varepsilon^-(t)|_p^p \le |X_\varepsilon^-(t)|_p \, |Y_\varepsilon^-(t)|_p^{p-1}$. Hence $|Y_\varepsilon^-(t)|_p \le |X_\varepsilon^-(t)|_p$ and so

$$\int_{\mathscr{O}} Y_\varepsilon^-(t)(X_\varepsilon^-(t))^{p-1} d\xi \le |X_\varepsilon^-(t)|_p^{p-1} \, |Y_\varepsilon^-(t)|_p \le |X_\varepsilon^-(t)|_p^p.$$

Inserting the latter into (3.31) and taking into account that $Y_\varepsilon X_\varepsilon^- \ge -Y_\varepsilon^- X_\varepsilon^-$ we see that $\mathbb{E}|X_\varepsilon^-(t)|_p^p = 0$, for a.e. $t \ge 0$ i.e, $X_\varepsilon^- = 0$ and therefore $X_\varepsilon \ge 0$ a.e. on $(0, \infty \times \mathscr{O})$. Taking into account Lemma 3.4 we infer that $X \ge 0$ a.e. on $(0, \infty \times \mathscr{O})$, hence $\mathbb{P}$-a.s. $X(t) \ge 0$ (i.e. $X(t)$ is a nonnegative measure) for all $t \ge 0$, by the continuity of the sample paths of X in $H^{-1}(\mathscr{O})$. This completes the proof in the case when β is strictly monotone. □

To treat the general case of β satisfying (1.3) we shall associate to (1.5) the equation

$$\begin{cases} dX^\lambda(t) + A^\lambda X^\lambda(t) = \sigma(X^\lambda(t))dW(t), & t \ge 0, \\ X^\lambda(0) = x, \end{cases} \tag{3.33}$$

where

$$A^\lambda(x) = -\Delta(\beta(x) + \lambda x), \quad \lambda > 0$$

and

$$D(A^\lambda) = \{x \in H^{-1}(\mathscr{O}) \cap L^1(\mathscr{O}) : \ \beta(x) + \lambda x \in H_0^1(\mathscr{O})\}$$

According to the first part of the proof, for each $x \in L^p(\mathscr{O}), x \ge 0$ and $\lambda > 0$, equation (3.33) has a unique strong solution X^λ which is nonnegative a.e. on $\Omega \times (0, T) \times \mathscr{O}$.

On the other hand, applying the Itô formula from [11, Theorem I 3.2] to the equation

$$d(X^\lambda(t) - X(t)) + (A^\lambda X^\lambda(t) - AX(t))dt = (X^\lambda(t) - X(t))dW(t)$$

where X is the solution to (1.1), we get after some calculations that

$$\frac{1}{2}\,\mathbb{E}|X^\lambda(t)-X(t)|^2_{-1}+\lambda\mathbb{E}\int_0^t\langle X^\lambda(s),X^\lambda(s)-X(s)\rangle_{-1}ds$$

$$\leq\frac{1}{2}\,\sum_{k=1}^\infty\mu_k^2\,\mathbb{E}\int_0^t|(X^\lambda(s)-X(s))e_k|^2_{-1}ds.$$

This yields (see (1.9)), since

$$\langle X^\lambda(s),X^\lambda(s)-X(s)\rangle_{-1}\geq\langle X(s),X^\lambda(s)-X(s)\rangle_{-1},$$

$$\mathbb{E}|X^\lambda(t)-X(t)|^2_{-1}\leq C\mathbb{E}\int_0^t|X^\lambda(s)-X(s)|^2_{-1}ds+\lambda^2\mathbb{E}\int_0^t|X(s)|^2_{-1}ds.$$

Since $X\in C_W([0,T];L^2(\Omega,L^2(\mathscr{O}))$, we infer via Gronwall's lemma that

$$\lim_{X^\lambda\to 0}X^\lambda=X\quad\text{in}\ \ C_W([0,T];L^2(\Omega,L^2(\mathscr{O}))$$

and so $X\geq 0$ a.e. in $\Omega\times(0,T)\times\mathscr{O}$ as claimed.

The final part of the assertion in Theorem 2.2 follows by the continuity of sample paths, since $L^p(\mathscr{O})$ is dense in $H^{-1}(\mathscr{O})$ and the continuity of solutions $X=X(t,x)$ with respect to the initial data x. Indeed by Itô, formula we see that

$$\frac{1}{2}\,\mathbb{E}|X(t,x)-X(t,y)|^2_{-1}+\mathbb{E}\int_0^t(\beta(X(s,x))-\beta(X(s,y))),X(s,x)-X(s,y))ds$$

$$\frac{1}{2}\,|x-y|^2_{-1}+\frac{1}{2}\,\mathbb{E}\sum_{j,k=1}^\infty\int_0^t\mu_k^2((X(s,x)-X(s,y))e_k,|e_j|^{-1}_{-1}e_j)^2_{-1}ds$$

$$\frac{1}{2}\,|x-y|^2_{-1}+C\mathbb{E}\int_0^t|(X(s,x)-X(s,y)|^2_{-1}ds.$$

This yields

$$\mathbb{E}|X(t,x)-X(t,y)|^2_{-1}\leq C_1|x-y|^2_{-1},$$

as claimed. □

4 Concluding remarks

1) Condition $1\leq n\leq 3$ is unnecessarily restrictive and was taken for convenience only taking into account that all physical models represented by

equation (1.1) correspond to this situations. However, Theorem 2.2 remains true for any $n \in \mathbb{N}$ assuming that instead of (1.8) the following condition holds

$$\sum_{k=1}^{\infty} \mu_k^2 (|e_k|_\infty + \lambda_k |e_k|_{\frac{4n}{n+6}})^2 < \infty. \tag{4.1}$$

Indeed, by Sobolev's embedding theorem we have

$$\begin{aligned} |xe_k|_{-1}^2 &\leq C\left(|e_k|_\infty + \lambda_k |\nabla e_k|_{\frac{4n}{n+2}}\right)^2 \\ &\leq C\left(|e_k|_\infty + \lambda_k |e_k|_{\frac{4n}{n+6}}\right)^2, \quad k \in \mathbb{N}, \end{aligned}$$

and so $\sigma(x)$, defined by (1.2) makes sense and it is in $L(L^2(\mathscr{O}), H^{-1}((\mathscr{O})))$.

The condition $n \leq 3$ was also used in the proof of Lemma 3.5 to show that $\varphi_\lambda \in C^2(L^2(\mathscr{O}))$, but, replacing φ_λ by $\varphi((1+\lambda A_0)^{-k}x)$ with k sufficiently large, the same conclusion follows for any $n \geq 1$.

2) One might speculate that Theorem 2.2 remains valid for multiplicative noise $\sigma(X)W$ where σ is a smooth nonlinear function such that $\sigma(X^+)\sigma(X^-) = 0$. However, the extension of the our arguments to this situation is not straightforward. The main difficulty is that such a function is not well defined on $H = H^{-1}(\mathscr{O})$ which is the natural space for porous media equations. An alternative way maight by to try to work on the space $L^p(\mathscr{O})$.

3) Theorem 2.2 and its proof remain valid for time–dependent nonlinear functions $\beta = \beta(t,x)$ where β is monotonically increasing in x, satisfies (1.3) uniformly with respect to t and is continuous in t.

4) One might speculate however that nonnegativity of $X(t,x)$ for $x \geq 0$ follows directly in $H^{-1}(\mathscr{O})$ by taking instead of $\varphi(x) = \frac{1}{p}|x^-|_p^p$ a suitable C^2-function on $H^{-1}(\mathscr{O})$ which is zero on the cone of positive $x \in H^{-1}(\mathscr{O})$ but so far we failed to find such a function.

References

[1] D.G. Aronson, *The porous medium equation,* Lecture Notes Math. Vol. 1224, Springer, Berlin, 1–46, 1986.

[2] V.Barbu, *Analysis and Control of Infinite Dimensional Systems,* Academic Press, 1993.

[3] V. Barbu and G. Da Prato *The two phase stochastic Stefan problem*, Probab. Theory Relat. Fields, **124**, 544–560, 2002.

[4] V. Barbu, V.I. Bogachev, G. Da Prato and M. Röckner, *Weak solution to the stochastic porous medium equations: the degenerate case,* J. Functional Analysis, **235** (2), 430-448, 2006.

[5] V.I. Bogachev, G. Da Prato and M. Röckner, *Invariant measures of stochastic porous medium type equations*, Doklady Math. Russian Acad. Sci. 396, 1, 7–11 (Russian), 2004.

[6] H. Brézis, *Operatéurs maximaux monotones*, North-Holland, 1973.

[7] G. Da Prato, M. Röckner, B. L. Rozovskii and Feng-Yu Wang, *Strong Solutions of Stochastic Generalized Porous Media Equations: Existence, Uniqueness and Ergodicity*, Comm. Partial Differential Equations **31**, no. 1-3, 277–291, 2006

[8] G. Da Prato and M. Röckner, *Weak solutions to stochastic porous media equations,* J. Evolution Equ. 4, 249–271, 2004.

[9] G. Da Prato and M. Röckner, *Invariant measures for a stochastic porous medium equation,* Advanced Studies in Pure Mathematics 41, 2004. Stochastic Analysis and related topics, 13–29, 2004.

[10] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, Cambridge University Press, 1992.

[11] N.V. Krylov and B.L. Rozovskii, *Stochastic evolution equations,* Translated from Itogi Naukii Tekhniki, Seriya Sovremennye Problemy Matematiki 14(1979), 71–146, Plenum Publishing Corp. 1981.

[12] C. Prevot and M. Röckner, *A concise course on stochastic partial differential equations,* Monograph 2006, to appear in Lecture Notes in Mathematics, Springer.

[13] J. Ren, M. Röckner and Feng-Yu Wang, *Stochastic generalized porous media and fast diffusions equations*, BiBoS-preprint 2006, to appear in J. Diff. Eqn.