Stochastic variational inequalities and applications to the total variation flow perturbed by linear multiplicative noise

Viorel Barbu* **Michael Röckner†**

Abstract

In this work, we introduce a new method to prove the existence and uniqueness of a variational solution to the stochastic nonlinear diffusion equation $dX(t) = \text{div}\left[\frac{\nabla X(t)}{|\nabla X(t)|}\right] dt + X(t)dW(t)$ in $(0, \infty) \times \mathcal{O}$, where $\mathcal{O}$ is a bounded and open domain in $\mathbb{R}^N$, $N \geq 1$, and $W(t)$ is a Wiener process of the form $W(t) = \sum_{k=1}^{\infty} \mu_k e_k \beta_k(t)$, $e_k \in C^2(\overline{\mathcal{O}}) \cap H_0^1(\mathcal{O})$, and β_k, $k \in \mathbb{N}$, are independent Brownian motions. This is a stochastic diffusion equation with a highly singular diffusivity term and one main result established here is that, for all initial conditions in $L^2(\mathcal{O})$, it is well posed in a class of continuous solutions to the corresponding stochastic variational inequality. Thus one obtains a stochastic version of the (minimal) total variation flow. The new approach developed here also allows to prove the finite time extinction of solutions in dimensions $1 \leq N \leq 3$, which is another main result of this work.

Keywords: stochastic diffusion equation, Brownian motion, bounded variation, convex functions, bounded variation flow.

AMS Subject Classification: 60H15, 35K55.

---

*Octav Mayer Institute of Mathematics (Romanian Academy), 700506 Iaşi, Romania (vb41@uaic.ro). This author was partially supported by a grant of the Romanian National Authority for Scientific Research, CNCS-UEFISCDI Project PN-II-ID-PCE-2011-3-0027 and the BiBoS – Research Center.

†Fakultät für Mathematik, Universität Bielefeld, D-33501 Bielefeld, Germany (roeckner@math.uni-bielefeld.de). This research was also supported by the DFG through CRC 701.

1 Introduction

We are concerned here with the stochastic nonlinear diffusion equation

$$\begin{aligned} & dX(t) = \mathrm{div}[\mathrm{sgn}(\nabla X(t))]dt + X(t)dW(t) \quad \text{in } (0,\infty)\times\mathcal{O}, \\ & X = 0 \quad \text{on } (0,\infty)\times\partial\mathcal{O}, \\ & X(0) = x \quad \text{in } \mathcal{O}, \end{aligned} \tag{1.1}$$

where $\mathcal{O}$ is a bounded and convex open domain in $\mathbb{R}^N$, $N \geq 1$, with smooth boundary $\partial\mathcal{O}$ and $W(t)$ is a Wiener process of the form

$$W(t) = \sum_{k=1}^{\infty} \mu_k e_k \beta_k(t), \quad t \geq 0, \quad \text{in } \mathcal{O}, \tag{1.2}$$

where μ_k are real numbers, $e_k \in C^2(\overline{\mathcal{O}}) \cap H_0^1(\mathcal{O})$ forming an orthonormal basis in $L^2(\mathcal{O})$ and $\{\beta_k\}_{k=1}^{\infty}$ are independent Brownian motions on a stochastic basis $\{\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}\}$. For simplicity, let us assume that e_k, $k \in \mathbb{N}$, are the eigenfunctions of the Dirichlet Laplacian:

$$-\Delta e_k = \lambda_k e_k \text{ in } \mathcal{O};\ e_k = 0 \text{ on } \partial\mathcal{O},$$

(but cf. Remark 2.1 (iii) below).

Throughout the paper, we assume

(H1) $C_\infty^2 := \sum_{k=1}^{\infty} \mu_k^2 |e_k|_\infty^2 < \infty,$

and

(H2) $D_\infty := \sum_{k=1}^{\infty} \mu_k |\nabla e_k|_\infty < \infty,$

where $|\cdot|_\infty$ denotes supremum norm in $C(\overline{\mathcal{O}})$.

Define

$$\mu(\xi) := \sum_{k=1}^{\infty} \mu_k^2 e_k^2(\xi),\ \xi \in \mathcal{O}. \tag{1.3}$$

The multi–valued graph $\mathrm{sgn} : \mathbb{R}^N \to 2^{\mathbb{R}^N}$ is defined by

$$\mathrm{sgn}\ r = r|r|^{-1} \ \text{ if } \ r \neq 0; \ \ \mathrm{sgn}\ 0 = \{r \in \mathbb{R}^N;\ |r| \leq 1\}, \tag{1.4}$$

and $|\cdot|$ is the Euclidean norm of $\mathbb{R}^N$. By the same symbol $|x|$, we shall denote the absolute value of $x \in \mathbb{R}$. It should be emphasized that the homogeneous boundary condition arising in (1.1) is in a certain sense formal because (1.1) is not well posed in the classical Sobolev spaces with zero trace on the boundary.

In nonlinear diffusion theory, equation (1.1) is derived from the continuity equation perturbed by a Gaussian process proportional to the density $X(t)$ of the material, that is,

$$dX(t) = \operatorname{div} J(\nabla X(t))dt + X(t)dW(t),$$

where $J = \operatorname{sgn}$ is the flux of the diffusing material. (See [23], [24], [25].)

Equation (1.1) is also relevant as a mathematical model for faceted crystal growth under a stochastic perturbation as well as in material sciences (see [26] for the deterministic model and complete references on the subject). As a matter of fact, these models are based on differential gradient systems corresponding to a convex and nondifferentiable potential (energy).

Other recent applications refer to the PDE approach to image recovery (see, e.g., [18] and also [6], [19]). In fact, if $x \in L^2(\mathcal{O})$ is the blurred image, one might find the restored image via the total variation flow $X = X(t)$ generated by the stochastic equation

$$\begin{aligned} & dX(t) = \operatorname{div}\left(\frac{\nabla X(t)}{|\nabla X(t)|}\right) dt + X(t)dW(t) \quad \text{in } (0,\infty)\times\mathcal{O}, \\ & X(0) = x \ \text{ in } \mathcal{O}. \end{aligned} \tag{1.5}$$

In its deterministic form, this is the so-called *total variation based image restoration model* and its stochastic version (1.5) arises naturally in this context as perturbation of the *total variation flow* by a Gaussian (Wiener) noise (which explains the title of the paper).

It should be said that, due to its high singularity, equation (1.1) does not have a solution in the standard sense for every initial condition in $L^2(\mathcal{O})$, that is, as an Itô integral equation, and this happens in the deterministic case, too. However, this equation has a natural formulation in the framework of stochastic variational inequalities (SVI) (see Definition 3.1 below) and, as we show later on, it is well posed in this generalized sense. Below, we shall call solutions to such (SVI) *variational solutions* and solutions to standard Itô-integral equations, as e.g. the solutions to the approximating equation (1.7) below (see Proposition 5.1 (i), *ordinary variational solutions.*

In [8], a complete existence and uniqueness result was proved for variational solutions to (1.1) in the case of additive noise, that is,

$$\begin{aligned} &dX(t) - \operatorname{div}[\operatorname{sgn}(\nabla X(t))]dt = dW(t) \quad \text{in } (0,\infty)\times\mathcal{O}, \\ &X(0) = x \ \text{ in } \mathcal{O}, \qquad X(t) = 0 \ \text{ on } (0,\infty)\times\partial\mathcal{O}, \end{aligned} \tag{1.6}$$

if $1 \leq N \leq 2$. For the multiplicative noise $X(t)dW(t)$, only the existence of a variational solution was proved and uniqueness remained open. (See, however, the work [22] for recent results on this line, if $x \in H_0^1(\mathcal{O})$.)

In this paper, we prove the existence and uniqueness of variational solutions to (1.1) in all dimensions $N \geq 1$ (see Theorem 3.2) and all initial conditions $x \in L^2(\mathcal{O})$. We would like to stress that one main difficulty is when $x \in L^2(\mathcal{O}) \setminus H_0^1(\mathcal{O})$, while the case $x \in H_0^1(\mathcal{O})$ is more standard (see Remark 3.6 below). Furthermore, we prove the finite-time extinction of solutions with positive probability, if $N \leq 3$.

The approach we use here to prove the existence and uniqueness of (1.1) is obtained approximating equation (1.1) by

$$\begin{aligned} &dX - \Delta\widetilde{\psi}_\lambda(X)dt = X\,dW \ \text{ in } (0,T)\times\mathcal{O}, \\ &X = 0 \ \text{ on } (0,T)\times\partial\mathcal{O},\ X(0) = x \ \text{ in } \mathcal{O}, \end{aligned} \tag{1.7}$$

where $\widetilde{\psi}_\lambda(r) = \psi_\lambda(r) + \lambda r$ and ψ_λ is the Yosida approximation of the graph (1.4). By the substitution $Y = e^{-W}X$ ("scaling"), we reduce (1.1) and (1.7) to a random nonlinear diffusion equation (see (4.1) and cf. [9], [12], [13]) and, again, we reformulate this random equation as a (this time, deterministic) variational inequality (VI), but with random coefficients (see Definition 4.1). This equivalent formulation of (1.1) (respectively (1.7)) as a random partial differential equation (PDE) is crucial for the uniqueness proof of variational solutions to (1.1) (see Section 5) and allows to obtain sharper regularity results for (1.7) (see, e.g., Proposition 5.1(iii) and Lemma 5.4) than those obtained by a direct analysis of the stochastic equation as in [8], [11]. This approach which combines the analysis of approximating stochastic equations in connection with their equivalent random deterministic PDE versions is by our knowledge new in the general theory of stochastic PDE and represents one principal contribution of this work.

2 Preliminaries

For every $1 \leq p \leq \infty$, by $L^p(\mathcal{O})$ we denote the space of all Lebesgue p-integrable functions on $\mathcal{O}$ with norm $|\cdot|_p$. The scalar product in $L^2(\mathcal{O})$ is denoted by $\langle \cdot, \cdot \rangle$. $W^{1,p}(\mathcal{O})$ denotes the standard Sobolev space $\{u \in L^p(\mathcal{O}); \nabla u \in L^p(\mathcal{O})\}$ with the corresponding norm

$$\|u\|_{1,p} := \left(\int_{\mathcal{O}} |\nabla u|^p d\xi \right)^{1/p},$$

where $d\xi$ denotes the Lebesgue measure on $\mathcal{O}$. $W_0^{1,p}(\mathcal{O})$ denotes the space $\{u \in W^{1,p}(\mathcal{O});\ u = 0 \text{ on } \partial\mathcal{O}\}$. We set $H_0^1(\mathcal{O}) = W_0^{1,2}(\mathcal{O})$, $\|\cdot\|_1 = \|\cdot\|_{1,2}$ and $H^2(\mathcal{O}) = \{u \in L^2(\mathcal{O}) : D_{ij}^2 u \in L^2(\mathcal{O}),\ 1 \leq i,j \leq N\}$, with its usual norm $\|\cdot\|_{H^2(\mathcal{O})}$. $H^{-1}(\mathcal{O})$ with norm $\|\cdot\|_{-1}$ denotes the dual of $H_0^1(\mathcal{O}) = W_0^{1,2}(\mathcal{O})$. By $BV(\mathcal{O})$ we denote the space of functions u of bounded variation on $\mathcal{O}$ and by $\|Du\|$ the variation of u, that is,

$$\|Du\| = \sup \left\{ \int_{\mathcal{O}} u \text{ div } \varphi d\xi;\ \varphi \in C_0^\infty(\mathcal{O}; \mathbb{R}^N),\ |\varphi|_\infty \leq 1 \right\}. \tag{2.1}$$

By $BV^0(\mathcal{O})$ we denote the space of the functions $u \in BV(\mathcal{O})$ with vanishing trace on $\partial\mathcal{O}$.

Consider the function $\phi_0 : L^1(\mathcal{O}) \to \overline{\mathbb{R}} = (-\infty, +\infty]$

$$\phi_0(u) = \begin{cases} \|Du\| & \text{if } u \in BV^0(\mathcal{O}), \\ +\infty & \text{otherwise,} \end{cases}$$

and denote by $\text{cl}\,\phi_0$ the lower semicontinuous closure of ϕ_0 in $L^1(\mathcal{O})$, that is,

$$\text{cl}\,\phi_0(u) = \inf \left\{ \liminf \phi_0(u_n);\ u_n \to u \in L^1(\mathcal{O}) \right\}. \tag{2.2}$$

As in [4, p. 437] define, for $u \in L^1(\mathcal{O})$,

$$G(u) = \begin{cases} \displaystyle\int_{\mathcal{O}} |\nabla u| d\xi & \text{if } u \in W_0^{1,1}(\Omega), \\ +\infty & \text{otherwise.} \end{cases}$$

Then (e.g., by [1, Theorem 3.9]) it is easy to see that

$$\text{cl}\,\phi_0 = \text{cl}\ G.$$

Hence, by [4, Proposition 11.3.2], for $u \in L^1(\mathcal{O})$,

$$\operatorname{cl}\phi_0(u) = \begin{cases} \|Du\| + \displaystyle\int_{\partial\mathcal{O}} |\gamma_0(u)| d\mathcal{H}^{N-1} & \text{if } u \in BV(\mathcal{O}), \\ +\infty & \text{otherwise,} \end{cases}$$

where $\gamma_0(u)$ is the trace of u on the boundary and $d\mathcal{H}^{N-1}$ is the Hausdorff measure.

Let ϕ denote the restriction of $\operatorname{cl}\phi_0(u)$ to $L^2(\mathcal{O})$, i.e.,

$$\begin{aligned} &\phi(u) = \|Du\| + \int_{\partial\mathcal{O}} |\gamma_0(u)| d\mathcal{H}^{N-1} \quad \text{if } u \in BV(\mathcal{O}) \cap L^2(\mathcal{O}), \\ &\phi(u) = +\infty \quad \text{if } u \in L^2(\mathcal{O}) \setminus BV(\mathcal{O}). \end{aligned} \tag{2.3}$$

By $\partial\phi : D(\partial\phi) \subset L^2(\mathcal{O}) \to L^2(\mathcal{O})$, we denote the subdifferential of ϕ, that is,

$$\partial\phi(u) = \{\eta \in L^2(\mathcal{O});\ \phi(u) - \phi(v) \le \langle \eta, u - v \rangle,\ \forall v \in D(\phi)\}, \tag{2.4}$$

where

$$D(\phi) = \{u \in L^2(\mathcal{O});\ \phi(u) < \infty\} = BV(\mathcal{O}) \cap L^2(\mathcal{O}).$$

It turns out (see [1]) that $\eta \in \partial\phi(u)$ iff there is $z \in L^\infty(\mathcal{O}; \mathbb{R}^N)$ such that $\eta = -\operatorname{div} z$, $|z|_\infty \le 1$, and $\int_{\mathcal{O}} \eta u d\xi = \phi(u)$.(Here and everywhere in the following the derivatives are taken in the sense of distributions on $\mathcal{O}$.) The mapping $\partial\phi$ is not everywhere defined on $D(\phi)$, but it is maximal monotone in $L^2(\mathcal{O})$ and so generates a semigroup flow $u(t,x) = e^{-t\partial\phi}x$ which is the solution to the evolution equation (see [15], p. 72, [5], p. 47)

$$\frac{du}{dt}(t) + \partial\phi(u(t)) \ni 0, \quad \forall t \ge 0,\ u(0) = x, \tag{2.5}$$

for each $x \in \overline{D(\phi)} = L^2(\mathcal{O})$. More precisely, for $x \in L^2(\mathcal{O})$, there is a unique strong solution $u : [0, \infty] \to L^2(\mathcal{O})$ to (2.5) and, for each $T > 0$,

$$\sqrt{t}\,\frac{du}{dt} \in L^2(0,T; L^2(\mathcal{O})),\ t\phi(u(t)) \in L^\infty(0,T), \phi(u) \in L^1(0,T), \tag{2.6}$$

$$t\,\frac{du}{dt} \in L^\infty(0,T; L^2(\mathcal{O})),\ u \in C([0,T]; L^2(\mathcal{O})). \tag{2.7}$$

(See [5], p. 158.) In fact, if $u \in W_0^{1,1}(\mathcal{O})$ and $\eta \in \operatorname{div}[\operatorname{sgn}(\nabla u)] \cap L^2(\mathcal{O}) \neq \emptyset$, then it is easily seen that $u \in D(\partial\phi)$ and $\eta \in \partial\phi(u)$.

We can rewrite equation (1.1) as

$$\begin{aligned}&dX(t)+\partial\phi(X(t))dt\ni X(t)dW(t),\quad t\geq 0,\\&X(0)=x.\end{aligned}\tag{2.8}$$

However, since the multi-valued mapping $\partial\phi : L^2(\mathcal{O}) \to L^2(\mathcal{O})$ is highly singular, at present no general existence result for stochastic infinite dimensional equations of subgradient type is applicable to the present situation and so a direct approach should be used in order to get existence and uniqueness of solutions for (2.8).

In the following, $L^p(0,T;E)$, $1\leq p\leq\infty$, and E a Banach space, denotes the space of all Bochner measurable functions $u:(0,T)\to E$ with $\|u\|_E\in L^p(0,T)$. By $C([0,T];E)$ we denote the space of all the continuous E-valued functions on $[0,T]$. We also use the notation

$$W^{1,p}([0,T];E)=\left\{u\in L^p(0,T;E),\ \frac{du}{dt}\in L^p(0,T;E)\right\},$$

where $\frac{du}{dt}$ is taken in sense of E-valued distributions on $(0,T)$. (We recall that any $u\in W^{1,p}([0,T];E)$ is absolutely continuous and a.e. differentiable.)

The plan of the rest of the paper is the following.

In Section 3, one defines the variational solution to (1.1) through an SVI and one formulates the main existence result which is proved in Section 5, via the mentioned scaling method. In Section 6, we prove the positivity of solutions with nonnegative initial data and, in Section 7, we prove the finite time extinction of solutions.

We close this section with some remarks on our conditions (H1), (H2) and the stochastic integral in (1.1).

Remark 2.1

(i) It is easy to check that under (H1) the sum in (1.2) converges in $L^2(\Omega;C([0,T];C(\overline{\mathcal{O}})))$ and that under (H1) and (H2) the sum in (1.2) converges in $L^2(\Omega;C([0,T];C^1(\overline{\mathcal{O}})))$. In particular, for $\mathbb{P}$-a.e. $\omega\in\Omega$ the map $[0,T]\times\overline{\mathcal{O}}\ni(t,\xi)\longmapsto W(t,\xi)(\omega)\in\mathbb{R}$ is continuous and, for each $\xi\in\mathcal{O}$, the process $(W(t,\xi))_{t\geq 0}$ is a real-valued (not standard) $(\mathcal{F}_t)$-Brownian motion with quadratic variation $\mu(\xi)t$, with μ as defined in (1.3). Furthermore, by Fernique's theorem,

$$\exp\left(\sup_{0\le t\le T}|W(t)|_\infty\right)\in L^p(\Omega)\ \text{ for all } p\in(0,\infty). \tag{2.9}$$

(ii) Let $F : [0,T]\times\mathcal{O}\times\Omega\to\mathbb{R}$ be such that F (restricted to $[0,t]$) is $\mathcal{B}([0,t])\otimes\mathcal{B}(\mathcal{O})\otimes\mathcal{F}_t$–measurable and $F\in\mathcal{L}^2([0,T]\times\mathcal{O}\times\Omega)$, where $\mathcal{L}^2$ (instead of L^2) denotes square integrable functions (rather than equivalence classes thereof). Then, for $\xi\in\mathcal{O}$, we have $\mathbb{P}$-a.s.

$$\int_0^t F(s,\xi)dW(s,\xi)=\sum_{k=1}^\infty \mu_k e_k(\xi)\int_0^t F(s,\xi)d\beta_k(s),\ t\in[0,T], \tag{2.10}$$

where the sum on the right hand side converges in $L^2(\Omega;C([0,T];\mathbb{R}))$ for each $\xi\in\mathcal{O}$ and also in $L^2(\Omega;C([0,T];L^2(\mathcal{O})))$. Indeed, defining for $N\in\mathbb{N}$

$$W_N(t,\xi):=\sum_{k=1}^N \mu_k e_k(\xi)\beta_k(t),\ t\in[0,T],\ \xi\in\mathcal{O},$$

we have for fixed $\xi\in\mathcal{O}$, $N\in\mathbb{N}$, by Doob's inequality and Itô's isometry

$$\begin{aligned}&\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t F(s,\xi)d(W-W_N)(s,\xi)\right|^2\right]\\ &\qquad\le 2\mathbb{E}\left[\int_0^T |F(s,\xi)|^2ds\right]\sum_{k=N+1}^\infty \mu_k^2 e_k^2(\xi)\end{aligned}$$

and, similarly, for $N<M$,

$$\begin{aligned}&\mathbb{E}\left[\sup_{t\in[0,T]}\int_{\mathcal{O}}\left|\sum_{k=N}^M \mu_k e_k(\xi)\int_0^t F(s,\xi)\beta_k(ds)\right|^2 d\xi\right]\\ &\qquad\le 2\sum_{k,k'=N}^M \mu_k\mu_{k'}\int_{\mathcal{O}} e_k(\xi)e_{k'}(\xi)A_{k,k'}(\xi)d\xi,\end{aligned}$$

where

$$\begin{aligned}A_{k,k'}(\xi)&:=\mathbb{E}\left[\int_0^T F(s,\xi)d\beta_k(s)\int_0^T F(s,\xi)d\beta_{k'}(s)\right]\\ &=\delta_{k,k'}\mathbb{E}\left[\int_0^T |F(s,\xi)|^2ds\right]\end{aligned}$$

since $\beta_k, \beta_{k'}$ are independent. So, both claimed convergences follow from (H1).

(iii) The assumption that e_k, $k \in \mathbb{N}$, is an eigenbasis of the Dirichlet Laplacian is only used in the proof of Proposition 5.1 (ii) below. As it is pointed out there, this assumption is not necessary, provided the initial condition x is in $H_0^1(\mathcal{O})$. Since Proposition 5.1 (ii) is only used in this paper for $x \in H_0^1(\mathcal{O})$, we may drop the above assumption on e_k, $k \in \mathbb{N}$, and just assume that it is any orthonormal basis of $L^2(\mathcal{O})$ in $C^2(\overline{\mathcal{O}}) \cap H_0^1(\mathcal{O})$.

Remark 2.2 Let $H = L^2(\mathcal{O})$ with usual inner product $\langle \cdot, \cdot \rangle$ and norm $|\cdot|_2$. Under (H1), for all $h \in H$ we have

$$\sum_{k=1}^{\infty} |\mu_k| \, |\langle h, e_k \rangle| \, |e_k|_\infty \leq C_\infty |h|_2.$$

Hence

$$\sum_{k=1}^{\infty} \mu_k \, \langle h, e_k \rangle \, e_k \in C(\overline{\mathcal{O}})$$

and, for every $x \in H$, the following operator is well-defined

$$B(x)h := x \sum_{k=1}^{\infty} \mu_k \, \langle h, e_k \rangle \, e_k \quad \left(= \sum_{k=1}^{\infty} \mu_k \, \langle h, e_k \rangle \, (e_k \cdot x) \right), \;\; h \in H.$$

It is easy to check that $B(x) \in L_2(H, H)$ (= all Hilbert-Schmidt operators from H to H) and that

$$\|B(x)\|_{L_2(H,H)} = \left(\sum_{k=1}^{\infty} \mu_k^2 |e_k x|_2^2 \right)^{1/2} \leq C_\infty |x|_2. \tag{2.11}$$

Therefore, if we consider the cylindrical Wiener process

$$\widetilde{W}(t) := (\beta_k(t) e_k)_{k \in \mathbb{N}},$$

then, it is easy to check that, if F is as in Remark 2.1 (ii), hence $(s, \omega) \mapsto F(s, \cdot, \omega) \in L^2(\mathcal{O})$ progressively measurable, then we have the following identities of $L^2(\mathcal{O})$-valued stochastic integrals in $L^2(\Omega; C([0, T]; L^2(\mathcal{O})))$

$$\int_0^{\bullet} F(s) dW(s) = \int_0^{\bullet} B(X(s)) d\widetilde{W}(s) = \sum_{k=1}^{\infty} \mu_k \int_0^{\bullet} F(s) e_k d\beta_k(s), \tag{2.12}$$

where also the sum on the right hand side converges in $L^2(\Omega; C([0,T]; L^2(\mathcal{O})))$. In particular, the stochastic integral in (1.1) is a standard one. An easy application of the stochastic Fubini Theorem (cf. the proof of Claim 2 in the proof of Proposition 4.3) then shows that, by (2.12) and Remark 2.1 (ii), for every $t \in [0,T]$ and $\mathbb{P}$-a.e. $\omega \in \Omega$,

$$\xi \mapsto \int_0^t F(s,\xi)dW(s,\xi)(\omega)$$

(which is a real-valued stochastic integral) is a $d\xi$-version of $\int_0^t F(s)dW(s)$ (which is an $L^2(\mathcal{O})$-valued stochastic integral).

3 Definition of stochastic variational solutions and the main existence and uniqueness result

Definition 3.1 Let $0 < T < \infty$ and let $x \in L^2(\mathcal{O})$. A stochastic process $X : [0,T] \times \Omega \to L^2(\mathcal{O})$ is said to be a variational solution to (1.1) if the following conditions hold.

(i) X is $(\mathcal{F}_t)$-adapted, has $\mathbb{P}$-a.s. continuous sample paths in $L^2(\mathcal{O})$ and $X(0) = x$.

(ii) $X \in L^2([0,T] \times \Omega; L^2(\mathcal{O}))$, $\phi(X) \in L^1([0,T] \times \Omega)$.

(iii) For each $(\mathcal{F}_t)$- progressively measurable process $G \in L^2([0,T]\times\Omega; L^2(\mathcal{O}))$ and each $(\mathcal{F}_t)$-adapted $L^2(\mathcal{O})$-valued process Z with $\mathbb{P}$-a.s. continuous sample paths such that $Z \in L^2([0,T] \times \Omega; H_0^1(\mathcal{O}))$ and, solving the equation

$$Z(t) - Z(0) + \int_0^t G(s)ds = \int_0^t Z(s)dW(s), \quad t \in [0,T], \tag{3.1}$$

we have

$$\begin{aligned} &\frac{1}{2}\,\mathbb{E}|X(t) - Z(t)|_2^2 + \mathbb{E}\int_0^t \phi(X(\tau))d\tau \le \frac{1}{2}\,\mathbb{E}|x - Z(0)|_2^2 \\ &\quad +\mathbb{E}\int_0^t \phi(Z(\tau))d\tau + \frac{1}{2}\,\mathbb{E}\int_0^t\int_{\mathcal{O}} \mu(X(\tau) - Z(\tau))^2 d\xi\, d\tau \\ &\quad +\mathbb{E}\int_0^t \langle X(\tau) - Z(\tau), G(\tau)\rangle\, d\tau, \; t \in [0,T]. \end{aligned} \tag{3.2}$$

Here, ϕ is defined by (2.3), $\mu = \sum_{k=1}^{\infty} \mu_k^2 e_k^2$ and $\langle \cdot, \cdot \rangle$ is the duality pairing with pivot space $L^2(\mathcal{O})$. We also recall that (3.1) has a unique solution for a given initial condition in $L^2(\mathcal{O})$.

The relationship between (1.1) and (3.2) becomes more transparent if we recall that (1.1) can be rewritten as (2.8) and so we have

$$d(X - Z) + (\partial\phi(X) - G)dt \ni (X - Z)dW. \tag{3.3}$$

If we (formally) apply the Itô formula to $\frac{1}{2}|X - Z|_2^2$ in (3.3) and take into account (2.4), we obtain just (3.2) after taking expectation. It should be emphasized, however, that X arising in Definition 3.1 is not a strong solution to (1.1) (or (2.8)) in the standard sense, that is,

$$X(t) - x \in -\int_0^t \partial\phi(X(s))ds + \int_0^t X(s)dW(s),\ \forall t \in [0, T].$$

We also note that this concept of solution for equation (1.1) was already introduced in [8]. Theorem 3.2 below is our first main result.

Theorem 3.2 *Let $\mathcal{O}$ be a bounded and convex open subset of $\mathbb{R}^N$ with smooth boundary and $T > 0$. For each $x \in L^2(\mathcal{O})$ there is a variational solution X to equation* (1.1)*, such that, for all $p \in [2, \infty)$,*

$$\sup_{t\in[0,T]} \mathbb{E}[|X(t)|_2^p] \leq \exp\left[C_\infty^2 \frac{p}{2}(p-1)\right] \|x\|_2^p. \tag{3.4}$$

X is the unique solution in the class of all solutions X such that, for some $\delta > 0$,

$$X \in L^{2+\delta}(\Omega; L^2([0, T]; L^2(\mathcal{O}))). \tag{3.5}$$

Furthermore, if $x, x^* \in L^2(\mathcal{O})$ and X, X^* are the corresponding variational solutions with initial conditions x, x^*, respectively, then, for some positive constant $C = C(N, C_\infty^2)$,

$$\mathbb{E}\left[\sup_{\tau\in[0,T]} |X(\tau) - X^*(\tau)|_2^2\right] \leq 2|x - x^*|_2^2 e^{CT}. \tag{3.6}$$

In particular,

$$\mathbb{E}\left[\sup_{t\in[0,T]}|X(t)|_2^2\right]\le 2|x|_2^2e^{CT}, \tag{3.7}$$

and, moreover, $X\in L^2(\Omega;C([0,T];L^2(\mathcal{O})))$.

Remark 3.3 A similar result was established in [8] for the equation with additive noise where $N=1,2$. However, in the definition of the solution in [8], erroneously was taken the functional ϕ_0 instead of ϕ defined above, as it is correct. In this context we cite also the work [11], where this point was already clarified. Furthermore, by Remark 8.4 below, the convexity assumption on $\mathcal{O}$ can be relaxed. It is enough that $\partial\mathcal{O}$ can be parametrized locally by a convex C^2-map.

Remark 3.4 Apparently, the variational solution X defined above does not satisfy in any common sense the Dirichlet homogeneous condition on $\partial\mathcal{O}$, as written in (1.1). However, since $\phi(X)\in L^1((0,T)\times\Omega)$ and, as seen earlier, ϕ is just the closure in $L^2(\mathcal{O})$ of ϕ_0 and, in particular, of the norm $|\nabla u|_1$ of the space $W_0^{1,1}(\mathcal{O})$, we may regard X as a generalized solution to (1.1). For instance, if in (1.1) we replace the Dirichlet condition by the Neumann homogeneous condition, then in the above definition of the variational solution one should replace the function ϕ by

$$\phi_1(u)=||Du||,\ \forall u\in BV(\mathcal{O})\cap L^2(\mathcal{O});\ \ \phi_1(u)=+\infty \text{ otherwise.}$$

Remark 3.5 It follows from Lemma 7.3 below by Fatou's Lemma that, for $N\le 3$, in addition to (3.6) we also have, for some constant $C>0$,

$$\mathbb{E}\left[\sup_{\tau\in[0,T]}|X(\tau)-X^*(\tau)|_N^N\right]\le 2|x-x^*|_N^Ne^{CT}. \tag{3.8}$$

Remark 3.6 If $x\in H_0^1(\mathcal{O})$, it follows by Lemma 5.3 and Fatou's Lemma that, for some $C>0$ (independent of x),

$$\mathbb{E}\left[\sup_{t\in[0,T]}\|X(t)\|_1^2\right]\le C\|x\|_1^2,$$

hence $X\in L^2(\Omega;L^\infty([0,T];H_0^1(\mathcal{O})))$. From this, one can deduce that, if the initial condition x is in $H_0^1(\mathcal{O})$, then the corresponding solution X in

Theorem 3.2 is, in fact, an ordinary variational solution of the (multivalued) equation (1.1) (not just in the sense of SVI as in Definition 3.1). Our main point is, however, here to have existence and uniqueness for all starting points $x \in L^2(\mathcal{O})$. Therefore, we skip the details on the simpler case of special initial conditions in $H_0^1(\mathcal{O})$.

4 The random partial differential equation equivalent to (1.1)

Inspired by [9, Section 4] and [12], we would like to reduce equation (1.1) to the random differential equation

$$\begin{aligned} &\frac{\partial Y}{\partial t} = e^{-W}\mathrm{div}(\mathrm{sgn}(\nabla(e^W Y))) - \frac{1}{2}\,\mu Y, \ \mathbb{P}\text{-a.s. in } (0,T)\times\mathcal{O}, \\ &Y(0,\xi) = x(\xi), \quad \xi \in \mathcal{O}, \\ &Y = 0 \text{ on } (0,T)\times\partial\mathcal{O}, \end{aligned} \tag{4.1}$$

by the substitution $Y(t) = e^{-W(t)}X(t)$. The meaning of boundary condition in (4.1) is taken in the generalized sense as discussed in Remark 3.4. (We note that, in equation (1.1), XdW is meant to be an Itô differential, otherwise, i.e., if it is taken in the Stratonovich sense, then, in the corresponding equation (4.1), the linear term $\frac{1}{2}\,\mu Y$ would be missing.)

To do the reduction from (1.1) to (4.1) rigorously, our definitions of solutions for (4.1) must be again in the sense of a variational inequality, but this time a deterministic one, since the test processes $\widetilde{Z}$ (replacing Z in Definition 3.1) solve a deterministic PDE, however, with random coefficients.

Definition 4.1 Let $0 < T < \infty$ and let $x \in L^2(\mathcal{O})$. A stochastic process $Y : [0,T]\times\Omega \to L^2(\mathcal{O})$ is said to be a *variational solution* to (4.1) if the following conditions hold:

(i) Y is $(\mathcal{F}_t)$-adapted, has $\mathbb{P}$-a.s. continuous sample paths, and $Y(0) = x$.

(ii) $e^W Y \in L^2([0,T]\times\Omega; L^2(\mathcal{O}))$, $\phi(e^W Y) \in L^1([0,T]\times\Omega)$.

(iii) For each $(\mathcal{F}_t)$-progressively measurable process $G \in L^2([0,T]\times\Omega; L^2(\mathcal{O}))$ and each $(\mathcal{F}_t)$-adapted, $L^2(\mathcal{O})$-valued process $\widetilde{Z}$ with $\mathbb{P}$-a.s. continuous sample paths such that $e^W\widetilde{Z} \in L^2([0,T]\times\Omega; H_0^1(\mathcal{O}))$ and solving the equation

$$\widetilde{Z}(t)-\widetilde{Z}(0)+\int_0^t e^{-W(s)}G(s)ds+\frac{1}{2}\int_0^t \mu\widetilde{Z}(s)ds=0, \qquad t\in[0,T],\ \mathbb{P}\text{-a.s.}, \tag{4.2}$$

we have

$$\begin{aligned}
&\frac{1}{2}\,\mathbb{E}|e^{W(t)}(Y(t)-\widetilde{Z}(t))|_2^2+\mathbb{E}\int_0^t \phi(e^{W(\tau)}Y(\tau))d\tau\\
&\quad\le \frac{1}{2}\,\mathbb{E}|x-\widetilde{Z}(0)|_2^2+\mathbb{E}\int_0^t \phi(e^{W(\tau)}\widetilde{Z}(\tau))d\tau\\
&\quad+\frac{1}{2}\,\mathbb{E}\int_0^t\int_{\mathcal{O}}\mu e^{2W(\tau)}(Y(\tau)-\widetilde{Z}(\tau))^2d\xi\,d\tau\\
&\quad+\mathbb{E}\int_0^t \left\langle e^{W(\tau)}(Y(\tau)-\widetilde{Z}(\tau)),G(\tau)\right\rangle d\tau,\ t\in[0,T].
\end{aligned}\tag{4.3}$$

We recall that the deterministic equation (4.2) has a unique solution for a given initial condition in $L^2(\mathcal{O})$ for $\mathbb{P}$-a.e. given $\omega\in\Omega$.

Proposition 4.2 *$X:[0,T]\times\Omega\to L^2(\mathcal{O})$ is a variational solution to equation* (1.1) *if and only if $Y:=e^{-W}X$ is a variational solution to* (4.1).

The above proposition is an immediate consequence of Proposition 4.3(iii) below, which addresses a technical, but very important issue. To be precise (and make our point) in its proof, we have to distinguish between the space $\mathcal{L}^2(\mathcal{O})$ of square integrable functions and $L^2(\mathcal{O})$, i.e., the corresponding $d\xi$-classes.

Proposition 4.3 *Let $G\in L^2([0,T]\times\Omega;L^2(\mathcal{O}))$ be $(\mathcal{F}_t)$-progressively measurable and $Z(0)\in L^2(\Omega,\mathcal{F}_0;L^2(\mathcal{O}))$. Let G^0 be a $(dt\otimes d\xi\otimes\mathbb{P})$-version of G such that $(t,\omega)\longmapsto G^0(t,\xi,\omega)$ is $(\mathcal{F}_t)$-progressively measurable and in $L^2([0,T]\times\Omega)$ for every $\xi\in\mathcal{O}$. Furthermore, let $Z^0(0)$ be a $(d\xi\otimes\mathbb{P})$-version of $Z(0)$ such that $\omega\longmapsto Z^0(0)(\xi,\omega)$ is $\mathcal{F}_0$-measurable for all $\xi\in\mathcal{O}$.*

(i) *Fix $\xi\in\mathcal{O}$. Then*

$$\begin{aligned}
Z_\xi^0(t)&:=e^{W(t,\xi)-\frac{1}{2}\mu(\xi)t}Z^0(0)(\xi)\\
&\quad-e^{W(t,\xi)-\frac{1}{2}\mu(\xi)t}\int_0^t e^{-W(s,\xi)+\frac{1}{2}\mu(\xi)s}G^0(s,\xi)ds,\ t\in[0,T],
\end{aligned}\tag{4.4}$$

is a real-valued continuous solution to the stochastic differential equation

$$\begin{aligned} dZ^0_\xi(t) &= -G^0(t,\xi)dt + Z^0_\xi(t)dW(t,\xi),\ t\in[0,T],\\ Z^0_\xi(0) &= Z^0(0,\xi), \end{aligned} \tag{4.5}$$

which is $\mathcal{B}([0,t])\otimes\mathcal{B}(\mathcal{O})\otimes\mathcal{F}_t$*-measurable for each* $t\in[0,T]$.

Furthermore, the map $[0,T]\ni t\longmapsto Z^0_\cdot(t)\in\mathcal{L}^2(\mathcal{O})$ *is* $\mathbb{P}$*-a.s. continous. Hence the corresponding* $d\xi$*-classes* $Z(t)\in L^2(\mathcal{O})$, $t\in[0,T]$, *form the unique solution to* (3.1).

(ii) *Fix* $\xi\in\mathcal{O}$. *Then*

$$\widetilde{Z}^0_\xi(t) := e^{-\frac12\mu(\xi)t}Z^0(0)(\xi) - e^{-\frac12\mu(\xi)t}\int_0^t e^{-W(s,\xi)+\frac12\mu(\xi)s}G^0(s,\xi)ds,\qquad t\in[0,T],$$

is a real-valued continuous solution to the differential equation

$$d\widetilde{Z}^0_\xi(t) = -e^{-W(t,\xi)}G^0(t,\xi)dt - \frac12\,\mu(\xi)\widetilde{Z}^0_\xi(t)dt,\ \ \widetilde{Z}^0_\xi(0) = Z^0(0,\xi),$$

which is $\mathcal{B}([0,t])\otimes\mathcal{B}(\mathcal{O})\otimes\mathcal{F}_t$*-measurable for each* $t\in[0,T]$.

Furthermore, the map $[0,T]\ni t\longmapsto \widetilde{Z}^0_\cdot(t)\in\mathcal{L}^2(\mathcal{O})$ *is* $\mathbb{P}$*-a.s. continuous. Hence the corresponding* $d\xi$*-classes* $\widetilde{Z}_\cdot(t)\in L^2(\mathcal{O})$, $t\in[0,T]$, *form the unique solution of the deterministic equation* (4.4) *for* $\mathbb{P}$*-a.e. given* $\omega\in\Omega$.

(iii) *An* $(\mathcal{F}_t)$*-adapted* $\mathbb{P}$*-a.s. continuous* $L^2(\mathcal{O})$*-valued process* $(Z(t))_{t\in[0,T]}$ *is a solution to the stochastic equation* (3.1) *if and only if* $(e^{-W(t)}Z(t))_{t\in[0,T]}$ *is a solution to the deterministic equation* (4.1) *for* $\mathbb{P}$*-a.e. given* $\omega\in\Omega$.

Proof. (iii) is an immediate consequence of (i) and (ii). (ii) is more or less well-known since it is about a deterministic equation and the proof is anyway similar to that of (i). Therefore, we only prove (i).

First, we note that applying a mollifier in ξ and taking the limsup of a properly chosen subsequence, the mentioned version of G^0 and $Z^0(0)$ always exist. Obviously, Z^0_ξ is a well-defined, $(\mathcal{F}_t)$-adapted, $\mathbb{P}$-a.s. continuous real-valued process, and applying Itô's product formula we obtain that it solves (4.5). Furthermore, the stated continuity in $\mathcal{L}^2(\mathcal{O})$ is obvious. So, it remains to show the last part of the assertion, which follows from the following two claims.

Claim 1. Let $t \in [0,T]$. Then $\mathbb{P}$-a.e. $\xi \longmapsto \int_0^t G^0(s,\xi)ds$, $\xi \in \mathcal{O}$, is a $d\xi$-version of the $L^2(\mathcal{O})$-valued Bochner integral $\int_0^t G(s)ds$.

Claim 2. Let $t \in [0,T]$. Then $\mathbb{P}$-a.s. $\xi \longmapsto \int_0^t Z^0_\xi(s)dW(s,\xi)$, $\xi \in \mathcal{O}$, is a $d\xi$-version of the $L^2(\mathcal{O})$-valued stochastic integral $\int_0^t Z(s)dW(s)$.

Claim 1 is a trivial consequence of Fubini's theorem. So, we only prove Claim 2 whose proof is similar, but based on the stochastic Fubini theorem.

Proof of Claim 2. Let $i \in \mathbb{N}$. Then $\mathbb{P}$-a.s. for every $t \in [0,T]$, by Remark 2.1 (ii),

$$\begin{aligned}
&\int_{\mathcal{O}} e_i(\xi) \int_0^t Z^0_\xi(s)dW(s,\xi)d\xi \\
&= \sum_{k=1}^{\infty} \mu_k \int_{\mathcal{O}} e_i(\xi)e_k(\xi) \int_0^t Z^0_\xi(s)d\beta_k(s)d\xi \\
&= \sum_{k=1}^{\infty} \mu_k \int_0^t \int_{\mathcal{O}} e_i(\xi)e_k(\xi)Z^0_\xi(s)d\xi d\beta_k(s) \\
&= \sum_{k=1}^{\infty} \mu_k \int_0^t \langle e_i, e_k Z(s)\rangle \, d\beta_k(s) \\
&= \sum_{k=1}^{\infty} \mu_k \left\langle e_i, \int_0^t e_k Z(s)d\beta_k(s) \right\rangle \\
&= \left\langle e_i, \int_0^t Z(s)dW(s) \right\rangle,
\end{aligned}$$

where we used the stochastic Fubini theorem in the second equality. Now, Claim 2 follows. $\square$

Remark 4.4 Proposition 4.3 justifies to apply Itô's formula for a solution $Z(t)$, $t \in [0,T]$, to (3.1) for $d\xi$-a.e. $\xi \in \mathcal{O}$ to the process $Z(t)(\xi)$, $t \in [0,T]$, by taking the version $Z^0_\xi(t)$, $t \in [0,T]$, from Proposition 4.3(i). We stress that for Proposition 4.3 we only used (H1), not (H2) (see Remark 2.1).

In particular, by Theorem 3.2, Proposition 4.2 and (5.12) below, we have the following existence result for (4.1), which has an intrinsic interest.

Proposition 4.5 *Under the assumptions of Theorem* 3.2*, for each* $x \in L^2(\mathcal{O})$*, there is a variational solution* Y *to* (4.1)*, which is unique in the class of all solutions* Y *such that, for some* $\delta > 0$*,*

$$Y \in L^{2+\delta}(\Omega; L^2([0,T]; L^2(\mathcal{O}))).$$

Moreover,

$$Y \in L^2(\Omega; C([0,T]; L^2(\mathcal{O}))).$$

5 Proof of Theorem 3.2

It should be said that, for the proof of the uniqueness part of Theorem 3.2, as well as for the finite-time extinction property of the solutions to (1.1), it is convenient and apparently necessary to replace (1.1) by (4.1) and to construct approximating schemes for both equations.

We approximate (1.1) by

$$\begin{aligned} &dX_\lambda = \mathrm{div}\ \widetilde{\psi}_\lambda(\nabla X_\lambda)dt + X_\lambda dW \ \text{ in } \ (0,T)\times\mathcal{O},\\ &X_\lambda(0) = x \ \text{ in } \ \mathcal{O},\\ &X_\lambda = 0 \ \text{ on } \ (0,T)\times\partial\mathcal{O}, \end{aligned} \tag{5.1}$$

and the corresponding rescaled equation (4.1) by

$$\begin{aligned} &\frac{dY_\lambda}{dt} = e^{-W}\mathrm{div}(\widetilde{\psi}_\lambda(\nabla(e^W Y_\lambda))) - \frac{1}{2}\,\mu Y_\lambda\\ &\qquad\qquad\qquad\qquad \text{in } (0,T)\times\mathcal{O},\ \mathbb{P}\text{-a.s.},\\ &Y_\lambda(0) = x \text{ in } \mathcal{O},\quad Y_\lambda = 0 \text{ on } (0,T)\times\partial\mathcal{O}, \end{aligned} \tag{5.2}$$

where $\lambda \in (0,1]$, $\widetilde{\psi}_\lambda(u) = \psi_\lambda(u) + \lambda u,\ \forall u \in \mathbb{R}^N$.

In (5.2), $\frac{d}{dt}Y_\lambda \in L^2(0,T;H^{-1}(\mathcal{O}))$ is the strong derivative of $t \to Y_\lambda(t)$ and the operator div is taken in sense of distributions on $\mathcal{O}$.

Here, ψ_λ is the Yosida approximation of the function $\psi(u) = \mathrm{sgn}\ u$, that is (see, e.g., [5]),

$$\psi_\lambda(u) = \begin{cases} \dfrac{1}{\lambda}\,u & \text{if } |u| \le \lambda,\\ \dfrac{u}{|u|} & \text{if } |u| > \lambda. \end{cases} \tag{5.3}$$

Let $j_\lambda(u) = \inf\limits_v \left\{ \dfrac{|u-v|^2}{2\lambda} + |v| \right\}$ be the Moreau–Yosida approximation of the function $v \to |v|$. We recall that $\nabla j_\lambda = \psi_\lambda$, $\forall \lambda > 0$ (see, e.g., [5], p. 48). We first prove the existence of a strong solution Y_λ to (5.2).

It should be emphasized that, for the existence and uniqueness part of the proof, it is convenient to analyze equation (5.1) while, for getting sharp estimates on the variational solutions X to (1.1), it is necessary to work directly with the random equation (5.2) instead of (5.1). As regards the existence and uniqueness for (5.1), (5.2), we have:

Proposition 5.1

(i) *For each $\lambda \in (0,1]$ and each $x \in L^2(\mathcal{O})$, there is a unique strong solution X_λ to (5.1) which satisfies $X_\lambda(0) = x$, that is, X_λ is $\mathbb{P}$-a.s. continuous in $L^2(\mathcal{O})$ and $\{\mathcal{F}_t\}$-adapted such that*

$$\begin{aligned} &X_\lambda \in L^2([0,T]\times\Omega; H_0^1(\mathcal{O})),\\ &X_\lambda(t) = x + \int_0^t \operatorname{div} \widetilde{\psi}_\lambda\,(\nabla X_\lambda(s))\,ds + \int_0^t X_\lambda(s)dW(s), \qquad (5.4)\\ &\qquad\qquad t\in[0,T],\ \mathbb{P}\text{-a.s.} \end{aligned}$$

Furthermore, $X_\lambda \in L^2(\Omega; C([0,T]; L^2(\mathcal{O})))$ and, for all $p \in [2,\infty)$,

$$\sup_{t\in[0,T]} \mathbb{E}\,[|X_\lambda(t)|_2^p] \le \exp\left[C_\infty^2\,\frac{p}{2}\,(p-1)\right] |x|_2^p, \qquad (5.5)$$

and, if $x, x^* \in L^2(\mathcal{O})$ and X_λ and X_λ^* are the corresponding solutions with initial conditions x, x^*, respectively, then, for some positive constant $C = C(N, C_\infty^2)$,

$$\mathbb{E}\left[\sup_{\tau\in[0,T]} |X_\lambda(\tau) - X_\lambda^*(\tau)|_2^2\right] \le 2|x-x^*|_2^2 e^{CT}. \qquad (5.6)$$

(ii) *$Y_\lambda = e^{-W} X_\lambda$ is an $(\mathcal{F}_t)$-adapted process $Y_\lambda : [0,T]\times\Omega \to L^2(\mathcal{O})$ with $\mathbb{P}$-a.s. continuous paths which is the unique solution of (5.2), i.e., it satisfies $\mathbb{P}$-a.s. equation (5.2) with $Y_\lambda(0) = x$ and*

$$Y_\lambda \in L^2([0,T]; H_0^1(\mathcal{O})) \cap C([0,T]; L^2(\mathcal{O})) \cap W^{1,2}([0,T]; H^{-1}(\mathcal{O})), \qquad (5.7)$$

a.e. $t \in [0,T]$.

(iii) *If $x \in H_0^1(\mathcal{O})$, then $\mathbb{P}$-a.s.*

$$X_\lambda \in C([0,T]; H_0^1(\mathcal{O})) \tag{5.8}$$

and

$$X_\lambda \in L^2(\Omega; L^\infty([0,T]; H_0^1(\mathcal{O}))) \cap L^2([0,T] \times \Omega; H^2(\mathcal{O})). \tag{5.9}$$

Remark 5.2 It is readily seen that, by Itô's formula, X_λ is also a variational solution to (5.1) in the sense of Definition 3.1, where $\phi(y)$ is replaced by

$$\widetilde{\phi}_\lambda(y) = \int_{\mathcal{O}} \left(j_\lambda(\nabla y) + \frac{\lambda}{2} |\nabla y|^2 \right) d\xi.$$

Proof of Proposition 5.1. Consider the operator $A_\lambda : H_0^1(\mathcal{O}) \to H^{-1}(\mathcal{O})$ defined by

$$\langle A_\lambda y, \varphi \rangle = \int_{\mathcal{O}} \widetilde{\psi}_\lambda(\nabla y) \cdot \nabla \varphi d\xi, \ \forall \varphi \in H_0^1(\mathcal{O}), \tag{5.10}$$

and note that A_λ is demicontinuous (see, for instance, [5], p. 81).

Moreover, we have

$$\begin{aligned} \|A_\lambda y\|_{-1} &\leq \lambda \|y\|_1 + \left(\int_{\mathcal{O}} d\xi \right)^{\frac{1}{2}}, \quad && \forall y \in H_0^1(\mathcal{O}), \\ \langle A_\lambda y_1 - A_\lambda y_2, y_1 - y_2 \rangle &\geq \lambda \|y_1 - y_2\|_1^2, && \forall y \in H_0^1(\mathcal{O}). \end{aligned}$$

On the other hand, equation (5.1) can be rewritten as

$$\begin{aligned} & dX_\lambda + A_\lambda X_\lambda dt = X_\lambda dW, \ t \in [0,T], \\ & X_\lambda(0) = x. \end{aligned} \tag{5.11}$$

Then, by the standard existence theory for stochastic differential equations associated with nonlinear monotone and demicontinuous operators in a duality pair (V, V') ([27], [29], [30]) equation (5.11) (equivalently, (5.1)) has a unique strong solution X_λ satisfying (5.4) and (5.6). (5.5) is then an easy consequence of Itô's formula for $|X_\lambda|_2^2$ (see, e.g., [30]).

To prove (ii), below we use $\langle \cdot, \cdot \rangle_2$ to denote the inner product in $L^2(\mathcal{O})$, in order to avoid confusion with the quadratic variation process.

Let $\varphi \in H_0^1(\mathcal{O}) \cap L^\infty(\mathcal{O})$. Then, for every $t \in [0,T]$,

$$\left\langle \varphi, e^{-W(t)} X_\lambda(t) \right\rangle_2 = \sum_{j=1}^{\infty} \left\langle e_j, e^{-W(t)} \varphi \right\rangle_2 \langle e_j, X_\lambda(t) \rangle_2 .$$

Furthermore, by Itô's formula we have $d\xi \otimes \mathbb{P}$-a.e. that, for all $\xi \in \mathcal{O}$, $t \in [0,T]$,

$$e^{-W(t,\xi)} = 1 - \int_0^t e^{-W(s,\xi)} dW(s,\xi) + \frac{1}{2}\mu(\xi) \int_0^t e^{-W(s,\xi)} ds.$$

Now, fix $j \in \mathbb{N}$. Then, by Remark 2.1 (ii) and Remark 2.2, we have $\mathbb{P}$-a.e. that, for all $t \in [0,T]$,

$$\begin{aligned}
\left\langle e_j, e^{-W(t)} \varphi \right\rangle_2 &= \langle e_j, \varphi \rangle_2 - \sum_{k=1}^{\infty} \mu_k \int_{\mathcal{O}} e_j(\xi)\varphi(\xi)e_k(\xi) \int_0^t e^{-W(s,\xi)} d\beta_k(s) d\xi \\
&\quad + \frac{1}{2} \int_0^t \left\langle e_j, \mu e^{-W(s)} \varphi \right\rangle_2 ds \\
&= \langle e_j, \varphi \rangle - \sum_{k=1}^{\infty} \mu_k \int_0^t \left\langle e_j, e_k e^{-W(s)} \varphi \right\rangle_2 d\beta_k(s) \\
&\quad + \frac{1}{2} \int_0^t \left\langle e_j, \mu e^{-W(s)} \varphi \right\rangle_2 ds,
\end{aligned}$$

where we used the stochastic Fubini Theorem in the second equality and the sums converge in $L^2(\Omega; C([0,T];\mathbb{R}))$. By Itô's product rule we hence obtain $\mathbb{P}$-a.s. that, for all $t \in [0,T]$,

$$\begin{aligned}
&\left\langle e_j, e^{-W(t)} \varphi \right\rangle_2 \langle e_j, X_\lambda(t) \rangle_2 = \langle e_j, \varphi \rangle_2 \langle e_j, x \rangle_2 \\
&\quad + \int_0^t \left\langle e_j, e^{-W(s)} \varphi \right\rangle_2 \left\langle e_j, \operatorname{div} \widetilde{\psi}_\lambda(\nabla X_\lambda(s)) \right\rangle ds \\
&\quad + \sum_{k=1}^{\infty} \mu_k \int_0^t \left\langle e_j, e^{-W(s)} \varphi \right\rangle_2 \langle e_j, X_\lambda(s) e_k \rangle_2 d\beta_k(s) \\
&\quad - \sum_{k=1}^{\infty} \mu_k \int_0^t \left\langle e_j, e_k e^{-W(s)} \varphi \right\rangle_2 \langle e_j, X_\lambda(s) \rangle_2 d\beta_k(s)
\end{aligned}$$

$$
+\frac{1}{2}\int_0^t \left\langle e_j, \mu e^{-W(s)}\varphi\right\rangle_2 \langle e_j, X_\lambda(s)\rangle_2\, ds
$$

$$
-\sum_{k=1}^{\infty}\mu_k^2 \int_0^t \langle e_j, X_\lambda(s)e_k\rangle_2 \left\langle e_j, e_k e^{-W(s)}\varphi\right\rangle_2 ds,
$$

where all the sums converge in $L^2(\Omega; C([0,T];\mathbb{R}))$ and interchanging the infinite sums with stochastic differentials is justified by Remark 2.1 (ii) and Remark 2.2, because of (5.5) and since, by (2.9),

$$
\sup_{(t,\xi)\in[0,T]\times\mathcal{O}} e^{-W(s,\xi)}\ |X_\lambda|_2 \in L^p([0,T]\times\Omega;\mathbb{R}) \text{ for all } p \geq 1. \tag{5.12}
$$

(We shall implicitly use both (5.5) and (5.12) several times in the rest of this paper without further notice.)

Now, we sum the above equation from $j = 1$ to $j = \infty$ and assume that we can interchange this summation both with the sum over k and with the deterministic and stochastic integrals (which we shall justify below). Then, because the two terms involving the stochastic integrals cancel, we obtain

$$
\left\langle \varphi, e^{-W(t)}X_\lambda(t)\right\rangle_2 = \langle \varphi, x\rangle_2 + \int_0^t \left\langle \varphi, e^{-W(s)}\text{div}\,\widetilde{\psi}_\lambda(\nabla X_\lambda(s))\right\rangle ds
$$

$$
+\frac{1}{2}\int_0^t \left\langle \varphi, \mu e^{-W(s)}X_\lambda(s)\right\rangle_2 ds - \sum_{k=1}^{\infty}\mu_k^2\int_0^t \left\langle \varphi, e_k^2 e^{-W(s)}X_\lambda(s)\right\rangle_2 ds,
$$

which immediately implies that $Y_\lambda = e^{-W}X_\lambda$ solves (5.2).

To justify interchanging sums and integrals, it suffices to note that, for the second term on the right hand side, this is true because $\{e_k\}$ is the eigenbasis of the Laplacian and that for the last term this is obvious because of (H1), while, for the two terms which cancel each other and involve stochastic integrals, this follows by applying the Burkholder–Davis–Gundy inequality and (H1). If, however, $x \in H_0^1(\mathcal{O})$, then, by Lemma 5.3 below, $\text{div}\,\widetilde{\psi}_\lambda(\nabla X_\lambda) \in L^2([0,T]\times\Omega; L^2(\mathcal{O}))$ (and not only in $L^2([0,T]\times\Omega; H^{-1}(\mathcal{O}))$). Hence, the above equality is true for any orthonormal basis e_k, $k \in \mathbb{N}$, of $L^2(\mathcal{O})$ in $C^2(\overline{\mathcal{O}})$.

It remains to prove the uniqueness. In fact, as it will be explained below, by standard methods one can prove directly the existence and uniqueness of a solution Y_λ to (5.2), which hence must be of the form $Y_\lambda = e^{-W}X_\lambda$. To this end, for each $\omega \in \Omega$, consider the operator

$$
\widetilde{A} = \widetilde{A}_\lambda(t,\omega) : H_0^1(\mathcal{O}) \to H^{-1}(\mathcal{O})
$$

defined by

$$\begin{aligned}\left\langle \widetilde{A}_\lambda(t)y,\varphi\right\rangle &= \int_{\mathcal{O}} \psi_\lambda(\nabla e^{W(t)}y))\cdot\nabla(e^{-W(t)}\varphi)d\xi \\ &+\lambda\int_{\mathcal{O}} \nabla(e^{W(t)}y)\cdot\nabla(e^{-W(t)}\varphi)d\xi+\frac{1}{2}\int_{\mathcal{O}}\mu y\varphi\,d\xi, \\ &\qquad \forall\varphi\in H_0^1(\mathcal{O}).\end{aligned} \tag{5.13}$$

In terms of $\widetilde{A}_\lambda$, equation (5.2) becomes

$$\begin{aligned}&\frac{dY_\lambda}{dt}+\widetilde{A}_\lambda(t)Y_\lambda(t)=0,\ \text{a.e. } t\in(0,T), \\ &X_\lambda(0)=x.\end{aligned} \tag{5.14}$$

It is easily seen that, for every $t\in[0,T]$ and $\mathbb{P}$-a.s., $\omega\in\Omega$, $\widetilde{A}_\lambda(t)=\widetilde{A}_\lambda(t)(\omega)$ is demicontinuous (that is, strongly-weakly continuous), coercive, that is,

$$\left\langle \widetilde{A}_\lambda(t)y,y\right\rangle \ge \lambda\|y\|_1^2-\alpha_t^\lambda|y|_2^2,\ \forall y\in H_0^1(\mathcal{O}), \tag{5.15}$$

bounded, that is,

$$\|\widetilde{A}_\lambda(t)y\|_{-1}\le C_t(1+\|y\|_1),\ \forall y\in H_0^1(\mathcal{O}), \tag{5.16}$$

and δ-monotone, that is,

$$\left\langle \widetilde{A}_\lambda(t)y-\widetilde{A}_\lambda(t)z,y-z\right\rangle+\delta_t^\lambda|y-z|_2^2\ge 0,\ \forall y,z\in H_0^1(\mathcal{O}), \tag{5.17}$$

where $C_t,\alpha_t^\lambda,\delta_t^\lambda:\Omega\to\mathbb{R}_+$, $t\in[0,T]$, are $(\mathcal{F}_t)$-adapted processes, $\mathbb{P}$-a.s. continuous on $[0,T]$. (Since, as pointed out before, we only need the uniqueness part, i.e., we only need (5.17), for the reader's convenience we include its proof in Appendix 2, i.e., Section 10.)

Hence, for each $x\in L^2(\mathcal{O})$, there is a unique solution Y_λ to (5.14) satisfying (5.7). (See, e.g., [6], p. 177). This completes the proof of (ii). To prove (iii), we need the following two lemmas:

Lemma 5.3 *Let* $x\in H_0^1(\mathcal{O})$. *Then,* $X_\lambda\in L^2(\Omega;L^\infty([0,T];H_0^1(\mathcal{O})))\cap L^2([0,T]\times\Omega;H^2(\mathcal{O}))$ *and*

$$\mathbb{E}\left[\sup_{t\in[0,T]}\|X_\lambda(t)\|_1^2\right]+\lambda\mathbb{E}\int_0^T|\Delta X_\lambda(t)|_2^2dt\le C\|x\|_1^2,\ \lambda\in(0,1]. \tag{5.18}$$

Proof of Lemma 5.3. In this proof, constants may change from line to line, though we continue to denote them by C. We set $A=-\Delta$, $D(A)=H_0^1(\mathcal{O})\cap H^2(\mathcal{O})$, $J_\varepsilon=(1+\varepsilon A)^{-1}$, $A_\varepsilon=AJ_\varepsilon=\frac{1}{\varepsilon}(I-J_\varepsilon)$ and note that, by virtue of Corollary 8.7 in Appendix 1, we have

$$\begin{aligned}-\langle A_\varepsilon X_\lambda,\operatorname{div}\psi_\lambda(\nabla X_\lambda)\rangle&=\frac{1}{\varepsilon}\int_{\mathcal{O}}(\nabla y-\nabla J_\varepsilon(y))\cdot\psi_\lambda(\nabla y)d\xi\\&\geq\frac{1}{\varepsilon}\int_{\mathcal{O}}(j_\lambda(\nabla y)-j_\lambda(\nabla J_\varepsilon(y)))d\xi\geq 0.\end{aligned}\tag{5.19}$$

Now, we apply Itô's formula to the function $\varphi(x)=\frac{1}{2}|A_\varepsilon^{\frac{1}{2}}x|_2^2$. We have $D\varphi=A_\varepsilon$, and so we get by Hypotheses (H1) and (H2) that

$$\begin{aligned}&\frac{1}{2}|A_\varepsilon^{\frac{1}{2}}X_\lambda(t)|_2^2+\lambda\int_0^t\langle A_\varepsilon X_\lambda(s),AX_\lambda(s)\rangle\,ds\\&\quad-\int_0^t\langle A_\varepsilon X_\lambda(s),\operatorname{div}\psi_\lambda(\nabla X_\lambda(s))\rangle\,ds\\&\quad\leq\frac{1}{2}|x|_2^2+C\int_0^t\|X_\lambda(s)\|_1^2ds\\&\quad+\int_0^t\langle A_\varepsilon X_\lambda(s),X_\lambda(s)dW(s)\rangle,\ t\in[0,T],\end{aligned}\tag{5.20}$$

since $|A_\varepsilon^{\frac{1}{2}}x|_2\leq\|x\|_1,\ \forall x\in H_0^1(\mathcal{O}),\ \varepsilon\in(0,1]$.

Now, keeping in mind that, for all $\varepsilon>0$,

$$\langle A_\varepsilon y,Ay\rangle\geq|A_\varepsilon y|_2^2,\ \forall y\in H_0^1(\mathcal{O}),$$

and, taking into account (5.19), we obtain by (5.20) that, for some $C>0$ independent of λ and ε,

$$\begin{aligned}|A_\varepsilon^{\frac{1}{2}}X_\lambda(t)|_2^2+\lambda\int_0^t|A_\varepsilon X_\lambda(s)|_2^2ds&\leq\frac{1}{2}|x|_2^2+C\int_0^t\|X_\lambda(s)\|_1^2ds\\&+\int_0^t\langle A_\varepsilon X_\lambda(s),X_\lambda(s)dW(s)\rangle,\ t\in[0,T],\ \forall\lambda,\varepsilon>0.\end{aligned}\tag{5.21}$$

In particular, for all $t\in[0,T]$

$$\begin{aligned}\sup_{r\in[0,t]}|A_\varepsilon^{\frac{1}{2}}X_\lambda(r)|_2^2&\leq\|x\|_1^2+C\int_0^t\sup_{r\in[0,s]}\|X_\lambda(r)\|_1^2ds\\&+\sup_{r\in[0,t]}\left|\int_0^r\langle A_\varepsilon X_\lambda(s),X_\lambda(s)dW(s)\rangle\right|.\end{aligned}$$

Hence, by the Burkholder-Davis-Gundy (for $p = 1$) and Gronwall's inequalities, we obtain that, for some $C > 0$, independent of λ and ε,

$$\mathbb{E}\left[\sup_{s\in[0,T]} |A_\varepsilon^{\frac{1}{2}} X_\lambda(s)|_2^2\right] \le 2\|x\|_1^2 e^{CT},\ \forall \lambda, \varepsilon \in (0,1].$$

Letting $\varepsilon \to 0$, we obtain

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|X_\lambda(t)\|_1^2\right] \le C\|x\|_1^2,\ \forall \lambda \in (0,1]. \tag{5.22}$$

Hence, taking expectation in (5.21) and letting $\varepsilon \to 0$, we obtain

$$\lambda \mathbb{E}\int_0^T |\Delta X_\lambda(s)|_2^2 ds \le C\|x\|_1^2,\ \forall \lambda \in (0,1].$$

This completes the proof of Lemma 5.3. □

Lemma 5.4 *Let $x \in H_0^1(\mathcal{O})$. Then, $Y_\lambda \in C([0,T]; H_0^1(\mathcal{O})) \cap L^2([0,T]; H^2(\mathcal{O}))$, $\mathbb{P}$-a.s.*

Proof. We rewrite (5.2) as the linear parabolic random equation

$$\begin{aligned}&\frac{\partial Y_\lambda}{\partial t} = \lambda \Delta Y_\lambda + f(t,\xi) \text{ in } (0,T)\times\mathcal{O},\\ &Y_\lambda = 0 \text{ on } (0,T)\times\partial\mathcal{O},\\ &Y_\lambda(0,\xi) = x(\xi) \text{ in } \mathcal{O},\end{aligned} \tag{5.23}$$

where

$$\begin{aligned}f(t,\xi) = {}& e^{-W(t,\xi)} \operatorname{div} \psi_\lambda(\nabla(e^{W(t,\xi)} Y_\lambda(t,\xi))) - \frac{1}{2}\,\mu(\xi) Y_\lambda(t,\xi)\\ &+2\lambda \nabla W(t,\xi)\cdot \nabla Y_\lambda(t,\xi) + \Delta W(t,\xi) Y_\lambda(t,\xi) + Y_\lambda |\nabla W|^2.\end{aligned}$$

Since, for $y \in H_0^1(\mathcal{O}) \cap H^2(\mathcal{O})$,

$$\operatorname{div} \psi_\lambda(\nabla y) = \begin{cases} \dfrac{1}{\lambda}\Delta y & \text{on } \{|\nabla y| \le \lambda\},\\[2mm] \dfrac{\Delta y}{|\nabla y|} - \dfrac{\nabla y \cdot \nabla |\nabla y|}{|\nabla y|^2} & \text{on } \{|\nabla y| > \lambda\},\end{cases} \tag{5.24}$$

by Lemma 5.3, we know that

$$f(t) \in L^2(0,T;L^2(\mathcal{O})),\ \mathbb{P}\text{-a.s.} \tag{5.25}$$

Then, by the general theory of linear parabolic equations (see, e.g. [14]), we have, for each $\omega \in \Omega$,

$$Y_\lambda \in C([0,T];H_0^1(\mathcal{O})) \cap L^2([0,T];H^2(\mathcal{O})), \tag{5.26}$$

and the lemma is proved. □

Lemmas 5.3, 5.4 and part (ii) now imply part (iii) and the proof of Proposition 5.1 is complete. □

Proof of Theorem 3.2 (continued). It is enough to prove the existence for initial conditions $x \in H_0^1(\mathcal{O})$, provided one can also prove (3.6) for such solutions with initial conditions $x, x^* \in H_0^1(\mathcal{O})$. Indeed, if we have that we can extend our solutions for arbitrary $x \in L^2(\mathcal{O})$, since $H_0^1(\mathcal{O})$ is dense in $L^2(\mathcal{O})$ and (3.2) is obviously stable under taking limits in X_n replacing X, with X_n converging in $L^2(\Omega; C([0,T];L^2(\mathcal{O})))$ (since ϕ is lower semicontinuous on $L^2(\mathcal{O})$).

Hence, let $x \in H_0^1(\mathcal{O})$. Using the Itô formula in (5.1) (or, equivalently, in (5.4)), we obtain that

$$\begin{aligned}\mathbb{E}|X_\lambda(t)|_2^2 &+2\mathbb{E}\int_0^t \int_{\mathcal{O}} j_\lambda(\nabla(X_\lambda(s,\xi)))d\xi\, ds + \lambda\mathbb{E}\int_0^t |\nabla(X_\lambda(s))|_2^2 ds \\ &\le |x|_2^2 + \frac{1}{2}\ \mathbb{E}\int_0^t \int_{\mathcal{O}} \sum_{j=1}^{\infty} \mu_j^2 |X_\lambda e_j|^2 d\xi\, ds,\ t \in [0,T],\end{aligned}$$

because $\widetilde{\psi}_\lambda(u)\cdot u \ge j_\lambda(u) + \lambda|u|^2, \forall u \in \mathbb{R}^N$.

This yields (via Gronwall's lemma)

$$\begin{aligned}&\mathbb{E}|X_\lambda(t)|_2^2 + 2\mathbb{E}\int_0^t \int_{\mathcal{O}} j_\lambda(\nabla X_\lambda(s,\xi))d\xi\, ds \\ &\quad +\lambda\mathbb{E}\int_0^t \int_{\mathcal{O}} |\nabla X_\lambda|^2 d\xi\, ds \le C_1,\quad \forall \lambda > 0,\ t \in [0,T],\end{aligned} \tag{5.27}$$

where $C_1 = e^{2C_\infty^2}|x|_2^2$.

Moreover, we have, for all $t \in [0,T]$,

$$\mathbb{E}\int_0^t \phi(X(t))dt \le \liminf_{\lambda\to 0}\mathbb{E}\int_0^t\int_{\mathcal{O}} j_\lambda(\nabla(X_\lambda(t)))d\xi\,dt < \infty, \tag{5.28}$$

where ϕ is defined by (2.3). Indeed, we have

$$|j_\lambda(\nabla u) - |\nabla u|\,| \le \frac{1}{2}\,\lambda, \tag{5.29}$$

and this yields

$$\left|\mathbb{E}\left(\int_0^t\int_{\mathcal{O}} j_\lambda(\nabla X_\lambda(t))d\xi\,dt - \int_0^t \phi(X_\lambda(t))dt\right)\right| \le C\lambda,\ \forall \lambda\in(0,1]. \tag{5.30}$$

On the other hand, we have

$$\lim_{\lambda\to 0}\mathbb{E}\left[\sup_{t\in[0,T]}|X_\lambda(t)-X(t)|_2^2\right] = 0. \tag{5.31}$$

Indeed, by Itô's formula, we have

$$\begin{aligned}
&\frac{1}{2}\,d|X_\lambda(t)-X_\varepsilon(t)|_2^2\\
&\qquad+\langle\psi_\lambda(\nabla X_\lambda(t))-\psi_\varepsilon(\nabla X_\varepsilon(t)),\nabla(X_\lambda(t)-X_\epsilon(t))\rangle\\
&\qquad+\langle\lambda\nabla X_\lambda(t)-\varepsilon\nabla X_\epsilon(t),\nabla(X_\lambda(t)-X_\varepsilon(t))\rangle\\
&\qquad=\frac{1}{2}\int_0^t\int_{\mathcal{O}}\sum_{j=1}^\infty \mu_j^2|(X_\lambda(s)-X_\varepsilon(s))e_j|^2 ds\,d\xi\\
&\qquad+\int_0^t\langle X_\lambda - X_\varepsilon,(X_\lambda-X_\varepsilon)dW(s)\rangle,\ t\in[0,T].
\end{aligned}$$

Taking into account that, by the definition of ψ_λ,

$$(\psi_\lambda(u)-\psi_\epsilon(v))\cdot(u-v)\ge(\lambda\psi_\lambda(u)-\epsilon\psi_\varepsilon(v))\cdot(\psi_\lambda(u)-\psi_\varepsilon(v))\ge-(\lambda+\varepsilon)$$

and that

$$\begin{aligned}
&\langle\lambda\nabla X_\lambda(t)-\varepsilon\nabla X_\varepsilon(t),\nabla(X_\lambda(t)-X_\varepsilon(t))\rangle\\
&\qquad=-\langle\lambda\Delta X_\lambda(t)-\varepsilon\Delta X_\varepsilon(t),X_\lambda(t)-X_\varepsilon(t)\rangle\\
&\qquad\ge-(\lambda^2|\Delta X_\lambda(t)|_2^2+\varepsilon^2|\Delta X_\varepsilon(t)|_2^2)-\frac{1}{2}\,|X_\lambda-X_\varepsilon|_2^2,
\end{aligned}$$

we get, for some constant $C>0$ and all $t\in[0,T]$,

$$
\begin{aligned}
|X_\lambda(t)-X_\varepsilon(t)|_2^2 &\le (C_\infty^2+1)\int_0^t |X_\lambda(s)-X(s)|_2^2 ds + M_{\lambda,\varepsilon}(t)\\
&+2(\lambda+\varepsilon)t\int_{\mathcal{O}} d\xi + 2\lambda^2\int_0^t |\Delta X_\lambda(s)|_2^2 ds + 2\varepsilon^2\int_0^t |\Delta X_\varepsilon(s)|_2^2 ds,
\end{aligned}
$$

where

$$M_{\lambda,\varepsilon}(t)=2\int_0^t \langle X_\lambda - X_\varepsilon, (X_\lambda - X_\varepsilon)dW(s)\rangle,\ t\in[0,T],$$

is a local real-valued $(\mathcal{F}_t)$-martingale. Then, by the Burkholder-Davis-Gundy inequality (for $p=1$), we get (see [9], (3.12)-(3.13)), for some constant $C>0$,

$$
\begin{aligned}
\mathbb{E}\sup_{0\le s\le t}|X_\lambda(s)-X_\varepsilon(s)|_2^2 &\le C(\lambda+\varepsilon)+C\int_0^t \mathbb{E}\sup_{0\le s\le t}|X_\lambda(s)-X_\varepsilon(s)|_2^2 ds\\
&+C\lambda^2\mathbb{E}\int_0^t |\Delta X_\lambda(s)|_2^2 ds + C\varepsilon^2\mathbb{E}\int_0^t |\Delta X_\varepsilon(s)|_2^2 ds,\ t\in[0,T],
\end{aligned}
$$

and, by Lemma 5.3 and Gronwall's lemma, it follows that $\{X_\lambda\}_\lambda$ is Cauchy in $L^2(\Omega;C([0,T];L^2(\mathcal{O})))$, which completes the proof of (5.31).

Now, recalling that ϕ is lower-semicontinuous in $L^1(\mathcal{O})$ (see (2.2)), we have by (5.31) and Fatou's lemma that

$$\liminf_{\lambda\to 0}\mathbb{E}\int_0^t \phi(X_\lambda(t))dt \ge \mathbb{E}\int_0^t \phi(X(t))dt,\ \forall t\in[0,T],$$

which, by virtue of (5.30), implies (5.28), as claimed.

We note that (5.31) and (5.6) imply (3.6), and that (3.4) then follows from (5.5) and Fatou's lemma.

It remains to prove (3.2). By Itô's formula, we have, for all the processes Z satisfying Definition 3.1(iii) and (3.1), (cf. Remark 5.2),

$$
\begin{aligned}
&\frac{1}{2}\,\mathbb{E}|(X_\lambda(t)-Z(t))|_2^2+\mathbb{E}\int_0^t\int_{\mathcal{O}} j_\lambda(\nabla X_\lambda(\tau))d\xi\,d\tau\\
&\quad\le \frac{1}{2}\,\mathbb{E}|x-Z(0)|_2^2+\mathbb{E}\int_0^t\int_{\mathcal{O}} j_\lambda(\nabla Z(\tau))d\xi\,d\tau\\
&\quad+\frac{1}{2}\,\mathbb{E}\int_0^t \langle X_\lambda(\tau)-Z(\tau),G(\tau)\rangle\,d\tau\\
&\quad+\frac{1}{2}\,\mathbb{E}\int_0^t\int_{\mathcal{O}} \mu(X_\lambda(\tau)-Z(\tau))^2 d\xi\,d\tau,\ t\in[0,T].
\end{aligned}
\tag{5.32}
$$

Now, letting λ tend to zero, it follows by (5.28), (5.29) and (5.31) that (3.2) holds. This completes the proof of the existence. $\square$

Uniqueness. Let X^* be an arbitrary variational solution to (3.1) with $X^*(0) = x^* \in L^2(\mathcal{O})$ and satisfying (3.5).

Let $x \in H_0^1(\mathcal{O})$ and X be the solution constructed in the existence part of the proof, but with $X(0) = x$. Set $Y^* := e^{-W}X^*$ and $Y := e^{-W}X$. We set $Y_\lambda^\varepsilon = J_\varepsilon(Y_\lambda)$, where Y_λ is the solution to (5.2), but with initial condition $x \in H_0^1(\mathcal{O})$. On the basis of (H2) and Lemma 5.3 it follows that $e^{-W}Y_\lambda^\varepsilon \in L^2([0,T] \times \Omega; H_0^1(\mathcal{O}))$. Clearly, it is also a $\mathbb{P}$-a.s. continuous $(\mathcal{F}_t)$-adapted process in $L^2(\mathcal{O})$. Hence, in (4.4), (4.3), we may choose $\widetilde{Z} = Y_\lambda^\varepsilon$ and we obtain that for

$$G = G_\lambda^\varepsilon = -J_\varepsilon(\text{div } \widetilde{\psi}_\lambda(\nabla(e^W Y_\lambda))) + \eta_\lambda^\varepsilon,$$

where

$$\begin{aligned}\eta_\lambda^\varepsilon = \frac{1}{2}\, & e^W(J_\varepsilon(\mu Y_\lambda) - \mu J_\varepsilon(Y_\lambda)) \\ & + J_\varepsilon(\text{div}\, \widetilde{\psi}_\lambda(\nabla(e^W Y_\lambda))) - e^W J_\varepsilon(e^{-W}\text{div}\, \widetilde{\psi}_\lambda(\nabla(e^W Y_\lambda))),\end{aligned}$$

the function $\widetilde{Z}$ satisfies (4.4).

Then, by (4.3), we have

$$\begin{aligned}&\frac{1}{2}\,\mathbb{E}|e^{W(t)}(Y^*(t) - Y_\lambda^\varepsilon(t))|_2^2 + \mathbb{E}\int_0^t \phi(e^{W(\tau)}Y^*(\tau))d\tau \\ &\quad \le \frac{1}{2}\,|x^* - x|_2^2 + \mathbb{E}\int_0^t \phi(e^{W(\tau)}Y_\lambda^\varepsilon(\tau))d\tau \\ &\quad + \frac{1}{2}\,\mathbb{E}\int_0^t \int_{\mathcal{O}} \mu e^{2W(\tau)}(Y^*(\tau) - Y_\lambda^\varepsilon(\tau))^2 d\xi\, d\tau \\ &\quad + \mathbb{E}\int_0^t \left\langle e^{W(\tau)}(Y^*(\tau) - Y_\lambda^\varepsilon(\tau)), G_\lambda^\varepsilon \right\rangle d\tau, \\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{a.e. } t \in [0,T],\ \lambda > 0.\end{aligned} \tag{5.33}$$

By Green's formula, we have

$$\begin{aligned}&\left\langle e^W(Y^* - Y_\lambda^\varepsilon), G_\lambda^\varepsilon \right\rangle \\ &= \left\langle \psi_\lambda(\nabla(e^W Y_\lambda)) + \lambda\nabla(e^W Y_\lambda), \nabla J_\varepsilon(e^W Y^*) - \nabla(e^W Y_\lambda) \right\rangle \\ &+ \left\langle \psi_\lambda(\nabla(e^W Y_\lambda)) + \lambda\nabla(e^W Y_\lambda), \zeta_\lambda^\varepsilon \right\rangle + \left\langle e^W(Y^* - Y_\lambda^\varepsilon), \eta_\lambda^\varepsilon \right\rangle\end{aligned}$$

where

$$\zeta_\lambda^\varepsilon = \nabla(e^W Y_\lambda) - \nabla J_\varepsilon(e^W Y_\lambda^\varepsilon).$$

Taking into account that

$$\psi_\lambda(u)\cdot(u-v) \geq j_\lambda(u) - j_\lambda(v),\ \forall u,v\in\mathbb{R}^d,$$

this yields

$$\begin{aligned}
&\left\langle e^W(Y^*-Y_\lambda^\varepsilon), G_\lambda^\varepsilon\right\rangle \leq \phi_\lambda(J_\varepsilon(e^W Y^*)) - \phi_\lambda(e^W Y_\lambda) - \lambda|\nabla(e^W Y_\lambda)|_2^2\\
&\qquad -\lambda\left\langle \Delta(e^W Y_\lambda), J_\varepsilon(e^W Y^*)\right\rangle + \left\langle e^W(Y^*-Y_\lambda^\varepsilon),\eta_\lambda^\varepsilon\right\rangle\\
&\qquad +\left\langle \psi_\lambda(\nabla(e^W Y_\lambda)) + \lambda\nabla(e^W Y_\lambda), \zeta_\lambda^\varepsilon\right\rangle.
\end{aligned}$$

Here, ϕ_λ is the function

$$\phi_\lambda(z) = \int_{\mathcal{O}} j_\lambda(\nabla z)d\xi,\quad \forall z\in H_0^1(\mathcal{O}).$$

Substituting into (5.33), we obtain that

$$\begin{aligned}
&\frac{1}{2}\,\mathbb{E}|e^{W(t)}(Y^*(t)-Y_\lambda^\varepsilon(t))|_2^2 + \mathbb{E}\int_0^t \phi(e^{W(\tau)}Y^*(\tau))d\tau\\
&\quad +\mathbb{E}\int_0^t \phi_\lambda(e^{W(\tau)}Y_\lambda(\tau))d\tau + \lambda\mathbb{E}\int_0^t |\nabla(e^{W(\tau)}Y_\lambda(\tau))|_2^2 d\tau\\
&\quad \leq \frac{1}{2}\,|x^*-x|_2^2 + \mathbb{E}\int_0^t \phi(e^{W(\tau)}Y_\lambda^\varepsilon(\tau))d\tau\\
&\quad +\mathbb{E}\int_0^t \phi_\lambda(J_\varepsilon(e^W Y^*(\tau)))d\tau\\
&\quad -\lambda\mathbb{E}\int_0^t \left\langle \Delta(e^{W(\tau)}Y_\lambda(\tau)), J_\varepsilon(e^{W(\tau)}Y^*(\tau))\right\rangle d\tau\\
&\quad +\mathbb{E}\int_0^t \left(\left\langle e^W(Y^*-Y_\lambda^\varepsilon),\eta_\lambda^\varepsilon\right\rangle + \left\langle\psi_\lambda(\nabla(e^W Y_\lambda)) + \lambda\nabla(e^W Y_\lambda),\zeta_\lambda^\varepsilon\right\rangle\right) d\tau\\
&\quad +\frac{1}{2}\,C_\infty^2\mathbb{E}\int_0^t (e^{W(\tau)}(Y^*(\tau)-Y_\lambda^\varepsilon(\tau)))|_2^2 d\tau,\ t\in[0,T],\ \forall\lambda>0,
\end{aligned}\tag{5.34}$$

where C_∞^2 is as in (H1). Now, as seen earlier in (5.30), we have

$$|\phi(e^{W(\tau)}Y_\lambda(\tau)) - \phi_\lambda(e^{W(\tau)}Y_\lambda(\tau))| \leq C\lambda,\quad \forall\tau\in[0,T].\tag{5.35}$$

Similarly, we have also

$$\int_0^T |\phi_\lambda(J_\varepsilon(e^{W(\tau)}Y^*(\tau))) - \phi(J_\varepsilon(e^{W(\tau)}Y^*(\tau))|d\tau \le C\lambda. \tag{5.36}$$

Substituting (5.35), (5.36) in (5.34), yields

$$\begin{aligned}
&\frac{1}{2}\,\mathbb{E}|e^{W(t)}(Y^*(t)-Y_\lambda^\varepsilon(t))|_2^2 + \mathbb{E}\int_0^t \phi(e^{W(\tau)}Y^*(\tau))d\tau \\
&\le \frac{1}{2}\,|x^*-x|_2^2 + \mathbb{E}\int_0^t \phi(J_\varepsilon(e^{W(\tau)}Y^*(\tau)))d\tau \\
&+\mathbb{E}\int_0^t (\phi(e^{W(\tau)}Y_\lambda^\varepsilon(\tau)) - \phi(e^{W(\tau)}Y_\lambda(\tau)))d\tau \\
&+\frac{1}{2}\,C_\infty^2\mathbb{E}\int_0^t |e^{W(\tau)}(Y^*(\tau)-Y_\lambda^\varepsilon(\tau))|_2^2 d\tau \\
&-\lambda\mathbb{E}\int_0^t \left\langle \Delta(e^{W(\tau)}Y_\lambda(\tau)), J_\varepsilon(e^{W(\tau)}Y^*(\tau))\right\rangle d\tau \\
&+C_{\lambda,\varepsilon}\left(\mathbb{E}\int_0^t |\zeta_\lambda^\varepsilon(\tau)|_2^2 d\tau\right)^{1/2} + C_{\lambda,\varepsilon}\left(\mathbb{E}\left(\int_0^t |\eta_\lambda^\varepsilon(\tau)|_2^2 d\tau\right)^{r/2}\right)^{1/r},
\end{aligned} \tag{5.37}$$

where δ is as in (3.5), $r=\frac{\delta+2}{\delta+1}$ and

$$\begin{aligned}
C_{\lambda,\varepsilon} = 4&\left(\mathbb{E}\left(\int_0^T |e^W(Y^*-Y_\lambda^\varepsilon)|_2^2 d\tau\right)^{\frac{2+\delta}{2}}\right)^{\frac{1}{2+\delta}} \\
&+4+4\left(\mathbb{E}\int_0^T \lambda|\nabla(e^W Y_\lambda)|_2^2 d\tau\right)^{1/2}.
\end{aligned}$$

Now, recalling that, by Corollary 8.5,

$$\mathbb{E}\int_0^t \phi(J_\varepsilon(e^{W(\tau)}Y^*(\tau)))d\tau \le \mathbb{E}\int_0^t \phi(e^{W(\tau)}Y^*(\tau))d\tau,\ \forall \varepsilon>0,$$

letting $\varepsilon\to 0$ in (5.37) yields

$$\begin{aligned}
\mathbb{E}|e^{W(t)}&(Y^*(t)-Y_\lambda(t))|_2^2 \le |x^*-x|_2^2 \\
&+C_\infty^2\mathbb{E}\int_0^t |e^{W(\tau)}(Y^*(\tau)-Y_\lambda(\tau))|^2 d\tau \\
&-\lambda\mathbb{E}\int_0^t \left\langle \Delta(e^{W(\tau)}Y_\lambda(\tau)), e^{W(\tau)}Y^*(\tau)\right\rangle d\tau.
\end{aligned} \tag{5.38}$$

because

$$\lim_{\varepsilon\to 0}\mathbb{E}\left(\int_0^T |\eta_\lambda^\varepsilon(\tau)|_2^2 d\tau\right)^{r/2}=0,$$
$$\lim_{\varepsilon\to 0}\mathbb{E}\int_0^T |\zeta_\lambda^\varepsilon(\tau)|_2^2 d\tau=0,$$
$$\sup\{C_{\lambda,\varepsilon};\ \varepsilon\in(0,1)\}<\infty,$$
$$e^W Y_\lambda = X_\lambda \in L^2([0,T]\times\Omega;H^2(\mathcal{O})) \text{ by Lemma 5.3 and}$$
$$e^W Y^* \in L^2([0,T]\times\Omega;L^2(\mathcal{O})).$$

To check all this is pretty routine. The main problem is to justify the interchange of "$\lim_{\varepsilon\to 0}$" with the integral with respect to $d\tau\otimes\mathbb{P}$, i.e., to find an integrable uniformly dominating function. As an exemplary case, we show how this is done for the last summand in the definition of η_λ^ε:

Clearly, since J_ε is a contraction on $L^2(\mathcal{O})$, it follows by (5.24) that there exists a constant $c>0$ such that, for all $\varepsilon\in(0,1)$,

$$|e^W J_\varepsilon(e^{-W}\mathrm{div}\,\widetilde{\psi}_\lambda(\nabla(e^W Y_\lambda)))|_2^2 \le c\cdot\exp\left(4\sup_{\tau\in[0,T]}|W(\tau)|_\infty\right)\|X_\lambda\|_{H^2(\mathcal{O})}^2.$$

Hence, applying Hölder's inequality with $p=\frac{2}{r}$ (>1), $q=\frac{2}{2-r}$ to the expectation, we obtain

$$\begin{aligned}&\mathbb{E}\left(\int_0^T \exp\left(4\sup_{\tau\in[0,T]}|W(\tau)|_\infty\right)\|X_\lambda\|_{H^2(\mathcal{O})}^2 d\tau\right)^{r/2}\\ &\quad\le\left(\mathbb{E}\exp\left(\frac{8r}{2-r}\sup_{\tau\in[0,T]}|W(\tau)|_\infty\right)\right)^{\frac{2-r}{2}}\left(\mathbb{E}\int_0^T\|X_\lambda(\tau)\|_{H^2(\mathcal{O})}^2 d\tau\right)^{r/2},\end{aligned}$$

which is finite by (2.9) and Lemma 5.3.

Now, by Lemma 5.3, we have

$$\lim_{\lambda\to 0}\lambda\mathbb{E}\int_0^t\left\langle\Delta(e^{W(\tau)}Y_\lambda(\tau)),e^{W(\tau)}Y^*(\tau)\right\rangle d\tau=0.$$

Then, letting $\lambda\to 0$ in (5.38), we obtain via Gronwall's lemma

$$\mathbb{E}|X^*(t)-X(t)|_2^2=\mathbb{E}|e^{W(t)}(Y^*(t)-Y(t))|_2^2\le|x^*-x|_2^2 e^{C_\infty^2 T}.$$

Now, letting $x \to x^*$ in $L^2(\mathcal{O})$, we see by (3.6) that X^* coincides with the solution starting at x^* constructed in the existence part of the proof, which is hence unique. $\square$

Remark 5.5 We did not succeed in proving the uniqueness for Theorem 3.2 directly for the original equation (1.1). The reason is that, regularizing (1.1) by J_ε destroys the special form of the noise. Therefore, we had to use equation (4.1) and Proposition 4.2.

6 Positivity of solutions

It should be emphasized that physical models of nonlinear diffusion are concerned in general with nonnegative solutions of equation (1.1). In this context, we have the following result.

Theorem 6.1 *In Theorem* 3.2 *assume further that* $x \geq 0$*, a.e. in* $\mathcal{O}$*. Then*

$$X(t,\xi) \geq 0 \quad a.e. \ in \ (0,T) \times \mathcal{O} \times \Omega. \tag{6.1}$$

Proof. It suffices to show that the solution X_λ to (5.1) is a.e. nonnegative on $[0,T] \times \mathcal{O} \times \Omega$. By (5.6) we may assume that $x \in L^4(\mathcal{O})$. Below we only give a heuristic argument to prove the assertion (e.g., apply Itô's formula in an informal way), which can be made rigorous by regularization. Since the latter is analogous as in the proof of Theorem 2.2 in [7] or can be done similarly as in the proof of Theorem 7.1 below, we omit the details.

We apply the Itô formula in (5.1) to the function $x \to \frac{1}{4}\,|x^-|_4^4$. We obtain

$$\begin{aligned}\frac{1}{4}\,\mathbb{E}\int_{\mathcal{O}} |X_\lambda^-(t,\xi)|^4 d\xi + \mathbb{E}\int_0^t \int_{\mathcal{O}} \widetilde{\psi}_\lambda(\nabla X_\lambda(s,\xi)) \cdot \nabla |X_\lambda^-(s,\xi)|^3 d\xi\, ds \\ = \frac{1}{4}\int_{\mathcal{O}} |x^-(\xi)|^4 d\xi + \mathbb{E}\int_0^t \int_{\mathcal{O}} \sum_{j=1}^\infty \mu_j^2 (X_\lambda e_j)^2 (X_\lambda^-)^2 d\xi\, ds.\end{aligned}$$

Recalling that $\nabla y \cdot \nabla y^- = -|\nabla y^-|^2$ a.e. in $\mathcal{O}$ for each $y \in H^1(\mathcal{O})$, it follows that

$$\mathbb{E}\int_{\mathcal{O}} |X_\lambda^-(t,\xi)|^4 d\xi \leq C\mathbb{E}\int_0^t \int_{\mathcal{O}} |X_\lambda^-(t,\xi)|^4 d\xi\, ds, \forall t \in [0,T],$$

which implies that $X_\lambda^- \equiv 0$, as claimed.

7 Extinction in finite-time

A striking feature of highly singular nonlinear diffusion equations is the extinction in finite time of the solution. In nonlinear diffusion phenomena, this is due to the singularity at level $X = 0$ of the diffusivity and this causes a fast loss of mass. (See [10] for the case of stochastic porous media equation and [9], [13] for stochastic self-organized criticality.) A similar phenomenon happens in the case of equation (1.1).

Theorem 7.1 *Let $2 \le N \le 3$. Let X be as in Theorem 3.2, with initial condition $x \in L^N(\mathcal{O})$, and let $\tau = \inf\{t \ge 0; |X(t)|_N = 0\}$. Then, we have*

$$\mathbb{P}[\tau \le t] \ge 1 - \rho^{-1} \left(\int_0^t e^{-C^* s} ds \right)^{-1} |x|_N, \ \forall t \ge 0. \tag{7.1}$$

Here $\rho = \inf\{|y|_{W_0^{1,1}(\mathcal{O})} / |y|_{\frac{N}{N-1}};\ y \in W_0^{1,1}(\mathcal{O})\}$ and $C^* = \frac{C_\infty^2}{2}(N-1)$. In particular, if $|x|_N < \rho/C^*$, then $\mathbb{P}[\tau < \infty] > 0$.

We shall prove Theorem 7.1 as stated, i.e., only for $2 \le N \le 3$. The case $N = 1$ is similar, but one proves extinction in $L^2(\mathcal{O})$-norm rather than $L^1(\mathcal{O})$-norm (see [11, Theorem 3] for details). We fix $\lambda \in (0, 1]$ and start with the following lemma, which is one of the main ingredients of the proof.

Before, we recall that, by (5.8), X_λ is $\mathbb{P}$-a.s. continuous in $H_0^1(\mathcal{O})$. For $K \in \mathbb{N}$, $K > \|x\|_1$, define the $\{\mathcal{F}_t\}$-stopping time

$$\tau_K := \inf\{t \ge 0; \|X_\lambda(t)\|_1 > K\}.$$

Lemma 7.2 *Let $x \in H_0^1(\mathcal{O})$. Then:*

(i) $e^{-NC^* t}|X_\lambda(t)|_N^N,\ t \ge 0$, *is an $\{\mathcal{F}_t\}$-supermartingale, and hence so is* $e^{-C^* t}|X_\lambda(t)|_N$, $t \ge 0$.

(ii) *We have $\mathbb{P}$-a.s.*

$$\begin{aligned} &|X_\lambda(t)|_N^N + N\rho \int_s^t |X_\lambda(r)|_N^{N-1} dr \\ &\le |X_\lambda(s)|_N^N + NC^* \int_s^t |X_\lambda(r)|_N^N dr + N(N-1)\lambda \int_s^t |X_\lambda(r)|_{N-2}^{N-2} dr \\ &+ N \int_s^t \left\langle |X_\lambda(r)|^{N-2} X_\lambda(r), X_\lambda(r) dW(r) \right\rangle,\ \forall s, t \in [0, T],\ s \le t. \end{aligned} \tag{7.2}$$

Proof. Since $N \leq 3$, we have by Sobolev embedding, $H_0^1(\mathcal{O}) \subset L^4(\mathcal{O})$ continuously, hence, for some constant $C > 0$,

$$\sup_{t \in [0,\tau_K]} |X_\lambda(t)|_N \leq CK \text{ on } \Omega. \tag{7.3}$$

We have by standard interpolation (see, e.g., [32, Theorem 2.1]) if $N = 3$

$$\begin{aligned}\mathbb{E}\int_0^{\tau_K} \|X_\lambda(t)\|_{1,3}^3 dt &\leq C\mathbb{E}\int_0^{\tau_K} \|X_\lambda(t)\|_{H^2(\mathcal{O})}^2 |X_\lambda(t)|_3 dt \\ &\leq CK\mathbb{E}\int_0^T \|X_\lambda(t)\|_{H^2(\mathcal{O})}^2 dt < \infty\end{aligned}$$

by (5.9), and if $N = 2$

$$\mathbb{E}\int_0^{\tau_K} \|X_\lambda(t)\|_1^2 dt < \infty.$$

Hence, by Theorem 2.1 in [28], applied with

$$\begin{aligned}f_t &:= \widetilde{\psi}_\lambda(\nabla X_\lambda(t)) \ (\leq 1 + \lambda|\nabla X_\lambda(t)|) \\ f_t^\circ &:= 0 \\ g_t^h &:= \mu_k e_k X_\lambda(t),\end{aligned}$$

we have the following Itô formula for the $L^N(\mathcal{O})$-norm $\mathbb{P}$-a.s.

$$\begin{aligned}&|X_\lambda(t\wedge\tau_K)|_N^N + N(N-1)\int_{s\wedge\tau_K}^{t\wedge\tau_K}\int_{\mathcal{O}} |X_\lambda(r)|^{N-2}\nabla X_\lambda(r)\cdot\widetilde{\psi}_\lambda(\nabla X_\lambda(r))d\xi\, dr \\ &\quad = |X_\lambda(s\wedge\tau_K)|_N^N + \frac{1}{2}N(N-1)\int_{s\wedge\tau_K}^{t\wedge\tau_K}\int_{\mathcal{O}} \mu|X_\lambda(r)|^N d\xi\, dr \\ &\quad +N\int_{s\wedge\tau_K}^{t\wedge\tau_K} \left\langle |X_\lambda(r)|^{N-2}X_\lambda(r), X_\lambda(r)dW(r)\right\rangle,\ \forall s,t\in[0,T],\ s\leq t.\end{aligned} \tag{7.4}$$

Since, by interpolation (cf. [32, Theorem 2.1])

$$\begin{aligned}\mathbb{E}\int_0^T |X_\lambda(r)|_3^6 dr &\leq C\mathbb{E}\int_0^T (\|X_\lambda(r)\|_{H^2(\mathcal{O})}^{\frac{3}{2}} |X_\lambda(r)|_2^{\frac{9}{2}})dr \\ &\leq C\mathbb{E}\int_0^T (\|X_\lambda(r)\|_{H^2(\mathcal{O})}^2 + |X_\lambda(r)|_2^9)dr,\end{aligned}$$

and, since by (5.9) and (5.5) the last term is finite, we can let $K \to \infty$ in (7.4) to obtain

$$\begin{aligned}
&|X_\lambda(t)|_N^N + N(N-1)\int_s^t \int_{\mathcal{O}} |X_\lambda(r)|^{N-2}\nabla X_\lambda(r)\cdot \widetilde{\psi}_\lambda(\nabla X_\lambda(r))d\xi\, dr \\
&\quad = |X_\lambda(s)|_N^N + \frac{1}{2}\,N(N-1)\int_s^t\int_{\mathcal{O}} \mu |X_\lambda(r)|^N d\xi\, dr \qquad (7.5)\\
&\quad +N\int_s^t \left\langle |X_\lambda(r)|^{N-2}X_\lambda(r), X_\lambda(r)dW(r)\right\rangle,\ \forall s,t\in[0,T],\ s\le t.
\end{aligned}$$

To prove (i), we recall that $\widetilde{\psi}_\lambda(u)\cdot u \ge 0$ for all $u\in\mathbb{R}^N$. Hence, (7.5) implies that

$$\begin{aligned}
&e^{-NC^*t}|X_\lambda(t)|_N^N \le e^{-NC^*s}|X_\lambda(s)|_N^N \\
&\quad +N\int_s^t e^{-NC^*r}\left\langle |X_\lambda(r)|^{N-2}X_\lambda(r), X_\lambda(r)dW(r)\right\rangle,\ \forall s,t\in[0,T],\ s\le t,
\end{aligned}$$

from which assertion (i) follows.

To prove (ii), we recall that $\widetilde{\psi}_\lambda(u)\cdot u \ge |u|-\lambda$. Hence, we have

$$\begin{aligned}
&(N-1)|X_\lambda|^{N-2}\nabla X_\lambda(r)\cdot\widetilde{\psi}_\lambda(\nabla X_\lambda) \\
&\quad \ge (N-1)|X_\lambda|^{N-2}(|\nabla X_\lambda|-\lambda) \qquad (7.6)\\
&\quad = |\nabla(|X_\lambda|^{N-1})| - (N-1)\lambda|X_\lambda|^{N-2}.
\end{aligned}$$

Hence, the second term on the left hand side of (7.5) is bigger than

$$N\rho\int_s^t\int_{\mathcal{O}} |X_\lambda(r)|_N^{N-1}dr - N(N-1)\lambda\int_s^t |X_\lambda(r)|_{N-2}^{N-2}dr,$$

where we used Sobolev's embedding theorem in $W_0^{1,1}(\mathcal{O})$, i.e.,

$$\rho|y|_{\frac{N}{N-1}} \le \|y\|_{1,1},\ \forall y\in W_0^{1,1}(\mathcal{O}),$$

in the last step. Plugging this into (7.5) implies the assertion of the lemma. □

Lemma 7.3 *Let $x, y \in H_0^1(\mathcal{O})$ and let X_λ^x, X_λ^y denote the solutions to* (5.1) *with initial conditions x, y, respectively. Then $\mathbb{P}$-a.s.*

$$\begin{aligned}
&|X_\lambda^x(t)-X_\lambda^y(t)|_N^N \le |X_\lambda^x(s)-X_\lambda^y(s)|_N^N + NC^* \int_s^t |X_\lambda^x(r)-X_\lambda^y(r)|_N^N dr \\
&+N\int_s^t \left\langle |X_\lambda^x(r)-X_\lambda^y(r)|^{N-2}(X_\lambda^x(r)-X_\lambda^y(r)), (X_\lambda^x(r)-X_\lambda^y(r))dW(r)\right\rangle, \\
&\hfill \forall s,t \in [0,T],\ s \le t.
\end{aligned} \tag{7.7}$$

Furthermore, for some constant C independent of λ

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_\lambda^x(t) - X_\lambda^y(t)|_N^N\right] \le 2|x-y|_N^N e^{CT},\ \forall x,y \in L^N(\mathcal{O}). \tag{7.8}$$

Proof. (7.7) follows analogously to (7.5), taking into account that $(\widetilde{\varphi}_\lambda(u) - \widetilde{\varphi}_\lambda(v)) \cdot (u-v) \ge 0$, $\forall u, v \in \mathbb{R}^N$. Then, for $x, y \in H_0^1(\mathcal{O})$, (7.8) follows by a standard application of the Burkholder-Davis-Gundy inequality (for $p = 1$). For arbitrary $x, y \in L^N(\mathcal{O})$, (7.8) then follows by (3.6), since $H_0^1(\mathcal{O}) \subset L^N(\mathcal{O})$ densely. $\square$

Proof of Theorem 7.1 (continued). Let $x \in H_0^1(\mathcal{O})$ and X_λ^x be the solution to (5.1) with initial condition x.

Since $x \in H_0^1(\mathcal{O})$, by Lemma 5.3, Remark 3.6 and interpolation we have, for $N = 3$ and some $C > 0$,

$$\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]} |X_\lambda^x(t) - X^x(t)|_N^2\right] \\
&\qquad \le C\left(\mathbb{E}\left[\sup_{t\in[0,T]} |X_\lambda^x(t) - X^x(t)|_2^2\right]\right)^{\frac{1}{2}} \|x\|_1,\ \forall \lambda \in (0,1].
\end{aligned} \tag{7.9}$$

where X^x is the solution to (1.1) with initial condition x.

Hence, by (5.31), for $2 \le N \le 3$,

$$\lim_{\lambda\to 0} \mathbb{E}\left[\sup_{t\in[0,T]} |X_\lambda^x(t) - X^x(t)|_N^2\right] = 0. \tag{7.10}$$

Furthermore, applying Itô's formula to (7.5) and the function $\varphi_\varepsilon(r) = (r+\varepsilon)^{\frac{1}{N}}$, $\varepsilon \in (0,1)$, and proceeding as in the proof of the previous lemma, we obtain $\mathbb{P}$-a.s.

$$\begin{aligned}
&\varphi_\varepsilon(|X_\lambda^x(t)|_N^N) + \rho \int_0^t |X_\lambda^x(r)|_N^{N-1}(|X_\lambda^x(r)|_N^N + \varepsilon)^{-\frac{N-1}{N}} dr \\
&\le \varphi_\varepsilon(|x|_N^N) + C^* \int_0^t |X_\lambda^x(r)|_N dr \\
&+\lambda(N-1) \int_0^t |X_\lambda^x(r)|_{N-2}^{N-2}(|X_\lambda^x(r)|_N^N + \varepsilon)^{-\frac{N-1}{N}} dr \\
&+ \int_0^t \left\langle X_\lambda^x(r)|X_\lambda^x(r)|^{N-2}(|X_\lambda^x(r)|_N^N + \varepsilon)^{-\frac{N-1}{N}}, X_\lambda^x(r)dW(r)\right\rangle, \ \forall t \ge 0.
\end{aligned} \tag{7.11}$$

Taking expectation in (7.11), we see that, by (7.10) and Fatou's Lemma we may then let $\lambda \to 0$ and, subsequently, let $\varepsilon \to 0$ to arrive at

$$e^{-C^* t}\mathbb{E}|X^x(t)|_N + \rho \int_0^t e^{-C^*\theta}\mathbb{P}[|X^x(\theta)|_N > 0]d\theta \le |x|_N, \ \forall t > 0. \tag{7.12}$$

Since for each $t > 0$,

$$\int_0^t e^{-C^*\theta}\mathbb{P}[|X^x(\theta)|_N > 0]d\theta = \sup_{\varepsilon>0} \int_0^t e^{-C^*}\mathbb{E}[|X^x(\theta)|_N(|X^x(\theta)|_N + \varepsilon)^{-1}]d\theta,$$

by (7.8) and Fatou's lemma, (7.12) extends to the solution X^x of (1.1) for arbitrary $x \in L^N(\mathcal{O})$. But, for $x \in L^N(\mathcal{O})$, by Lemma 7.2 (i), (7.8) and (7.10), the process $t \to e^{-C^*t}|X^x(t)|_N$ is an L^1-limit of supermartingales, hence itself a supermartingale. Hence

$$|X^x(t)|_N = 0 \text{ for } t \ge \tau = \inf\{t \ge 0 : |X^x(t)|_N = 0\},$$

and thus $\mathbb{P}[|X^x(\theta)|_N > 0] = \mathbb{P}[\tau > \theta]$. By (7.12), for X^x with $x \in L^N(\mathcal{O})$, this yields $\mathbb{P}[\tau > t] \le \left(\rho \int_0^t e^{-C^*\theta} d\theta\right)^{-1} |x|_N$, as claimed. □

Remark 7.4 In particular, taking $\mu_k = 0$ for all k, implying $C^* = 0$, we have $\tau \le \frac{|x|_N}{\rho}$ and recover the deterministic case ([2]).

8 Appendix 1

Proposition 8.1 below is due to H. Brezis ([16]) who answered a question we raised and we are grateful to him for this.

Proposition 8.1 *Let $\mathcal{O}$ be a bounded, convex domain of $\mathbb{R}^N$, $N \geq 1$, with smooth boundary (of class C^2). Let $J_\varepsilon = (I+\varepsilon A)^{-1}$, where $A = -\Delta$, $D(A) = H_0^1(\mathcal{O}) \cap H^2(\mathcal{O})$ and $\varepsilon > 0$. Then*

$$\int_{\mathcal{O}} |\nabla J_\varepsilon(y)| d\xi \leq \int_{\mathcal{O}} |\nabla y| d\xi, \quad \forall y \in W_0^{1,1}(\mathcal{O}). \tag{8.1}$$

Proof. For simplicity, we shall write here $|\nabla y|$ instead of $|\nabla y|_N$. Rescaling, we can assume $\varepsilon = 1$ and so reduce (8.1) to

$$\int_{\mathcal{O}} |\nabla u| d\xi \leq \int_{\mathcal{O}} |\nabla y| d\xi, \quad \forall y \in W_0^{1,1}(\mathcal{O}), \tag{8.2}$$

where

$$u - \Delta u = y \text{ in } \mathcal{O}; \quad u = 0 \ dS - \text{a.e. on } \partial\mathcal{O}, \tag{8.3}$$

and dS is the surface measure on $\partial\mathcal{O}$. Without loss of generality, we may also assume $y \in C_0^\infty(\mathcal{O})$.

We set

$$\begin{gathered} D_i = \frac{\partial}{\partial \xi_i}, \ D_{ij}^2 = \frac{\partial^2}{\partial \xi_i \partial \xi_j}, \ i,j = 1,...,N, \\ \varphi(\xi) = |\nabla u(\xi)| = \left(\sum_{i=1}^N |D_i u|^2 \right)^{\frac{1}{2}}, \ \varphi_\varepsilon(\xi) = \sqrt{\varepsilon^2 + |\nabla u(\xi)|^2}. \end{gathered} \tag{8.4}$$

We shall prove (8.2) following several steps.

Lemma 8.2 *We have*

$$\frac{\varphi^2}{\varphi_\varepsilon} - \Delta \varphi_\varepsilon \leq |\nabla y| \quad \text{in } \mathcal{O}. \tag{8.5}$$

Proof. By (8.4), we have

$$\varphi_\varepsilon D_j \varphi_\varepsilon = \sum_{i=1}^N D_i u D_{ij}^2 u,$$

which yields

$$(D_j\varphi_\varepsilon)^2 \le \frac{1}{\varphi_\varepsilon^2}\ |\nabla u|^2 \sum_{i=1}^N (D_{ij}^2 u)^2$$

and therefore

$$|\nabla\varphi_\varepsilon|^2 \le \frac{\varphi^2}{\varphi_\varepsilon^2}\ \sum_{i,j=1}^N |D_{ij}^2 u|^2 \le \sum_{i,j=1}^N |D_{ij}^2 u|^2 \ \text{ in } \mathcal{O}. \tag{8.6}$$

We also have

$$\begin{aligned}\varphi_\varepsilon\Delta\varphi_\varepsilon + |\nabla\varphi_\varepsilon|^2 &= \sum_{i,j=1}^N |D_{ij}^2 u|^2 + \sum_{i=1}^N D_i u \Delta D_i u \\ &= \sum_{i,j=1}^N |D_{ij}^2 u|^2 + \sum_{i=1}^N D_i u (D_i u - D_i y) \\ &= \sum_{i,j=1}^N |D_{ij}^2 u|^2 + |\nabla u|^2 - \nabla u\cdot\nabla y \ge |\nabla\varphi_\varepsilon|^2 + \varphi^2 - \varphi|\nabla y|,\end{aligned}$$

where the last inequality follows by (8.6). This yields

$$-\varphi_\varepsilon\Delta\varphi_\varepsilon + \varphi^2 \le \varphi|\nabla y|,$$

which implies (8.5), as claimed.

Assume that $0\in\partial\mathcal{O}$ and represent locally $\partial\mathcal{O} = \{(\xi',\xi_N);\ \xi_N = \gamma(\xi')\}$, where γ is a C^2-function in a neighborhood of 0 in $\mathbb{R}^{N-1}$ and $\gamma(0)=0$, $\nabla\gamma(0)=0$.

Lemma 8.3 *We have*

$$D_N\varphi_\varepsilon(0) = (D_N u)^2(0)(\varepsilon^2 + (D_N u)^2(0))^{-\frac{1}{2}}\Delta_{\xi'}\gamma(0). \tag{8.7}$$

Proof. By (8.4), we have

$$\varphi_\varepsilon D_N\varphi_\varepsilon = \sum_{i=1}^N D_i u D_{Ni} u. \tag{8.8}$$

Since $u=0$ on $\partial\mathcal{O}$, we have

$$u(\xi_1,\xi_2,...,\xi_{N-1},\gamma(\xi_1,\xi_2,...,\xi_{N-1})) = 0$$

and differentiating with respect to ξ_i, $i=1,...,N-1$, yields

$$D_i u + D_N u D_i \gamma \equiv 0,\ \ i=1,...,N-1, \tag{8.9}$$

$$D_{ii}^2 u + 2D_{iN} u D_i \gamma + D_{NN}^2 u (D_i \gamma)^2 + D_N u D_{ii} \gamma \equiv 0. \tag{8.10}$$

By (8.9), (8.10), we get in $\xi'=0,\ \xi_N=0$,

$$D_i u(0)=0, \quad D_{ii} u(0) + D_N u(0) D_{ii} \gamma(0) = 0. \tag{8.11}$$

By (8.11), we have

$$\Delta u(0) = D_{NN} u(0) - D_N u(0) \Delta_{\xi'} \gamma(0),$$

while, by (8.3), we have $\Delta u(0)=0$, which yields

$$D_{NN} u(0) = D_N(0) \Delta_{\xi'} \gamma(0). \tag{8.12}$$

Now, taking (8.8) in 0 and using (8.11), (8.12), we obtain

$$\sqrt{\varepsilon^2 + (D_N u(0))^2}\ D_N \varphi_\varepsilon(0) = D_N u(0) D_{NN} u(0) = (D_N u)^2(0) \Delta_{\xi'} \gamma(0),$$

as claimed.

Proof of Proposition 8.1. Let n be the outward normed to $\partial\mathcal{O}$. We have

$$\frac{\partial \varphi_\varepsilon}{\partial n}\,(0) = -D_N \varphi_\varepsilon(0) = \frac{-(D_N u(0))^2}{\sqrt{\varepsilon^2 + (D_N u)^2(0)}}\ \Delta_{\xi'} \gamma(0). \tag{8.13}$$

On the other hand, since $\mathcal{O}$ is convex, we have $\Delta \gamma(0) \geq 0$ and, therefore,

$$\frac{\partial \varphi_\varepsilon}{\partial n}\,(0) \leq 0. \tag{8.14}$$

Since 0 can be replaced by an arbitrary point of $\partial\mathcal{O}$, we have therefore

$$\frac{\partial \varphi_\varepsilon}{\partial n} \leq 0 \ \text{ on } \partial\mathcal{O}. \tag{8.15}$$

Integrating (8.5) over $\mathcal{O}$, we get

$$\int_{\mathcal{O}} \frac{\varphi^2}{\varphi_\varepsilon}\, d\xi - \int_{\partial\mathcal{O}} \frac{\partial \varphi_\varepsilon}{\partial n}\, dS \leq \int_{\mathcal{O}} |\nabla y| d\xi,$$

and so, by (8.15), we have

$$\int_{\mathcal{O}} \frac{\varphi^2}{\varphi_\varepsilon}\, d\xi \leq \int_{\mathcal{O}} |\nabla y| d\xi.$$

Then, letting $\varepsilon \to 0$, we get (8.2), thereby completing the proof.

Remark 8.4 Proposition 8.1, which has an interest in itself, amounts to saying that the heat flow on convex smooth domains $\mathcal{O}$ is nonexpansive in $W_0^{1,1}(\mathcal{O})$. Analyzing the previous proof, one sees that it remains true for domains with piecewise smooth and convex boundary.

Corollary 8.5 *Let $\mathcal{O}$ be a convex, bounded and open subset of $\mathbb{R}^N$. Then*

$$\phi(J_\varepsilon(u)) \le \phi(u), \quad \forall u \in BV(\mathcal{O}), \tag{8.16}$$

where ϕ is the functional (2.2)*.*

Proof. Let $u \in BV(\mathcal{O})$, $\phi(u) < \infty$. This means that there is $\{u_n\} \subset W_0^{1,1}(\mathcal{O})$ such that $u_n \to u$ in $L^1(\mathcal{O})$ and

$$\phi(u) \ge \limsup_{n\to\infty} \int_{\mathcal{O}} |\nabla u_n| d\xi. \tag{8.17}$$

while

$$\limsup_{n\to\infty} \int_{\mathcal{O}} |\nabla J_\varepsilon(y_n)| d\xi \ge \phi(J_\varepsilon(y)). \tag{8.18}$$

By (8.16), (8.17) and (8.1), it follows (8.16). □

Proposition 8.1 can be extended as follows.

Proposition 8.6 *Let $g : [0,\infty) \to [0,\infty)$ be a continuous and convex function of at most quadratic growth such that $g(0) = 0$. Then*

$$\int_{\mathcal{O}} g(|\nabla J_\varepsilon(y)|) d\xi \le \int_{\mathcal{O}} g(|\nabla y|) d\xi, \ \forall y \in H_0^1(\mathcal{O}). \tag{8.19}$$

Proof. Since g is of at most quadratic growth, as before we may assume that $y \in C_0^\infty(\mathcal{O})$. Furthermore, without loss of generality, we may assume that $g \in C^2([0,\infty))$. (This can be achieved by regularizing the function g.) As in the previous case, it suffices to prove (8.19) for $\varepsilon = 1$. We set

$$\phi(\xi) = g(\varphi(\xi)), \ \phi_\varepsilon(\xi) = g(\varphi_\varepsilon(\xi)), \ \xi \in \mathcal{O},$$

where φ and φ_ε are in (8.4). We have

$$\nabla \phi_\varepsilon = g'(\varphi_\varepsilon) \nabla \varphi_\varepsilon, \ \Delta \phi_\varepsilon = g'(\varphi_\varepsilon) \Delta \varphi_\varepsilon + g''(\varphi_\varepsilon) |\nabla \varphi_\varepsilon|^2, \ \xi \in \mathcal{O},$$

and so, by (8.5),

$$\begin{aligned}\frac{\phi^2}{\phi_\varepsilon}-\Delta\phi_\varepsilon &= g'(\varphi_\varepsilon)\left(\frac{\varphi^2}{\varphi_\varepsilon}-\Delta\varphi_\varepsilon\right)+\frac{g^2(\varphi)}{g(\varphi_\varepsilon)}\\ &\quad -g'(\varphi_\varepsilon)\,\frac{\varphi^2}{\varphi_\varepsilon}-g''(\varphi_\varepsilon)|\nabla\varphi_\varepsilon|^2 \qquad (8.20)\\ &\le g'(\varphi_\varepsilon)|\nabla y|+\frac{g^2(\varphi)}{g(\varphi_\varepsilon)}-g'(\varphi_\varepsilon)\,\frac{\varphi^2}{\varphi_\varepsilon}.\end{aligned}$$

Now, proceeding as in the proof of Proposition 8.1, we take $0\in\partial\mathcal{O}$ and represent locally $\partial\mathcal{O}$ as $\{(\xi',\xi_N);\ \xi_N=\gamma(\xi')\}$, where $\gamma\in C^2$, $\gamma(0)=0$, $\nabla\gamma(0)=0$. By (8.14) and since g is increasing, we have

$$D_N\phi_\varepsilon(0)=g'(\varphi_\varepsilon(0))D_N\varphi_\varepsilon(0)=-g'(\varphi_\varepsilon(0))\,\frac{\partial\varphi_\varepsilon}{\partial n}\,(0)\ge 0.$$

This yields $\frac{\partial\phi_\varepsilon}{\partial n}\,(0)=-D_N\phi_\varepsilon(0)\le 0$ and, therefore, replacing 0 by an arbitrary point of $\partial\mathcal{O}$, we obtain that

$$\frac{\partial\phi_\varepsilon}{\partial n}\le 0\ \text{ on } \partial\mathcal{O}.$$

Integrating (8.20) over $\mathcal{O}$, we therefore get

$$\int_{\mathcal{O}}\frac{\phi^2}{\phi_\varepsilon}\,d\xi\le\int_{\mathcal{O}}\left(g'(\varphi_\varepsilon)\left(|\nabla y|-\frac{\varphi^2}{\varphi_\varepsilon}\right)+\frac{g^2(\varphi)}{g(\varphi_\varepsilon)}\right)d\xi.$$

Letting $\varepsilon\longrightarrow 0$, we see that

$$\int_{\mathcal{O}}g(\varphi)d\xi\le\int_{\mathcal{O}}g'(\varphi)(|\nabla y|-\varphi)d\xi+\int_{\mathcal{O}}g(\varphi)d\xi\le\int_{\mathcal{O}}g(|\nabla y|)d\xi$$

because $g'(u)(u-v)\ge g(u)-g(v)$, $\forall u,v\in\mathbb{R}^+$. This completes the proof of (8.19). □

Let $j_\lambda:\mathbb{R}^N\to\mathbb{R}$ be the Moreau–Yosida approximation from Section 5. Then, since $\nabla j_\lambda=\psi_\lambda$, it is easy to check that

$$j_\lambda(u)=\begin{cases}\dfrac{1}{2\lambda}\,|u|_N^2 & \text{for } |u|_N\le\lambda,\\[2mm] |u|_N-\dfrac{\lambda}{2} & \text{for } |u|_N>\lambda.\end{cases}$$

Corollary 8.7 *For all $\varepsilon > 0$ and $\lambda > 0$, we have*

$$\int_{\mathcal{O}} j_\lambda(\nabla J_\varepsilon(y))d\xi \le \int_{\mathcal{O}} j_\lambda(\nabla y(\xi))d\xi,\ \forall y \in H_0^1(\mathcal{O}). \tag{8.21}$$

Proof. One applies Proposition 8.6 to the function

$$g(r) = \begin{cases} \dfrac{1}{2\lambda}\, r^2 & \text{for } 0 \le r \le \lambda, \\ r - \dfrac{\lambda}{2} & \text{for } r > \lambda. \quad \square \end{cases}$$

Remark 8.8 In Corollary 8.7, the quadratic growth condition on g can be relaxed. If, e.g., g grows at most of order $p \in [1, \infty)$, then

$$\int g(|\nabla J_\varepsilon(y)|)d\xi \le \int g(|\nabla y|)d\xi,\ \forall y \in W_0^{1,p}(\mathcal{O}).$$

In particular, applying Corollary 8.7 to $g(u) = |u|^p$, where $1 \le p < \infty$, we obtain that for each bounded and convex set $\mathcal{O} \subset \mathbb{R}^N$ with C^2-boundary, we have

$$|\nabla J_\varepsilon(y)|_p \le |\nabla y|_p,\ \forall y \in W_0^{1,p}(\mathcal{O}). \tag{8.22}$$

The case $p = \infty$ is also true and was earlier proved by Brezis and Stampacchia ([17]). In other words, the operator A is dissipative in $W_0^{1,p}(\mathcal{O})$ for all $1 \le p \le \infty$.

9 Appendix 2. (Proof of (5.17))

We have, for $y, z \in H_0^1(\mathcal{O})$,

$$\begin{aligned}
&\left\langle \widetilde{A}_\lambda(t)y - \widetilde{A}_\lambda(t)z, y - z \right\rangle \\
&= \int_{\mathcal{O}} (\psi_\lambda(\nabla(e^{W(t)}y)) - \psi_\lambda(\nabla(e^{W(t)}z))) \cdot \nabla(e^{-W(t)}(y - z))d\xi \\
&+\lambda \int_{\mathcal{O}} \nabla(e^{-W(t)} \cdot (y - z)) \cdot \nabla(e^{W(t)}(y - z))d\xi + \frac{1}{2}\int_{\mathcal{O}} \mu|y - z|^2 d\xi \\
&= \int_{\mathcal{O}} (\psi_\lambda(\nabla(e^{W(t)}y)) - \psi_\lambda(\nabla(e^{W(t)}z))) \cdot (\nabla(e^{W(t)}y) - \nabla(e^{W(t)}z))e^{-2W(t)}d\xi \\
&+ \int_{\mathcal{O}} (\psi_\lambda(\nabla(e^{W(t)}y)) - \psi_\lambda(\nabla(e^{W(t)}z)))e^{W(t)}(y - z) \cdot \nabla(e^{-2W(t)})d\xi \\
&+\frac{1}{2}\int_{\mathcal{O}} \mu|y - z|^2 d\xi + \lambda \int_{\mathcal{O}} \nabla(e^{-W(t)}(y - z)) \cdot \nabla(e^{W(t)}(y - z))d\xi
\end{aligned}$$

$$\geq -2\text{Lip}\psi_\lambda \int_{\mathcal{O}} |\nabla(e^{W(t)}(y-z))||y-z|\,|\nabla W(t)|\,e^{-W(t)}d\xi$$

$$+\lambda \int_{\mathcal{O}} |\nabla(y-z)|^2 d\xi - \lambda \int_{\mathcal{O}} (y-z)^2 |\nabla W(t)|^2 d\xi$$

$$\geq -2\text{Lip}\,\psi_\lambda \int_{\mathcal{O}} (|\nabla(y-z)||y-z||\nabla W(t)| + |y-z|^2|\nabla W(t)|^2)d\xi$$

$$+\lambda \int_{\mathcal{O}} (|\nabla(y-z)|^2 - (y-z)^2|\nabla W(t)|^2)d\xi \geq -\delta_t^\lambda(\omega)|y-z|_2^2,\ \forall t \in [0,T],$$

where

$$\delta_t^\lambda(\omega) = \left(\frac{1}{\lambda}(\text{Lip}\,\psi_\lambda)^2 + 2\right)|\nabla W(t)(\omega)|_\infty^2,$$

$\text{Lip}\,\psi_\lambda$ is the Lipschitz constant of ψ_λ and we have used the Young inequality in the last step. Then, (5.17) follows.

References

[1] L. Ambrosio, N. Fusco, D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford Universityt Press, New York, 2000.

[2] F. Andreu, V. Caselles, J. Díaz, J. Mazón, Some qualitative properties for the total variation flow, *J. Funct. Anal.*, 188 (2002), 516-547.

[3] F. Andreu-Vaillo, V. Caselles, J.M. Mazón, *Parabolic Quasilinear Equations Minimizing Linear Growth Functionals*, Progress in Mathematics 223, Birkhäuser, Basel, 2004.

[4] H. Attouch, G. Buttazzo, M. Gerard, *Variational Analysis in Sobolev Spaces and BV Spaces; Applications to PDES and Optimization*, SIAM Series on Optimization, Philadelphia, 2006.

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

[6] T. Barbu, V. Barbu, V. Biga, D. Coca, A PDE variational approach to image denoising and restoration, *Nonlinear Analysis Real World Applications,* 10 (2009), 1351-1361.

[7] V. Barbu, G. Da Prato, M. Röckner, Existence and uniqueness of nonnegative solutions to the stochastic porous media equation, *Indiana Univ. Math. J.*, 57 (2008), 187-212.

[8] V. Barbu, G. Da Prato, M. Röckner, Stochastic nonlinear diffusion equations with singular diffusivity, *SIAM J. Math. Anal.*, vol. 41, 3 (2009), 1106-1120.

[9] V. Barbu, G. Da Prato, M. Röckner, Stochastic porous media equations and self-organized criticality, *Comm. Math. Phys.* 285 (2009), 901-923.

[10] V. Barbu, G. Da Prato, M. Röckner, Finite time extinction of solutions to fast diffusion equation driven by linear multiplicative noise, *J. Math. Anal. Appl.* 389 (2012), 147-164.

[11] V. Barbu, G. Da Prato, M. Röckner, Addendum to "Stochastic nonlinear diffusion equations with singular diffusivity, *SIAM J. Math. Anal.*, vol. 41, 3 (2009), 1106-1120. CRC 701 - Preprint 2011.

[12] V. Barbu, M. Röckner, On a random scaled porous media equation, *J. Differential Equations*, 251 (2011), 2494-2514.

[13] V. Barbu, M. Röckner, Stochastic porous media equations and self-organized criticality: convergence to the critical state in all dimensions, *Commun. Math. Phys.*, 311 (2012), 539-555.

[14] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, 2010.

[15] H. Brezis, *Opérateurs maximaux monotones*, North–Holland, Amsterdam, 1973.

[16] H. Brezis, Personal communication and detailed paper in preparation.

[17] H. Brezis, G. Stampacchia, Sur la régularité de la solution d'inéquations elliptiques, *Bulletin de la SMF*, 96 (1968), 153-180.

[18] A. Chamballe, P.L. Lions, Image recovery via total variation minimization, *Numer. Math.*, 76 (1997), 17-31.

[19] T. Chan, S. Esedogly, F. Park, A. Yip, Total variation image restoration. Overview and recent developments, in *Handbook of Mathematical Models in Computer Vision*, Springer, New York, 2006.

[20] I. Ciotir, J.T. Tolle, Convergence of invariant measures for singular stochastic diffusion equations, *Stoch. Processes and Their Appl.*, 122 (2012), 1998-2017.

[21] G. Da Prato, J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, Cambridge, UK, 1992.

[22] B. Gess, J.M. Tölle, Multivalued singular stochastic evolution inclusion, http:arxiv.org/ads/112.5672

[23] Y. Giga, R. Kobayashi, On constrained equations with singular diffusivity, *Methods and Appl. Analysis*, vol. 10, 2 (2003), 253-278.

[24] M.H. Giga, Y. Giga, Generalized motion by non local curvature in the plane, *Arch. Rat. Mech. Anal.*, 159 (2001), 295-333.

[25] Y. Giga, R.V. Kohn, Scale invariant extinction time estimates for some singular diffusion equations, *Discrete Continuous Dynam. Systems*, Series A, 30 (2011), 509-535.

[26] R. Kobayashi, Y. Giga, Equations with singular diffusivity, *J. Statist. Phys.*, 95 (1999), 1187-1220.

[27] N.V. Krylov, B.L. Rozovskii, Stochastic evolution equations, *Current problems in mathematics*, vol. 14 (russian), *Acad. Nauk SSSR* (1979), 71-147.

[28] N.V. Krylov, Itô's formula for the L_p-norm of a stochastic W^1_p-valued process, *Probab. Theory Relat. Fields*, 147 (2010), 583-605.

[29] E. Pardoux, Equations aux dérivées partielles stochastiques non linéaires monotones, étude de soslutions fortes de type Ito, Thèse Université Paris Sud, Orsay, 1975.

[30] C. Prevot, M. Röckner, *A Concise Course on Stochastic Partial Differential Equations*, Lecture Notes in Math., 1905, Berlin, 2007.

[31] L.I. Rudin, S. Osher, E. Fatemi, Nonlinear total variational based noise removal algorithms, *Physica,* D, 60 (1992), 259-268.

[32] K. Taira, *Analytic Semigroups and Semilinear Initial Boundary Value Problems*, Cambridge University Press, 1995.