Well-posedness for a class of doubly nonlinear stochastic PDEs of divergence type*

Luca Scarpa

e-mail: luca.scarpa.15@ucl.ac.uk

Department of Mathematics, University College London

Gower Street, London WC1E 6BT, United Kingdom

Abstract

We prove well-posedness for doubly nonlinear parabolic stochastic partial differential equations of the form $dX_t - \operatorname{div}\gamma(\nabla X_t)\,dt + \beta(X_t)\,dt \ni B(t,X_t)\,dW_t$, where γ and β are the two nonlinearities, assumed to be multivalued maximal monotone operators everywhere defined on $\mathbb{R}^d$ and $\mathbb{R}$ respectively, and W is a cylindrical Wiener process. Using variational techniques, suitable uniform estimates (both pathwise and in expectation) and some compactness results, well-posedness is proved under the classical Leray-Lions conditions on γ and with no restrictive smoothness or growth assumptions on β. The operator B is assumed to be Hilbert-Schmidt and to satisfy some classical Lipschitz conditions in the second variable.

AMS Subject Classification: 60H15, 35R60, 35K55, 35D30, 47H05, 46N30.

Key words and phrases: doubly nonlinear stochastic equation, divergence, variational approach, existence of solutions, continuous dependence, multiplicative noise.

1 Introduction

In this work, we consider the boundary value problem with homogeneous Dirichlet conditions associated to a doubly nonlinear parabolic stochastic partial differential equation on an smooth bounded domain $D \subseteq \mathbb{R}^d$ of the type

$$dX_t - \operatorname{div}\gamma(\nabla X_t)\,dt + \beta(X_t)\,dt \ni B(t,X_t)\,dW_t \quad \text{in } D \times (0,T)\,, \tag{1.1}$$

$$X(0) = X_0 \quad \text{in } D\,, \tag{1.2}$$

$$X = 0 \quad \text{on } \partial D \times (0,T)\,, \tag{1.3}$$

where γ and β are two maximal monotone operators everywhere defined on $\mathbb{R}^d$ and $\mathbb{R}$, respectively, W is a cylindrical Wiener process, and B is a random time-dependent Hilbert-Schmidt operator (we will state the complete assumptions on the data in the next section).

***Acknowledgments.** The author is very grateful to Carlo Marinelli for suggesting the problem and for his expert advice and fundamental support throughout this project.

We prove existence of global solutions as well as a continuous dependence result using variational techniques (see e.g. the classical works [17, 22, 23] about the variational approach to SPDEs).

The problem (1.1)–(1.3) is very interesting from the mathematical point of view: as a matter of fact, the equation presents two strong nonlinearities. The first one is represented by γ within the divergence operator: in this case, we will need to assume some classical growth assumptions (the so-called Leray-Lions conditions) in order to recover a suitable coercivity on a natural Sobolev space. The other nonlinearity is represented by the operator β: this is treated as generally as possible, with no restriction on the growth and regularity. Because of this generality, the concept of solution and the appropriate estimates are more difficult to achieve, as we will see. We point also out that dealing with maximal monotone graphs makes our analysis absolutely exhaustive. As a matter of fact, in this way we include in our treatment any continuous increasing function β (with any order of growth), as well as every increasing function with a countable number of jumps: indeed, it is a standard matter to see that if β is an increasing function on $\mathbb{R}$ with jumps in $\{x_n\}_{n\in\mathbb{N}}$, one can obtain a maximal monotone graph by setting $\beta(x_n) = [\beta_-(x_n), \beta_+(x_n)]$. Finally, very mild assumptions on the noise are required, so that our results fit to any reasonable random time-dependent Hilbert-Schmidt operator B; in the case of multiplicative noise, only classical Lipschitz continuity hypotheses are in order.

The noteworthy feature of this paper is that problem (1.1)–(1.3) is very general and embraces a wide variety of specific sub-problems which are interesting on their own: consequently, we provide with our treatment a unifying analysis to several cases of parabolic SPDEs. Let us mention now about some of these and the main related literature.

If γ is the identity on $\mathbb{R}^d$, the resulting equation is the classical semilinear SPDE driven by the Laplace operator $dX - \Delta X\, dt + \beta(X)\, dt \ni B\, dW_t$, which has been widely studied. For example, in [21], global existence results of solutions are provided in the semilinear case, with the laplacian being generalized to any suitable linear operator: here, the idea is to doubly approximate the problem, in order to recover more regularity on β and B, to find then appropriate estimates on the approximated solutions and finally to pass to the limit in the equation. Moreover, in [13], mild solutions are obtained under the strong hypotheses that β is a polynomial of odd degree $m > 1$ and B can be written as $(-\Delta)^{-\frac{s}{2}}$ for a suitable s; in [3], existence of mild solutions is proved with no restrictive hypotheses on the growth of β, but imposing some strong continuity assumptions on the stochastic convolution. In [20], well-posedness is established for the semilinear problem in a L^q setting, with β having polynomial growth.

If γ is the monotone function on $\mathbb{R}^d$ given by $\gamma(x) = |x|^{p-2}x$, $x \in \mathbb{R}^d$, for a certain $p \geq 2$, then the term represented by the divergence in (1.1) is the usual p-laplacian: in this case, our equation becomes $dX - \Delta_p X\, dt + \beta(X)\, dt \ni B\, dW_t$, where $\Delta_p \cdot := \operatorname{div}(|\nabla\cdot|^{p-2}\nabla\cdot)$. This problem is far more interesting and complex than the semilinear case since $-\Delta_p$ is nonlinear for any $p > 2$ and consequently (1.1) becomes doubly nonlinear in turn. Among the great literature dealing with this problem, we can mention [18] for example, where the stochastic p-Laplace equation is studied in the singular case $p \in [1, 2)$, and [19] as well.

Let us now briefly outline the structure of the paper and the results that we present.

In section 2, we state the precise assumptions of the work and we accurately describe the general setting: here, the main hypotheses are stated and the variational setting is presented. Furthermore, we outline the four main results: the first theorem ensures that problem (1.1)–(1.3) admits global solutions in a suitable weak variational way in the case of additive noise, the second one is the very natural continuous dependence property with respect to the initial datum and B, the third is the existence result in case of multiplicative noise and the last one states the continuous dependence property with respect to the initial datum in case of multiplicative noise.

Section 3 contains the proof of the existence theorem with additive noise: the main idea is to introduce two approximations on the problem. The first approximation depends on a parameter λ and it is made on the maximal monotone operators β and γ, considering the Yosida approximations, as usual; moreover, a correction term is added in order to recover a suitable coercivity when λ is fixed, and that is going to vanish when taking the limit as $\lambda \searrow 0$. The second approximation depends on a parameter ε and is made on the operator B in order to gain more regularity on the noise. The double approximation is very similar to the one performed in [21]. The general idea is that given a fixed approximation in ε, the approximated noise is regular enough to allow us to pass to the limit pathwise in λ: once this first step is carried out, suitable probability estimates allow us to pass to the limit also in ε. More specifically, the proof of existence consists in obtaining uniform estimates on the approximated solutions, independently of the approximations, and then passing to the limit in the approximated problem. To this purpose, we will recover pathwise estimates which are uniform in λ (but not in ε), and global estimates also in expectation which are uniform both in λ and in ε. The passage to the limit is carried out in two steps: the first is on λ and it is made pathwise, while the second is made on ε and is made globally also in probability. The main idea is to use Itô's formula and some sharp testings to obtain L^1 estimates on the nonlinear terms in β and rely on the Dunford-Pettis theorem to recover a weak compactness, being inspired in this sense by some calculations performed in [3,21].

Section 4 is devoted to proving the continuous dependence result for the additive noise case, which easily follows from the definition of solution itself and a generalized Itô formula, which is accurately proved in the Appendix B.

Section 5 contains the proof of the main result, which ensures that the problem with multiplicative noise is well-posed: here, we build the global solutions step-by-step, proving at each iteration accurate contraction estimates and using classical fixed-points arguments. The continuous dependence follows from the generalized version of Itô's formula contained in Appendix B.

The appendixes A and B contain a version of a variational integration-by-parts formula and the generalized Itô formula, which are widely used throughout the paper: the first one is made pathwise and it is used when passing to the limit on λ in order to identify the limit of the nonlinearity in γ, while the second is a direct generalization of the classical Itô formula in a variational setting, and it is needed in the passage to the limit on ε and in the proof of the continuous dependence. The idea of the proof is to identify accurate approximations on the processes which have to satisfy appropriate conditions, such as linearity, smoothness properties and suitable asymptotical behaviours: in this sense, appropriate elliptic approximations are performed.

2 Setting and main results

In this section we state the precise assumptions on the data of the problem and the concept of solution. Moreover, we present the main results which will be proved in the subsequent sections.

In the entire work, $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a filtered probability space, where the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is assumed to satisfy the so-called "usual conditions" (i.e. it is saturated and right continuous) and $T > 0$ is the fixed final time; moreover, $D \subseteq \mathbb{R}^d$ is a smooth bounded domain and $Q := D \times (0, T)$ is the corresponding space-time cylinder. Furthermore, we set

$$H := L^2(D) \tag{2.1}$$

and we use the symbol $(\cdot, \cdot)$ for the standard inner product of H. Moreover, if U is a Banach space, we simply write $L^p(\Omega; U)$ (without specifying the σ-algebra) to indicate the usual class of Bochner-integrable functions $L^p(\Omega, \mathcal{F}, \mathbb{P}; U)$; when we are referring to the measure space with respect to a particular σ-algebra of the filtration, we write explicitly $L^p(\Omega, \mathcal{F}_t, \mathbb{P}; U)$ for any given $t \in [0, T]$. The symbol $C^0_w([0, T]; U)$ denotes the space of continuous functions from $[0, T]$ to the space U endowed with the weak topology: this means that $u \in C^0_w([0, T]; U)$ if for any $t \in [0, T]$ and $(t_n)_n \subseteq [0, T]$ with $t_n \to t$, then $u(t_n) \rightharpoonup u(t)$ in U. Furthermore, if U is a separable Hilbert Space, we will use the symbols $\mathscr{L}(U, H)$ and $\mathscr{L}_2(U, H)$ to indicate the spaces of the linear continuous operators and Hilbert-Schmidt operators from U to H, respectively.

We write "$\cdot$" for the usual scalar product in $\mathbb{R}^d$, while the symbols $\hookrightarrow$ and $\stackrel{c}{\hookrightarrow}$ indicate a continuous and a compact-continuous inclusion between Banach spaces, respectively. Moreover, for any constant appearing in the paper, we indicate in the subscript any quantity on which the constant depends: for example, we may use the notation $C_{a,b}$ to stress that the constant C only depends on a and b.

We can now specify the main hypotheses of our work. First of all, we introduce

$$\gamma : \mathbb{R}^d \to 2^{\mathbb{R}^d} \quad \text{maximal monotone}\,, \quad D(\gamma) = \mathbb{R}^d\,, \quad 0 \in \gamma(0) \tag{2.2}$$

$$\beta : \mathbb{R} \to 2^{\mathbb{R}} \quad \text{maximal monotone}\,, \quad D(\beta) = \mathbb{R}\,, \quad 0 \in \beta(0) \tag{2.3}$$

$$W \quad \text{cylindrical Wiener process on } U\,, \tag{2.4}$$

where U is a suitable separable Hilbert space. Now, thanks to definition (2.3), the function

$$j : \mathbb{R} \to [0, +\infty) \quad \text{proper, convex, lower semicontinuous}\,, \quad \partial j = \beta\,, \quad j(0) = 0 \tag{2.5}$$

is well defined; furthermore, we make the assumption that also γ is a subdifferential, i.e. that there exists

$$k : \mathbb{R}^d \to [0, +\infty) \quad \text{proper, convex, lower semicontinuous}\,, \quad \partial k = \gamma\,, \quad k(0) = 0\,. \tag{2.6}$$

We denote by k^* and j^* the convex conjugate functions of k and j, respectively, i.e.

$$k^* : \mathbb{R}^d \to [0, +\infty]\,, \quad k^*(r) := \sup_{y\in\mathbb{R}^d}\{r \cdot y - k(y)\}\,, \tag{2.7}$$

$$j^* : \mathbb{R} \to [0, +\infty]\,, \quad j^*(r) := \sup_{y\in\mathbb{R}}\{ry - j(y)\}\,. \tag{2.8}$$

The following facts from convex analysis are well-known (see for example [4, 11]):

$$k(z) + k^*(s) = s \cdot z \iff s \in \partial k(z)\,, \qquad j(y) + j^*(r) = ry \iff r \in \partial j(y)\,, \tag{2.9}$$

$$k(z) + k^*(s) \geq s \cdot z\,, \qquad j(y) + j^*(r) \geq ry \qquad \text{for all } y, r \in \mathbb{R}\,,\ z, s \in \mathbb{R}^d\,. \tag{2.10}$$

Throughout the paper, we will also assume that j is even, i.e.

$$j(x) = j(-x) \quad \text{for every } x \in \mathbb{R}. \tag{2.11}$$

Remark 2.1. Hypothesis (2.11) is needed in order to prove the generalized Itô formula for the solutions of our problem, which will be strongly used throughout the proofs. However, (2.11) can be weakened: the main point is that we only need j to grow at the same rate both at $+\infty$ and at $-\infty$ (cf. [5, p. 429]). In order to simplify the treatment we assume (2.11), but for sake of completeness we mention that we could have required a slightly weaker condition, namely

$$\limsup_{|x|\to+\infty} \frac{j(x)}{j(-x)} < +\infty\,.$$

Now, for every $\delta \in (0, 1)$, we introduce the resolvents and the Yosida approximations of γ and β as

$$J_\delta := (I_d + \delta\gamma)^{-1}\,, \qquad R_\delta := (I_1 + \delta\beta)^{-1}\,, \tag{2.12}$$

$$\gamma_\delta := \frac{I_d - J_\delta}{\delta}\,, \qquad \beta_\delta := \frac{I_1 - R_\delta}{\delta}\,, \tag{2.13}$$

where the symbol I_m stands for the identity in $\mathbb{R}^m$ for any $m \in \mathbb{N}$. Then, for every $\delta \in (0, 1)$, J_δ, R_δ, γ_δ and β_δ are single-valued, with the latter two being $\frac{1}{\delta}$-Lipschitz continuous, and

$$|J_\delta x| \leq |x| \quad \text{for all } x \in \mathbb{R}^d\,, \qquad |R_\delta x| \leq |x| \quad \text{for all } x \in \mathbb{R}\,, \tag{2.14}$$

$$\gamma_\delta(x) \in \gamma\left(J_\delta x\right) \quad \text{for all } x \in \mathbb{R}^d\,, \qquad \beta_\delta(x) \in \beta\left(R_\delta x\right) \quad \text{for all } x \in \mathbb{R} \tag{2.15}$$

(see for example [1, 11]).

As we have anticipated, we need to make some assumptions on the growth of γ, namely the so-called Leray-Lions conditions, which are widely required in the classical literature on elliptic and parabolic PDEs (the reader can refer here to [7–9] for classical examples). More in detail, we suppose that there are positive constants K, D_1, D_2 and an exponent $p \in [2, +\infty)$ such that

$$\sup\{|y| : y \in \gamma(r)\} \leq D_1 \left(1 + |r|^{p-1}\right) \quad \text{for every } r \in \mathbb{R}^d\,, \tag{2.16}$$

$$y \cdot r \geq K|r|^p - D_2 \quad \text{for every } r \in \mathbb{R}^d\,,\ y \in \gamma(r)\,. \tag{2.17}$$

In the sequel, we will write $q := \frac{p}{p-1} \in (1, 2]$ for the conjugate exponent of p.

Finally, we set

$$V := W_0^{1,p}(D) \tag{2.18}$$

and define the divergence operator in the variational sense:

$$-\operatorname{div} : L^q(D)^d \to V^*\,, \quad \langle -\operatorname{div} u, v\rangle := \int_D u \cdot \nabla v\,, \quad u \in L^q(D)^d\,,\ v \in V\,, \tag{2.19}$$

where we have used the symbol $\langle\cdot,\cdot\rangle$ for the duality pairing between V and V^*. Here and in the sequel, we make the natural identification $H \cong H^*$, so that H is continuously embedded in V^*: for every $u \in H$ and $v \in V$, we have $\langle u, v\rangle = (u, v)$. Taking these remarks into account, we have

$$V \overset{c}{\hookrightarrow} H \hookrightarrow V^*\,, \tag{2.20}$$

where the first inclusion is also dense. Moreover, we set

$$V_0 := H^k_0(D)\,, \quad k := \left[\max\left\{\frac{d}{2}, 1+\frac{d}{2}-\frac{d}{p}\right\}\right]+1\;: \tag{2.21}$$

note that with this particular choice of k, the classical results on Sobolev embeddings (see [4, Thm. 1.5] and [16, Thm. 219]) ensure that

$$V_0 \hookrightarrow V \quad \text{densely}\,, \quad V_0 \hookrightarrow L^\infty(D)\,,$$

so that we have

$$V_0 \hookrightarrow V \cap L^\infty(D)\,, \qquad V^*, L^1(D) \hookrightarrow V_0^*\,. \tag{2.22}$$

We can now state the four main results of the paper, which ensure that problem (1.1)–(1.2) is well-posed, both with additive and multiplicative noise.

Theorem 2.2. *In the setting* (2.1)–(2.22)*, assume that*

$$X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)\,, \tag{2.23}$$

$$B \in L^2(\Omega\times(0,T); \mathscr{L}_2(U,H)) \quad \textit{progressively measurable}\,, \tag{2.24}$$

$$\gamma \quad \textit{is single-valued}\;; \tag{2.25}$$

then there exist

$$X \in L^2(\Omega; L^\infty(0,T;H)) \cap L^p(\Omega\times(0,T);V)\,, \qquad X \in C^0_w([0,T];H) \quad \mathbb{P}\text{-}a.s.\,, \tag{2.26}$$

$$\eta \in L^q(\Omega\times(0,T)\times D)^d\,, \tag{2.27}$$

$$\xi \in L^1(\Omega\times(0,T)\times D)\,, \tag{2.28}$$

where X and ξ are predictable, η is adapted, and the following relations hold:

$$X(t) - \int_0^t \operatorname{div}\eta(s)\,ds + \int_0^t \xi(s)\,ds = X_0 + \int_0^t B(s)\,dW_s \quad \textit{in } L^1(D)\cap V^*\,, \quad \textit{for every } t\in[0,T]\,, \quad \mathbb{P}\textit{-almost surely}\,, \tag{2.29}$$

$$\eta \in \gamma(\nabla X) \quad \textit{a.e. in } \Omega\times(0,T)\times D\,, \tag{2.30}$$

$$\xi \in \beta(X) \quad \textit{a.e. in } \Omega\times(0,T)\times D\,, \tag{2.31}$$

$$j(X) + j^*(\xi) \in L^1(\Omega\times(0,T)\times D)\,. \tag{2.32}$$

Furthermore, if hypothesis (2.25) *is not assumed, then the same conclusion is true replacing conditions* (2.26) *and* (2.29) *with, respectively,*

$$X \in L^\infty\left(0,T;L^2(\Omega;H)\right) \cap L^p(\Omega\times(0,T);V) \cap C^0_w\left([0,T];L^2(\Omega;H)\right)\,, \tag{2.33}$$

$$X(t) - \int_0^t \operatorname{div}\eta(s)\,ds + \int_0^t \xi(s)\,ds = X_0 + \int_0^t B(s)\,dW_s \quad \textit{in } L^1(D)\cap V^*\,, \quad \mathbb{P}\textit{-almost surely}\,, \quad \textit{for every } t\in[0,T]\,. \tag{2.34}$$

Remark 2.3. The integral equation (2.29) is satisfied in the dual space V_0^*, but X is not V_0-valued, so that the results provided are not a direct generalization of the classical concept of variational solution (cf. [24]): we can define them as a weaker type of variational solution, in which the integral expression holds in a dual space V_0^*, but the solution takes values only in a space larger than V_0 (V in our case). Nevertheless, the integral formulation (2.29) can be seen as an identity in $L^1(D)$, so that the choice of V_0 turns out to be only a technical device *a posteriori*. The fact that one cannot expect classical variational solutions for this type of problem is due to fact that no hypotheses on the growth of β are assumed (in contrast to a large part of the literature).

Remark 2.4. Let us comment on hypothesis (2.25). The fact that γ is single-valued (thus a continuous function) is needed in order to prove uniqueness for our problem, which in turn ensures some reasonable measurability properties for the processes X, η and ξ, as we will show later on. On the other side, if we do not require (2.25), the measurability of the solutions cannot be shown using the same argument, but it has to be recovered in a different way: however, in this case, the formulation that one obtains is weaker than the previous one, since the passage to the limit has to be carried out in $\Omega \times D$, with $t \in [0, T]$ fixed, and the solution X is found is a larger space.

Theorem 2.5. *In the setting* (2.1)–(2.22)*, assume that*

$$X_0^1, X_0^2 \in L^2\left(\Omega, \mathcal{F}_0, \mathbb{P}; H\right), \tag{2.35}$$

$$B_1, B_2 \in L^2\left(\Omega \times (0,T); \mathscr{L}_2(U,H)\right) \quad \textit{progressively measurable}. \tag{2.36}$$

If hypothesis (2.25) *holds and* (X_1, η_1, ξ_1)*,* (X_2, η_2, ξ_2) *are any two corresponding solutions satisfying* (2.26)–(2.32)*, then there is a constant* $C > 0$ *(independent of the above quantities) such that*

$$\|X_1 - X_2\|_{L^2(\Omega; L^\infty(0,T;H))} \leq C \left\|X_0^1 - X_0^2\right\|_{L^2(\Omega;H)} + C \left\|B_1 - B_2\right\|_{L^2(\Omega\times(0,T);\mathscr{L}_2(U,H))}. \tag{2.37}$$

In this setting, if $X_0^1 = X_0^2$ *and* $B_1 = B_2$*, then* $X_1 = X_2$*,* $\eta_1 = \eta_2$ *and* $\xi_1 = \xi_2$*. Moreover, if hypothesis* (2.25) *is not assumed and* (X_1, η_1, ξ_1)*,* (X_2, η_2, ξ_2) *are any two corresponding solutions satisfying* (2.27)–(2.28) *and* (2.30)–(2.34)*, then*

$$\|X_1 - X_2\|_{L^\infty(0,T;L^2(\Omega;H))} \leq \left\|X_0^1 - X_0^2\right\|_{L^2(\Omega;H)} + \left\|B_1 - B_2\right\|_{L^2(\Omega\times(0,T);\mathscr{L}_2(U,H))}. \tag{2.38}$$

In this setting, if $X_0^1 = X_0^2$ *and* $B_1 = B_2$*, then* $X_1 = X_2$ *and* $-\operatorname{div} \eta_1 + \xi_1 = -\operatorname{div} \eta_2 + \xi_2$*.*

Remark 2.6. The uniqueness result strongly depends on the assumption (2.25). Indeed, if (2.25) is in order, uniqueness holds for the three solution components, separately; on the other side, if we do not assume (2.25), we can only recover uniqueness for X and the joint process $-\operatorname{div} \eta + \xi$. Moreover, note that the nonlinearity γ prevents us from finding a continuous dependence estimate also in the space $L^p(\Omega \times (0,T); V)$ for any $p > 2$. Nevertheless, if $p = 2$ and γ is the identity, the operator $-\Delta$ is linear and we can recover continuous dependence also in $L^2(\Omega \times (0,T); V)$, for which we refer to [21].

Theorem 2.7. *In the setting* (2.1)–(2.22)*, assume that*

$$X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H) \,, \tag{2.39}$$

$$B : \Omega \times [0,T] \times H \to \mathscr{L}_2(U,H) \quad \textit{progressively measurable}\,, \tag{2.40}$$

$$\begin{aligned} \exists\, L_B > 0 \ : \quad \|B(\omega,t,x_1) - B(\omega,t,x_2)\|_{\mathscr{L}_2(U,H)} &\leq L_B \|x_1 - x_2\|_H \\ \textit{for every} \quad (\omega,t) \in \Omega \times [0,T]\,, &\quad x_1, x_2 \in H\,, \end{aligned} \tag{2.41}$$

$$\exists\, R_B > 0 \ : \quad \|B(\omega,t,x)\|_{\mathscr{L}_2(U,H)} \leq R_B (1 + \|x\|_H) \quad \forall\, (\omega,t,x) \in \Omega \times [0,T] \times H\,. \tag{2.42}$$

If hypothesis (2.25) *holds, then there exists a triplet* (X, η, ξ) *satisfying conditions* (2.26)–(2.28)*,* (2.30)–(2.32) *and*

$$\begin{aligned} X(t) - \int_0^t \operatorname{div} \eta(s)\,ds + \int_0^t \xi(s)\,ds = X_0 + \int_0^t B(s,X(s))\,dW_s \quad \textit{in } L^1(D) \cap V^*\,, \\ \textit{for every } t \in [0,T]\,, \quad \mathbb{P}\textit{-almost surely}\,. \end{aligned} \tag{2.43}$$

If hypothesis (2.25) *is not assumed, then the same conclusion is true replacing* (2.26) *with* (2.33)*, and condition* (2.43) *with*

$$\begin{aligned} X(t) - \int_0^t \operatorname{div} \eta(s)\,ds + \int_0^t \xi(s)\,ds = X_0 + \int_0^t B(s,X(s))\,dW_s \quad \textit{in } L^1(D) \cap V^*\,, \\ \mathbb{P}\textit{-almost surely}\,, \quad \textit{for every } t \in [0,T]\,. \end{aligned} \tag{2.44}$$

Theorem 2.8. *In the setting* (2.1)–(2.22)*, let* X_0^1, X_0^2 *satisfy condition* (2.35)*. If hypothesis* (2.25) *holds and* (X_1, η_1, ξ_1)*,* (X_2, η_2, ξ_2) *are any two corresponding solutions satisfying* (2.26)–(2.28)*,* (2.30)–(2.32) *and* (2.43)*, then there is a constant* $C > 0$ *(independent of the above quantities) such that*

$$\|X_1 - X_2\|_{L^2(\Omega; L^\infty(0,T;H))} \leq C \left\|X_0^1 - X_0^2\right\|_{L^2(\Omega;H)} \,. \tag{2.45}$$

In this setting, if $X_0^1 = X_0^2$*, then* $X_1 = X_2$*,* $\eta_1 = \eta_2$ *and* $\xi_1 = \xi_2$*. Moreover, if hypothesis* (2.25) *is not assumed and* (X_1, η_1, ξ_1)*,* (X_2, η_2, ξ_2) *are any two corresponding solutions satisfying* (2.27)–(2.28)*,* (2.30)–(2.33) *and* (2.44)*, then there is a constant* $C > 0$ *(independent of the above quantities) such that*

$$\|X_1 - X_2\|_{L^\infty(0,T; L^2(\Omega;H))} \leq C \left\|X_0^1 - X_0^2\right\|_{L^2(\Omega;H)} \,. \tag{2.46}$$

In this setting, if $X_0^1 = X_0^2$*, then* $X_1 = X_2$ *and* $-\operatorname{div} \eta_1 + \xi_1 = -\operatorname{div} \eta_2 + \xi_2$*.*

Remark 2.9. It is worth recalling the classical approach to problem (1.1)–(1.3) in the deterministic case and the main differences with the stochastic case. The corresponding deterministic problem is

$$\frac{\partial u}{\partial t} - \operatorname{div} \gamma(\nabla u) + \beta(u) \ni f\,, \quad u(0) = u_0\,,$$

with homogeneous boundary conditions for u: here, the classical approach consists in proving that the sum of the two operators $-\operatorname{div}(\nabla \cdot)$ and $\beta(\cdot)$ is m-accretive in a suitable space. To this end, it is well-known that if (i) E is a Banach space with uniformly convex dual E^*, (ii) A and B are two m-accretive sets in $E \times E$, (iii) $D(A) \cap D(B) \neq \emptyset$, (iv)

$\langle Au, J(B_\lambda u)\rangle_E \geq 0$ for every $u \in D(A)$ and $\lambda \in (0,1)$ (where $J : E \to E^*$ is the duality mapping of E and B_λ is the Yosida approximation of B), then $A+B$ is m-accretive in $E \times E$ (see [4, Prop. 3.8]). If we take for example $E = L^s(D)$ for $1 < s < +\infty$, $A = -\operatorname{div}\gamma(\nabla\cdot)$, $B = \beta(\cdot)$ with their natural domains, we only need to check (iv), since (i)–(iii) are clearly satisfied. To this aim, we need to handle the term

$$\int_D -\operatorname{div}\gamma(\nabla u)\phi(\beta_\lambda(u))\,,$$

where $\phi(r) = |r|^{s-2}r$, $r \in \mathbb{R}$, using integration by parts. The first problem occurs if $s < 2$, since in this case the derivative of ϕ explodes at 0; if $s \geq 2$, we can proceed formally and recover

$$\int_D \phi'(\beta_\lambda(u))\beta_\lambda'(u)\gamma(\nabla u)\cdot\nabla u \geq 0\,.$$

The main difficulty is that β_λ is not differentiable, so that one needs to rely on some generalized chain-rules for Lipschitz functions or suitable mollifications of β_λ. The problem can be seen then as a particular case of the general one

$$\frac{\partial u}{\partial t} + Au \ni f\,,$$

with A purely nonlinear (multivalued) operator, for which one can rely on several classical well-posedness results. However, the corresponding general problem in the stochastic case, i.e.

$$du + Au\,dt \ni B\,dW_t\,,$$

does not have a direct counterpart in terms of existence and uniqueness: as a consequence, in our case the proof of m-accretivity is not sufficient to ensure well-posedness, so that one needs to deal with the problem "by hand". To this end, the variational approach is in order.

3 Existence with additive noise

In this section we prove the two existence results contained in Theorem 2.2: as already mentioned, we are going to approximate the problem using two different parameters. Uniform estimates are then proved and we obtain global solutions to the original problem by passing to the limit in a suitable topology.

3.1 The approximated problem

Thanks to (2.24), for every $\varepsilon \in (0,1)$ there exists an operator

$$B^\varepsilon \in L^2\left(\Omega \times (0,T); \mathscr{L}_2(U, V_0)\right) \tag{3.1}$$

such that:

$$B^\varepsilon \to B \quad \text{in } L^2\left(\Omega \times (0,T); \mathscr{L}_2(U,H)\right) \quad \text{as } \varepsilon \searrow 0\,, \tag{3.2}$$

$$\|B^\varepsilon\|_{L^2(\Omega\times(0,T);\mathscr{L}_2(U,H))} \leq \|B\|_{L^2(\Omega\times(0,T);\mathscr{L}_2(U,H))} \quad \text{for every } \varepsilon \in (0,1)\,. \tag{3.3}$$

Indeed, if k is chosen as in (2.21), then the operator $(I-\varepsilon\Delta)^{-k}$ maps H into V_0 for every $\varepsilon>0$, so that it suffices to take $B^\varepsilon := (I-\varepsilon\Delta)^{-k}B$. With this particular choice, using the fact that the operator $(I-\varepsilon\Delta)^{-k} : H \to H$ is a linear contraction converging to the identity in the strong operator topology as $\varepsilon \searrow 0$ and the ideal property of $\mathscr{L}_2(U;H)$ in $\mathscr{L}(U,H)$, we have that (3.1)–(3.3) are satisfied.

For every $\lambda \in (0,1)$ and $\varepsilon \in (0,1)$, let us consider the approximated problem

$$dX^\varepsilon_\lambda - \operatorname{div}[\gamma_\lambda(\nabla X^\varepsilon_\lambda) + \lambda\nabla X^\varepsilon_\lambda]\,dt + \beta_\lambda(X^\varepsilon_\lambda)\,dt = B^\varepsilon\,dW_t \quad \text{in } D\times(0,T)\,, \tag{3.4}$$

$$X^\varepsilon_\lambda(0) = X_0 \quad \text{in } D\,, \tag{3.5}$$

whose integral formulation is given by

$$\begin{aligned} &X^\varepsilon_\lambda(t) - \int_0^t \operatorname{div}[\gamma_\lambda(\nabla X^\varepsilon_\lambda(s))]\,ds - \lambda\int_0^t \Delta X^\varepsilon_\lambda(s)\,ds + \int_0^t \beta_\lambda(X^\varepsilon_\lambda(s))\,ds \\ &\quad = X_0 + \int_0^t B^\varepsilon(s)\,dW_s \quad \text{in } H^{-1}(D)\,, \quad \text{for every } t\in[0,T]\,, \quad \mathbb{P}\text{-almost surely}\,, \end{aligned} \tag{3.6}$$

where here $-\operatorname{div} : L^2(D)^d \to H^{-1}(D)$ and the laplacian is intended in the usual variational way, i.e.

$$-\Delta : H^1_0(D) \to H^{-1}(D)\,, \quad \langle -\Delta u, v\rangle_{H^1_0(D)} := \int_D \nabla u\cdot\nabla v\,, \quad u,v\in H^1_0(D)\,.$$

A unique solution to the approximated problem (3.6) can be easily obtained using the classical results contained in [17] (see also [24, Thm. 4.2.4]). In fact, the operator

$$A_\lambda : H^1_0(D) \to H^{-1}(D)\,, \quad A_\lambda : \phi \mapsto -\operatorname{div}[\gamma_\lambda(\nabla\phi) + \lambda\nabla\phi] + \beta_\lambda(\phi)\,, \tag{3.7}$$

is well-defined thanks to the Lipschitz continuity of β_λ and γ_λ, and problem (3.6) is the variational formulation with respect to the Gelfand triple $H^1_0(D) \hookrightarrow H \hookrightarrow H^{-1}(D)$ of the following:

$$dX^\varepsilon_\lambda + A_\lambda X^\varepsilon_\lambda\,dt = B^\varepsilon\,dW_t \quad \text{in } (0,T)\times D\,, \tag{3.8}$$

$$X^\varepsilon_\lambda(0) = X_0 \quad \text{in } D\,. \tag{3.9}$$

In this setting, we need to check that the operator A_λ satisfies the classical properties of hemicontinuity, monotonicity, coercivity and boundedness, in order to recover solutions of (3.6). The following lemma is straightforward.

Lemma 3.1. *The following conditions are satisfied for every $\lambda\in(0,1)$.*

(H1) (Hemicontinuity). For all $u,v,x\in H^1_0(D)$, the following map is continuous:

$$s \mapsto \langle A_\lambda(u+sv), x\rangle_{H^1_0(D)}\,, \quad s\in\mathbb{R}\,.$$

(H2) (Monotonicity). For all $u,v\in H^1_0(D)$,

$$\langle A_\lambda u - A_\lambda v, u-v\rangle_{H^1_0(D)} \geq 0\,.$$

(H3) (Coercivity). There exists $C_1 > 0$ such that, for all $v \in H^1_0(D)$,

$$\langle A_\lambda v, v\rangle_{H^1_0(D)} \geq C_1 \|v\|^2_{H^1_0(D)} .$$

(H4) (Boundedness). There exists $C_2 > 0$ such that, for all $v \in H^1_0(D)$,

$$\|A_\lambda v\|_{H^{-1}(D)} \leq C_2 \|v\|_{H^1_0(D)} .$$

Proof. For all $u, v, x \in H^1_0(D)$ we have

$$\langle A_\lambda(u+sv), x\rangle_{H^1_0(D)} = \int_D \gamma_\lambda(\nabla(u+sv))\cdot\nabla x + \lambda\int_D \nabla(u+sv)\cdot\nabla x + \int_D \beta_\lambda(u+sv)x\,,$$

so that (H1) is satisfied thanks to the Lipschitz continuity of γ_λ and β_λ. Secondly, (H2) trivially holds using the monotonicity of γ_λ and β_λ. Moreover, for all $v \in H^1_0(D)$, thanks to the monotonicity of γ_λ and β_λ, and the fact that $\gamma(0) \ni 0$ and $\beta(0) \ni 0$, we have

$$\langle A_\lambda v, v\rangle_{H^1_0(D)} = \int_D \gamma_\lambda(\nabla v)\cdot\nabla v + \lambda\int_D |\nabla v|^2 + \int_D \beta_\lambda(v)v \geq \lambda\int_D |\nabla v|^2\,,$$

so that (H3) holds true thanks to the Poincaré inequality. Finally, using the Lipschitz continuity of β_λ and γ_λ and the Hölder inequality, we have for all $u, v \in H^1_0(D)$

$$\begin{aligned}\langle A_\lambda v, u\rangle_{H^1_0(D)} &= \int_D \gamma_\lambda(\nabla v)\cdot\nabla u + \lambda\int_D \nabla v\cdot\nabla u + \int_D \beta_\lambda(v)u \\ &\leq \left(\frac{1}{\lambda}+\lambda\right)\|\nabla v\|_H\|\nabla u\|_H + \frac{1}{\lambda}\|v\|_H\|u\|_H \leq \left(\frac{2}{\lambda}+\lambda\right)\|v\|_{H^1_0(D)}\|u\|_{H^1_0(D)}\,,\end{aligned}$$

from which (H4) follows. □

Lemma 3.1 ensures that, for all $\varepsilon, \lambda \in (0,1)$, there exists a unique adapted process

$$X^\varepsilon_\lambda \in L^2\left(\Omega; C^0([0,T]; H)\right) \cap L^2\left(\Omega\times(0,T); H^1_0(D)\right) \tag{3.10}$$

such that

$$\begin{aligned}&X^\varepsilon_\lambda(t) - \int_0^t \operatorname{div}[\gamma_\lambda(\nabla X^\varepsilon_\lambda(s))]\,ds - \lambda\int_0^t \Delta X^\varepsilon_\lambda(s)\,ds + \int_0^t \beta_\lambda(X^\varepsilon_\lambda(s))\,ds \\ &\quad= X_0 + \int_0^t B^\varepsilon(s)\,dW_s \quad \text{in } H^{-1}(D)\,, \quad \text{for every } t\in[0,T]\,, \quad \mathbb{P}\text{-almost surely}\,.\end{aligned} \tag{3.11}$$

3.2 A priori estimates I

Here we prove uniform pathwise estimates on X^ε_λ, independent of λ (but not of ε), which will allow us to pass to the limit as $\lambda \searrow 0$ in the approximated problem (3.11) with ε fixed.

Let us define, for any $\varepsilon \in (0,1)$,

$$W^\varepsilon_B(t) := \int_0^t B^\varepsilon(s)\,dW_s\,, \quad t\in[0,T]\,. \tag{3.12}$$

Thanks to the Burkholder-Davis-Gundy inequality and condition (2.24) we deduce

$$W_B^\varepsilon \in L^2\left(\Omega; L^\infty(0,T;V_0)\right) . \tag{3.13}$$

In particular, recalling (2.22), we have that

$$W_B^\varepsilon(\omega) \in L^p(0,T;V) \cap L^\infty(Q) \quad \text{for } \mathbb{P}\text{-almost every } \omega \in \Omega . \tag{3.14}$$

Equation (3.11) can be rewritten as

$$\partial_t \left(X_\lambda^\varepsilon - W_B^\varepsilon\right)(t) - \operatorname{div}\left[\gamma_\lambda(\nabla X_\lambda^\varepsilon(t)) + \lambda \nabla X_\lambda^\varepsilon(t)\right] + \beta_\lambda(X_\lambda^\varepsilon(t)) = 0 \quad \text{in } H^{-1}(D)$$

for every $t \in [0,T]$, for any ω out of a set of probability 0 (the symbol ∂_t for the derivative with respect to time makes sense only if applied to the difference $X_\lambda^\varepsilon - W_B^\varepsilon$). Fix now ω and test by $X_\lambda^\varepsilon(t) - W_B^\varepsilon(t)$ (see [2, §1.3]): we obtain

$$\begin{aligned}
\frac{1}{2}\left\|X_\lambda^\varepsilon(t) - W_B^\varepsilon(t)\right\|_H^2 &+ \int_0^t \int_D \gamma_\lambda(\nabla X_\lambda^\varepsilon(s)) \cdot \nabla (X_\lambda^\varepsilon(s) - W_B^\varepsilon(s))\, ds \\
&+ \lambda \int_0^t \int_D \nabla X_\lambda^\varepsilon(s) \cdot \nabla\left(X_\lambda^\varepsilon(s) - W_B^\varepsilon(s)\right)\, ds \\
&+ \int_0^t \int_D \beta_\lambda(X_\lambda^\varepsilon(s))(X_\lambda^\varepsilon(s) - W_B^\varepsilon(s))\, ds = \frac{1}{2}\left\|X_0\right\|_H^2 .
\end{aligned} \tag{3.15}$$

Using the identity $I_d = \lambda\gamma_\lambda + J_\lambda$ and rearranging terms in the previous relation, we have

$$\begin{aligned}
&\frac{1}{2}\left\|X_\lambda^\varepsilon(t) - W_B^\varepsilon(t)\right\|_H^2 + \int_0^t \int_D \gamma_\lambda(\nabla X_\lambda^\varepsilon(s)) \cdot J_\lambda\left(\nabla X_\lambda^\varepsilon(s)\right)\, ds + \lambda \int_0^t \int_D \left|\gamma_\lambda\left(\nabla X_\lambda^\varepsilon(s)\right)\right|^2\, ds \\
&\qquad + \lambda \int_0^t \int_D \left|\nabla X_\lambda^\varepsilon(s)\right|^2 ds + \int_0^t \int_D \beta_\lambda(X_\lambda^\varepsilon(s))(X_\lambda^\varepsilon(s) - W_B^\varepsilon(s))\, ds \\
&= \frac{1}{2}\left\|X_0\right\|_H^2 + \int_0^t \int_D \gamma_\lambda\left(\nabla X_\lambda^\varepsilon(s)\right) \cdot \nabla W_B^\varepsilon(s)\, ds + \lambda \int_0^t \int_D \nabla X_\lambda^\varepsilon(s) \cdot \nabla W_B^\varepsilon(s)\, ds .
\end{aligned}$$

Using the generalized Young inequality of the form $ab \leq \delta\frac{p-1}{p}a^{\frac{p}{p-1}} + C_{\delta,p}b^p$ (for any $a, b, \delta > 0$ and a certain $C_{\delta,p} > 0$) on the second term on the right-hand side, thanks also to hypotheses (2.16)–(2.17) and condition (2.15) we deduce for every $t \in [0,T]$ that

$$\begin{aligned}
&\frac{1}{2}\left\|X_\lambda^\varepsilon(t) - W_B^\varepsilon(t)\right\|_H^2 + K \int_0^t \left\|J_\lambda\left(\nabla X_\lambda^\varepsilon(s)\right)\right\|_{L^p(D)}^p ds + \lambda \int_0^t \left\|\gamma_\lambda\left(\nabla X_\lambda^\varepsilon(s)\right)\right\|_H^2 ds \\
&\qquad + \lambda \int_0^t \left\|\nabla X_\lambda^\varepsilon(s)\right\|_H^2 ds + \int_0^t \int_D \beta_\lambda(X_\lambda^\varepsilon(s))(X_\lambda^\varepsilon(s) - W_B^\varepsilon(s))\, ds \\
&\leq C' + \frac{1}{2}\left\|X_0\right\|_H^2 + \delta\frac{(p-1)D_1}{p} \int_0^t \left\|J_\lambda\left(\nabla X_\lambda^\varepsilon(s)\right)\right\|_{L^p(D)}^p ds + C_{\delta,p} \int_0^t \left\|\nabla W_B^\varepsilon(s)\right\|_{L^p(D)}^p ds \\
&\qquad + \frac{\lambda}{2} \int_0^t \left\|\nabla X_\lambda^\varepsilon(s)\right\|_H^2 ds + \frac{\lambda}{2} \int_0^t \left\|\nabla W_B^\varepsilon(s)\right\|_H^2 ds
\end{aligned}$$

for a positive constants C' independent of λ and ε. Hence, choosing $\delta = \frac{Kp}{2D_1(p-1)}$, we get that, for every $t \in [0,T]$,

$$\begin{aligned}
&\|X_\lambda^\varepsilon(t)\|_H^2 + \frac{K}{2}\int_0^t \|J_\lambda(\nabla X_\lambda^\varepsilon(s))\|_{L^p(D)}^p\,ds + \lambda\int_0^t \|\gamma_\lambda(\nabla X_\lambda^\varepsilon(s))\|_H^2\,ds \\
&\qquad + \frac{\lambda}{2}\int_0^t \|\nabla X_\lambda^\varepsilon(s)\|_H^2\,ds + \int_0^t\int_D \beta_\lambda(X_\lambda^\varepsilon(s))(X_\lambda^\varepsilon(s) - W_B^\varepsilon(s))\,ds \\
&\quad \leq C' + \frac{1}{2}\|X_0\|_H^2 + C_p\|W_B^\varepsilon\|_{L^p(0,T;V)}^p + \frac{1}{2}\|W_B^\varepsilon\|_{L^\infty(0,T;H)}^2 + \frac{1}{2}\|W_B^\varepsilon\|_{L^2(0,T;H_0^1(D))}^2
\end{aligned} \tag{3.16}$$

for a positive constant C_p independent of λ and ε. Denoting by $j_\lambda : \mathbb{R} \to [0,+\infty)$ the proper, convex, lower semicontinuous function such that $\beta_\lambda = \partial j_\lambda$ and $j_\lambda(0) = 0$, one has that $j_\lambda \leq j$ and $j_\lambda(x) \nearrow j(x)$ for every $x \in \mathbb{R}$ (recall that $\mathbb{R} = D(\beta) \subseteq D(j)$). Hence, for every $x, y \in \mathbb{R}$ we have that

$$\beta_\lambda(x)(x-y) \geq j_\lambda(x) - j_\lambda(y) \geq j_\lambda(x) - j(y)\,.$$

Applying this inequality to the last term on the left-hand side of (3.16), we deduce that, for every $t \in [0,T]$,

$$\begin{aligned}
&\|X_\lambda^\varepsilon(t)\|_H^2 + \frac{K}{2}\int_0^t \|J_\lambda(\nabla X_\lambda^\varepsilon(s))\|_{L^p(D)}^p\,ds + \lambda\int_0^t \|\gamma_\lambda(\nabla X_\lambda^\varepsilon(s))\|_H^2\,ds \\
&\qquad + \frac{\lambda}{2}\int_0^t \|\nabla X_\lambda^\varepsilon(s)\|_H^2\,ds + \int_0^t\int_D j_\lambda(X_\lambda^\varepsilon(s))\,ds \\
&\leq C\left(1 + \|X_0\|_H^2 + \|W_B^\varepsilon\|_{L^p(0,T;V)}^2 + \|W_B^\varepsilon\|_{L^\infty(0,T;H)}^2 + \|W_B^\varepsilon\|_{L^2(0,T;H_0^1(D))}^2 + \int_Q j(W_B^\varepsilon)\right)
\end{aligned}$$

for a certain constant $C > 0$. Note that all the terms on the right-hand side are finite $\mathbb{P}$-almost surely: for the first five, this is immediate thanks to (2.23) and (3.14), while $j(W_B^\varepsilon) \in L^1(Q)$ since $W_B^\varepsilon \in L^\infty(Q)$. Using the positivity of j_λ we deduce that for $\mathbb{P}$-almost every $\omega \in \Omega$ there exists a positive constant $M = M_{\omega,\varepsilon}$, independent of λ, such that, for every $\lambda \in (0,1)$,

$$\|X_\lambda^\varepsilon(\omega)\|_{L^\infty(0,T;H)} \leq M_{\omega,\varepsilon}\,, \tag{3.17}$$

$$\|J_\lambda(\nabla X_\lambda^\varepsilon(\omega))\|_{L^p(Q)} \leq M_{\omega,\varepsilon}\,, \tag{3.18}$$

$$\lambda^{1/2}\|\gamma_\lambda(\nabla X_\lambda^\varepsilon(\omega))\|_{L^2(Q)} \leq M_{\omega,\varepsilon}\,, \tag{3.19}$$

$$\lambda^{1/2}\|\nabla X_\lambda^\varepsilon(\omega)\|_{L^2(Q)} \leq M_{\omega,\varepsilon}\,. \tag{3.20}$$

Finally, by (2.17) and (2.15) we also have

$$\int_Q |\gamma_\lambda(\nabla X_\lambda^\varepsilon)|^q \leq D_1\int_Q (1 + |J_\lambda(\nabla X_\lambda^\varepsilon)|)^p\,,$$

so that by (3.18) it follows (possibly redefining $M_{\omega,\varepsilon}$) that, for every $\lambda \in (0,1)$,

$$\|\gamma_\lambda(\nabla X_\lambda^\varepsilon(\omega))\|_{L^q(Q)} \leq M_{\omega,\varepsilon}\,. \tag{3.21}$$

3.3 A priori estimates II

In this section we prove some estimates in expectation on X_λ^ε independent both of λ and ε. The main tool is a version of Itô's formula in a variational framework.

Thanks to conditions (2.23)–(2.24) and (3.10)–(3.11), we can apply Itô's formula (see [24, Thm. 4.2.5]), obtaining

$$\begin{aligned}\frac{1}{2}\|X_\lambda^\varepsilon(t)\|_H^2 &+ \int_0^t\int_D \gamma_\lambda(\nabla X_\lambda^\varepsilon(s))\cdot\nabla X_\lambda^\varepsilon(s)\,ds + \lambda\int_0^t\int_D |\nabla X_\lambda^\varepsilon(s)|^2\,ds\\ &+ \int_0^t\int_D \beta_\lambda(X_\lambda^\varepsilon(s))X_\lambda^\varepsilon(s)\,ds\\ &= \frac{1}{2}\|X_0\|_H^2 + \frac{1}{2}\int_0^t \|B^\varepsilon(s)\|_{\mathscr{L}_2(U,H)}^2\,ds + \int_0^t (X_\lambda^\varepsilon(s), B^\varepsilon(s)\,dW_s)\end{aligned} \tag{3.22}$$

for every $t\in[0,T]$, $\mathbb{P}$-almost surely, which yields, by definition of γ_λ and conditions (2.17) and (2.15),

$$\begin{aligned}\frac{1}{2}\|X_\lambda^\varepsilon(t)\|_H^2 &+ K\int_0^t \|J_\lambda(\nabla X_\lambda^\varepsilon(s))\|_{L^p(D)}^p\,ds + \lambda\int_0^t \|\gamma_\lambda(\nabla X_\lambda^\varepsilon(s))\|_H^2\,ds\\ &+ \lambda\int_0^t \|\nabla X_\lambda^\varepsilon(s)\|_H^2\,ds + \int_0^t\int_D \beta_\lambda(X_\lambda^\varepsilon(s))X_\lambda^\varepsilon(s)\,ds\\ &\le C'' + \frac{1}{2}\|X_0\|_H^2 + \frac{1}{2}\|B^\varepsilon(s)\|_{L^2(0,T;\mathscr{L}_2(U,H))}^2 + \sup_{t\in[0,T]}\left|\int_0^t (X_\lambda^\varepsilon(s), B^\varepsilon(s)\,dW_s)\right|\end{aligned}$$

for a constant $C''>0$, independent of ε and λ. Thanks to Davis' inequality, the Hölder and Young inequalities, and condition (3.3), we have for some $c,\tilde c>0$

$$\begin{aligned}\mathbb{E}\sup_{t\in[0,T]}\left|\int_0^t (X_\lambda^\varepsilon(s), B^\varepsilon(s)\,dW_s)\right| &\le c\mathbb{E}\left[\left(\int_0^T \|X_\lambda^\varepsilon(s)\|_H^2\|B^\varepsilon(s)\|_{\mathscr{L}_2(U,H)}^2\,ds\right)^{1/2}\right]\\ &\le c\mathbb{E}\left[\|X_\lambda^\varepsilon\|_{L^\infty(0,T;H)}\|B^\varepsilon\|_{L^2(0,T;\mathscr{L}_2(U,H))}\right]\\ &\le \frac{1}{4}\|X_\lambda^\varepsilon\|_{L^2(\Omega;L^\infty(0,T;H))}^2 + \tilde c\|B\|_{L^2(\Omega\times(0,T);\mathscr{L}_2(U,H))}^2\,;\end{aligned}$$

consequently, taking the supremum in $t\in[0,T]$ and expectations, we obtain

$$\begin{aligned}\frac{1}{4}\|X_\lambda^\varepsilon\|_{L^2(\Omega;L^\infty(0,T;H))}^2 &+ K\|J_\lambda(\nabla X_\lambda^\varepsilon)\|_{L^p(\Omega\times(0,T)\times D)}^p + \lambda\|\gamma_\lambda(\nabla X_\lambda^\varepsilon)\|_{L^2(\Omega\times(0,T)\times D)}^2\\ &+ \lambda\|\nabla X_\lambda^\varepsilon\|_{L^2(\Omega\times(0,T)\times D)}^2 + \int_{\Omega\times Q}\beta_\lambda(X_\lambda^\varepsilon)X_\lambda^\varepsilon\\ &\le C'' + \frac{1}{2}\|X_0\|_{L^2(\Omega;H)}^2 + \frac{3}{2}\|B\|_{L^2(\Omega\times(0,T);\mathscr{L}_2(U,H))}^2\,.\end{aligned} \tag{3.23}$$

We infer that there exists a constant $N>0$, independent of λ and ε, such that

$$\|X_\lambda^\varepsilon\|_{L^2(\Omega;L^\infty(0,T;H))} \le N\,, \tag{3.24}$$

$$\|J_\lambda(\nabla X_\lambda^\varepsilon)\|_{L^p(\Omega\times(0,T)\times D)} \le N\,, \tag{3.25}$$

$$\lambda^{1/2}\|\gamma_\lambda(\nabla X_\lambda^\varepsilon)\|_{L^2(\Omega\times(0,T)\times D)} \le N\,, \tag{3.26}$$

$$\lambda^{1/2}\|\nabla X_\lambda^\varepsilon\|_{L^2(\Omega\times(0,T)\times D)} \le N\,, \tag{3.27}$$

for every $\varepsilon, \lambda \in (0,1)$. Finally, by (2.17) and (2.15) we also have

$$\int_{\Omega\times Q} |\gamma_\lambda(\nabla X_\lambda^\varepsilon)|^q \leq D_1 \int_{\Omega\times Q} (1 + |J_\lambda(\nabla X_\lambda^\varepsilon)|)^p ,$$

so that by (3.25) it follows (possibly redefining N) that, for every $\varepsilon, \lambda \in (0,1)$,

$$\|\gamma_\lambda(\nabla X_\lambda^\varepsilon)\|_{L^q(\Omega\times(0,T)\times D)} \leq N . \tag{3.28}$$

3.4 A priori estimates III

In this section we prove uniform estimates on the term $\beta_\lambda(X_\lambda^\varepsilon)$, independent of λ (with ε fixed), which are useful to recover a suitable weak compactness. We rely on some computations performed in [3] to obtain some L^1 estimates, the classical results by de la Vallée-Poussin about uniform integrability and on the Dunford-Pettis theorem.

Firstly, let us fix $\omega \in \Omega$. Property (2.9), conditions (2.14)–(2.15) and the monotonicity of β_λ imply that

$$j(R_\lambda X_\lambda^\varepsilon) + j^*(\beta_\lambda(X_\lambda^\varepsilon)) = \beta_\lambda(X_\lambda^\varepsilon) R_\lambda X_\lambda^\varepsilon \leq |\beta_\lambda(X_\lambda^\varepsilon)|\,|X_\lambda^\varepsilon| = \beta_\lambda(X_\lambda^\varepsilon) X_\lambda^\varepsilon .$$

Consequently, from inequality (3.16) evaluated at time T and the previous relation, recalling (3.14) and using the generalized Young inequality of the form $ab \leq j(2a) + j^*(b/2)$ for any $a, b \in \mathbb{R}$ (see (2.10)), we deduce that $\mathbb{P}$-almost surely we have

$$\begin{aligned}
\int_Q j^*(\beta_\lambda(X_\lambda^\varepsilon)) &\leq \int_Q \beta_\lambda(X_\lambda^\varepsilon) X_\lambda^\varepsilon \leq C' + \frac{1}{2}\|X_0\|_H^2 + C_p \|W_B^\varepsilon\|_{L^p(0,T;V)}^p \\
&\qquad + \frac{1}{2}\|W_B^\varepsilon\|_{L^\infty(0,T;H)}^2 + \frac{1}{2}\|W_B^\varepsilon\|_{L^2(0,T;H_0^1(D))}^2 + \int_Q \beta_\lambda(X_\lambda^\varepsilon) W_B^\varepsilon \\
&\leq C' + \frac{1}{2}\|X_0\|_H^2 + C_p \|W_B^\varepsilon\|_{L^p(0,T;V)}^p + \frac{1}{2}\|W_B^\varepsilon\|_{L^\infty(0,T;H)}^2 \\
&\qquad + \frac{1}{2}\|W_B^\varepsilon\|_{L^2(0,T;H_0^1(D))}^2 + \|j\left(2W_B^\varepsilon\right)\|_{L^1(Q)} + \frac{1}{2}\int_Q j^*\left(\beta_\lambda(X_\lambda^\varepsilon)\right) .
\end{aligned}$$

All the terms on the right hand side are finite thanks to (2.23) and (3.14): hence, since j^* is even by assumption, we have proved that

$$\|j^*\left(|\beta_\lambda(X_\lambda^\varepsilon(\omega))|\right)\|_{L^1(Q)} = \|j^*\left(\beta_\lambda(X_\lambda^\varepsilon(\omega))\right)\|_{L^1(Q)} \leq \int_Q \beta_\lambda(X_\lambda^\varepsilon(\omega)) X_\lambda^\varepsilon(\omega) \leq M_{\omega,\varepsilon} \tag{3.29}$$

for $\mathbb{P}$-almost every $\omega \in \Omega$; moreover, since $D(\beta) = \mathbb{R}$ by (2.3), we have that

$$\lim_{|r|\to+\infty} \frac{j^*(r)}{|r|} = +\infty$$

(see for example [4, 11]). Hence, using then the criterion by de la Vallée-Poussin for uniform integrability combined with the Dunford-Pettis theorem, we deduce that, for $\mathbb{P}$-almost every $\omega \in \Omega$ and for every $\varepsilon \in (0,1)$,

$$\{\beta_\lambda(X_\lambda^\varepsilon)(\omega)\}_{\lambda\in(0,1)} \quad \text{is weakly relatively compact in } L^1(Q) . \tag{3.30}$$

Finally, let us obtain the corresponding information also in expectation. It easily follows from (3.23) that there exists a constant $N > 0$, independent of λ and ε, such that

$$\|\beta_\lambda(X_\lambda^\varepsilon)X_\lambda^\varepsilon\|_{L^1(\Omega\times(0,T)\times D)} \leq N \quad \text{for every } \varepsilon, \lambda \in (0,1)\,;$$

hence, in analogy to the derivation of (3.29), we get

$$\int_{\Omega\times Q} j^*(\beta_\lambda(X_\lambda^\varepsilon)) \leq \|\beta_\lambda(X_\lambda^\varepsilon)X_\lambda^\varepsilon\|_{L^1(\Omega\times(0,T)\times D)} \leq N \quad \text{for every } \varepsilon, \lambda \in (0,1)\,. \tag{3.31}$$

Since j^* is even and superlinear at infinity, the criterion by de la Vallée-Poussin and the Dunford-Pettis theorem imply that

$$\{\beta_\lambda(X_\lambda^\varepsilon)\}_{\varepsilon,\lambda\in(0,1)} \quad \text{is weakly relatively compact in } L^1(\Omega\times(0,T)\times D)\,. \tag{3.32}$$

3.5 Passage to the limit as $\lambda \searrow 0$

In this section, we pass to the limit as $\lambda \searrow 0$ in the approximated problem (3.11) with $\varepsilon \in (0,1)$ being fixed: the idea is to pass to the limit pathwise as $\lambda \searrow 0$. Throughout the section, $\varepsilon \in (0,1)$ and $\omega \in \Omega$ are fixed.

First of all, conditions (3.17)–(3.21) and (3.30) ensure that there exist

$$X^\varepsilon(\omega) \in L^\infty(0,T;H)\,, \tag{3.33}$$

$$Y^\varepsilon(\omega) \in L^p(Q)^d\,, \tag{3.34}$$

$$\eta^\varepsilon(\omega) \in L^q(Q)^d\,, \tag{3.35}$$

$$\xi^\varepsilon(\omega) \in L^1(Q) \tag{3.36}$$

and a sequence $\{\lambda_n\}_{n\in\mathbb{N}}$ (which clearly depends on ε and ω as well) such that as $n \to \infty$

$$X_{\lambda_n}^\varepsilon(\omega) \overset{*}{\rightharpoonup} X^\varepsilon(\omega) \quad \text{in } L^\infty(0,T;H)\,, \tag{3.37}$$

$$J_{\lambda_n}\left(\nabla X_{\lambda_n}^\varepsilon(\omega)\right) \rightharpoonup Y^\varepsilon(\omega) \quad \text{in } L^p(Q)^d\,, \tag{3.38}$$

$$\gamma_{\lambda_n}(\nabla X_\lambda^\varepsilon(\omega)) \rightharpoonup \eta^\varepsilon(\omega) \quad \text{in } L^q(Q)^d\,, \tag{3.39}$$

$$\beta_{\lambda_n}(X_{\lambda_n}^\varepsilon(\omega)) \rightharpoonup \xi^\varepsilon(\omega) \quad \text{in } L^1(Q) \tag{3.40}$$

and also as $\lambda \searrow 0$ that

$$\lambda\gamma_\lambda(\nabla X_\lambda^\varepsilon(\omega)) \to 0 \quad \text{in } L^2(Q)^d\,, \tag{3.41}$$

$$\lambda\nabla X_\lambda^\varepsilon(\omega) \to 0 \quad \text{in } L^2(Q)^d\,. \tag{3.42}$$

In particular, since $\lambda^2|\gamma_\lambda(\nabla X_\lambda^\varepsilon)|^2 = |\nabla X_\lambda^\varepsilon - J_\lambda(\nabla X_\lambda^\varepsilon)|^2$, from (3.41) we have that

$$\int_Q |\nabla X_\lambda^\varepsilon - J_\lambda(\nabla X_\lambda^\varepsilon)|^2(\omega) \to 0 \quad \text{as } \lambda \searrow 0\,,$$

which together with (3.38) implies that $\nabla X_{\lambda_n}^\varepsilon(\omega) \rightharpoonup Y^\varepsilon$ in $L^2(Q)^d$; hence, we deduce

$$X^\varepsilon(\omega) \in L^p(0,T;V)\,, \tag{3.43}$$

$Y^\varepsilon = \nabla X^\varepsilon$ and as a consequence (possibly renominating $\{\lambda_n\}_{n\in\mathbb{N}}$)

$$J_{\lambda_n}\left(\nabla X^\varepsilon_{\lambda_n}(\omega)\right) \rightharpoonup \nabla X^\varepsilon(\omega) \quad \text{in } L^p(Q)^d\,, \tag{3.44}$$

$$\nabla X^\varepsilon_{\lambda_n}(\omega) \rightharpoonup \nabla X^\varepsilon(\omega) \quad \text{in } L^2(Q)^d\,. \tag{3.45}$$

The second step is to prove a strong convergence for X^ε_λ. To this purpose, equation (3.11) can be rewritten on the path starting from ω as

$$\partial_t\left(X^\varepsilon_\lambda - W^\varepsilon_B\right)(t) - \operatorname{div}\gamma_\lambda(\nabla X^\varepsilon_\lambda(t)) - \lambda\Delta X^\varepsilon_\lambda(t) + \beta_\lambda(X^\varepsilon_\lambda(t)) = 0 \quad \text{in } H^{-1}(D)$$

for every $t \in [0,T]$: we estimate the different terms of the previous relation in the larger space $L^1(0,T;V_0^*)$. Recalling that $L^1(D), H^{-1}(D), V^* \hookrightarrow V_0^*$, using the fact that $\|-\operatorname{div} v\|_{V^*} \leq \|v\|_{L^q(D)}$ for every $v \in L^q(D)^d$ (thanks to definition (2.19)) and that $\|-\Delta v\|_{H^{-1}(D)} \leq \|\nabla v\|_{L^2(D)}$ for every $v \in H^1_0(D)$, using conditions (3.20)–(3.21) and (3.30), we deduce that for every $\lambda \in (0,1)$

$$\|-\operatorname{div}\gamma_\lambda(\nabla X^\varepsilon_\lambda(\omega))\|_{L^1(0,T;V_0^*)} \leq c\,\|\gamma_\lambda(\nabla X^\varepsilon_\lambda(\omega))\|_{L^q(Q)} \leq M_{\omega,\varepsilon}\,,$$

$$\|-\lambda\Delta X^\varepsilon_\lambda\|_{L^1(0,T;V_0^*)} \leq c\lambda\,\|\nabla X^\varepsilon_\lambda\|_{L^2(Q)} \leq M_{\omega,\varepsilon}\,,$$

$$\|\beta_\lambda(X^\varepsilon_\lambda(\omega))\|_{L^1(0,T;V_0^*)} \leq c\,\|\beta_\lambda(X^\varepsilon_\lambda(\omega))\|_{L^1(Q)} \leq M_{\omega.\varepsilon}\,,$$

for a certain constant $c > 0$ and renominating the constant $M_{\omega,\varepsilon}$ at each passage. Hence, we deduce by difference that

$$\|\partial_t\left(X^\varepsilon_\lambda - W^\varepsilon_B\right)(\omega)\|_{L^1(0,T;V_0^*)} \leq M_{\omega,\varepsilon} \quad \text{for every } \lambda \in (0,1)\,. \tag{3.46}$$

At this point, we can recover a strong convergence using some classical compactness results with $\omega \in \Omega$ being fixed. The proposition that we are going to use is the following (the reader can refer to [25, Cor. 4, p. 85]).

Proposition 3.2. *Let $A_1 \overset{c}{\hookrightarrow} A_2 \hookrightarrow A_3$ be three Banach spaces and let $F \subseteq L^r(0,T;A_1)$ be a bounded set such that $\frac{\partial F}{\partial t} := \{\partial_t f : f \in F\}$ is bounded in $L^1(0,T;A_3)$ for a given $r \geq 1$. Then F is relatively compact in $L^r(0,T;A_2)$.*

In our setting, we make the natural choices $A_1 = H^1_0(D)$, $A_2 = H$, $A_3 = V_0^*$, $r = 2$ and $F = \{(X^\varepsilon_{\lambda_n} - W^\varepsilon_B)(\omega)\}_{n\in\mathbb{N}}$: since by (3.45) the family $\{X^\varepsilon_{\lambda_n}(\omega)\}_{n\in\mathbb{N}}$ is bounded in $L^2(0,T;H^1_0(D))$, thanks also to (3.46) we can apply Proposition 3.2 to recover that the set F is relatively compact in $L^2(0,T;H)$. Hence, there exists $X^\varepsilon_B(\omega) \in L^2(0,T;H)$ such that

$$(X^\varepsilon_{\lambda_n} - W^\varepsilon_B)(\omega) \to X^\varepsilon_B(\omega) \quad \text{in } L^2(0,T;H) \quad \text{as } n \to \infty\,,$$

possibly updating the sequence $\{\lambda_n\}_{n\in\mathbb{N}}$. Using condition (3.37) and the fact that W^ε_B is fixed with respect to λ, we infer that

$$(X^\varepsilon_{\lambda_n} - W^\varepsilon_B)(\omega) \overset{*}{\rightharpoonup} (X^\varepsilon - W^\varepsilon_B)(\omega) \quad \text{in } L^\infty(0,T;H) \quad \text{as } n \to \infty\,,$$

and for uniqueness of the weak limit we have $X^\varepsilon_B(\omega) = (X^\varepsilon - W^\varepsilon_B)(\omega)$ a.e. in Q. As a consequence, we have that

$$X^\varepsilon_{\lambda_n}(\omega) \to X^\varepsilon(\omega) \quad \text{in } L^2(0,T;H) \quad \text{as } n \to \infty\,. \tag{3.47}$$

We are now ready to pass to the limit as $\lambda \searrow 0$ in (3.11): in particular, we are going to show that for every $\varepsilon \in (0,1)$ we have

$$X^\varepsilon(t) - \int_0^t \operatorname{div} \eta^\varepsilon(s)\,ds + \int_0^t \xi^\varepsilon(s)\,ds = X_0 + \int_0^t B^\varepsilon(s)\,dW_s \quad \text{in } V_0^*\,, \qquad \text{for every } t \in [0,T]\,, \quad \mathbb{P}\text{-almost surely}\,, \tag{3.48}$$

$$\eta^\varepsilon \in \gamma(\nabla X^\varepsilon) \quad \text{a.e. in } Q\,, \quad \mathbb{P}\text{-almost surely}\,, \tag{3.49}$$

$$\xi^\varepsilon \in \beta(X) \quad \text{a.e. in } Q\,, \quad \mathbb{P}\text{-almost surely}\,, \tag{3.50}$$

$$j(X^\varepsilon) + j^*(\xi^\varepsilon) \in L^1(Q)\,, \quad \mathbb{P}\text{-almost surely}\,. \tag{3.51}$$

Firstly, let $\varepsilon \in (0,1)$ and $\omega \in \Omega$ be fixed as usual. Let $w \in V_0$ and recall the fact that $V_0 \hookrightarrow L^\infty(D) \cap V$: then, thanks to (3.37), (3.39), (3.42) and (3.40), for almost every $t \in (0,T)$ we have

$$\int_D X^\varepsilon_{\lambda_n}(t) w \to \int_D X^\varepsilon(t) w\,,$$

$$\int_0^t \int_D \gamma_{\lambda_n}(\nabla X^\varepsilon_{\lambda_n}(s)) \cdot \nabla w\,ds \to \int_0^t \int_D \eta^\varepsilon(s) \cdot \nabla w\,ds\,,$$

$$\lambda_n \int_0^t \int_D \nabla X^\varepsilon_{\lambda_n}(s) \cdot \nabla w\,ds \to 0\,,$$

$$\int_0^t \int_D \beta_{\lambda_n}(X^\varepsilon_{\lambda_n}(s)) w\,ds \to \int_0^t \int_D \xi^\varepsilon(s) w\,ds\,,$$

as $n \to \infty$. Hence, taking these remarks into account, letting $n \to \infty$ in equation (3.11) evaluated with λ_n, we obtain exactly

$$X^\varepsilon(t) - \int_0^t \operatorname{div} \eta^\varepsilon(s)\,ds + \int_0^t \xi^\varepsilon(s)\,ds = X_0 + \int_0^t B^\varepsilon(s)\,dW_s \qquad \text{in } V_0^*$$
$$\text{for almost every } t \in (0,T)\,, \quad \mathbb{P}\text{-almost surely}\,.$$

Since all the terms except the first are continuous with respect to time, we deduce a posteriori that $X^\varepsilon(\omega) \in C^0([0,T];V_0^*)$ $\mathbb{P}$-almost surely. Recall now that for any two Banach spaces E_1, E_2 with E_1 reflexive and $E_1 \hookrightarrow E_2$ continuously and densely, we have $L^\infty(0,T;E_1) \cap C^0_w([0,T];E_2) = C^0_w([0,T];E_1)$ (see [26, Thm. 2.1]). Hence, since also $X^\varepsilon(\omega) \in L^\infty(0,T;H)$, we deduce that

$$X^\varepsilon \in C^0_w([0,T];H) \quad \mathbb{P}\text{-almost surely}\,. \tag{3.52}$$

Hence, the last integral relation holds for every $t \in [0,T]$ and (3.48) is proved.

Secondly, let us show (3.50): to this end, we will need the following lemma, due to Brezis (see [10, Thm. 18, p. 126] for a detailed reference).

Lemma 3.3. *Let α be a maximal monotone graph in $\mathbb{R}\times\mathbb{R}$ with $D(\alpha) = \mathbb{R}$ and $0 \in \alpha(0)$. Assume that the sequences $(y_n)_{n\in\mathbb{N}}$, $(g_n)_{n\in\mathbb{N}}$ of real-valued measurable functions on a finite measure space $(Y, \mathscr{A}, \mu)$ are such that $y_n \to y$ μ-a.e. as $n \to \infty$, $g_n \in \alpha(y_n)$ μ-a.e. for all $n \in \mathbb{N}$, and $(g_n y_n)$ is a bounded subset of $L^1(Y, \mathscr{A}, \mu)$. Then there exists $g \in L^1(Y, \mathscr{A}, \mu)$ and a subsequence n' such that $g_{n'} \to g$ weakly in $L^1(Y, \mathscr{A}, \mu)$ as $n' \to \infty$ and $g \in \alpha(y)$ μ-almost everywhere.*

By (3.47) we can assume that $X^\varepsilon_{\lambda_n}(\omega) \to X^\varepsilon(\omega)$ a.e. in Q as $k \to \infty$, from which, since R_{λ_n} is a contraction, we deduce also that $R_{\lambda_n} X^\varepsilon_{\lambda_n}(\omega) \to X^\varepsilon(\omega)$ a.e. in Q. Moreover, by (2.15) and (3.40), we also know that $\beta_{\lambda_n}(X^\varepsilon_{\lambda_n}(\omega)) \in \beta(R_{\lambda_n} X^\varepsilon_{\lambda_n}(\omega))$ and $\beta_{\lambda_n}(X^\varepsilon_{\lambda_n}(\omega)) \rightharpoonup \xi^\varepsilon(\omega)$ in $L^1(Q)$. Consequently, since $\{\beta_{\lambda_n}(X^\varepsilon_{\lambda_n}(\omega)) X^\varepsilon_{\lambda_n}(\omega)\}_{n\in\mathbb{N}}$ is bounded in $L^1(Q)$ thanks to (3.29), we can apply the result contained in Lemma 3.3, with the choices $Y = Q$, μ the Lebesgue measure on Q, $y_n = X^\varepsilon_{\lambda_n}$ and $g_n = R_{\lambda_n} X^\varepsilon_{\lambda_n}$, to infer (3.50).

Furthermore, by definition of β_{λ_n} we have $X^\varepsilon - R_{\lambda_n} X^\varepsilon_{\lambda_n} = (X^\varepsilon - X^\varepsilon_{\lambda_n}) + \lambda_n \beta_{\lambda_n}(X^\varepsilon_{\lambda_n})$, so that thanks to (3.40) and (3.47) we deduce that $R_{\lambda_n} X^\varepsilon_{\lambda_n}(\omega) \to X^\varepsilon(\omega)$ in $L^1(Q)$: hence, by the weak lower semicontinuity of the convex integrals and conditions (3.40), (2.15), (2.9) and (3.29), we have that

$$\begin{aligned}\int_Q [j(X^\varepsilon(\omega)) + j^*(\xi^\varepsilon(\omega))] &\leq \liminf_{n\to\infty} \int_Q \left[j(R_{\lambda_n} X^\varepsilon_{\lambda_n}(\omega)) + j^*(\beta_{\lambda_n}(X^\varepsilon_{\lambda_n})(\omega))\right] \\ = \liminf_{n\to\infty} \int_Q R_{\lambda_n} X^\varepsilon_{\lambda_n}(\omega) \beta_{\lambda_n}(X^\varepsilon_{\lambda_n}(\omega)) &\leq \liminf_{n\to\infty} \int_Q X^\varepsilon_{\lambda_n}(\omega) \beta_{\lambda_n}(X^\varepsilon_{\lambda_n}(\omega)) \leq M_{\omega,\varepsilon}\,,\end{aligned}$$

so that also (3.51) is proved. Let us also point out that conditions (3.50) and (2.9) imply $\xi^\varepsilon X^\varepsilon = j(X^\varepsilon) + j^*(\xi^\varepsilon)$ almost everywhere on Q, so that from the very last calculations, using the fact that R_{λ_n} is a contraction and the monotonicity of β_λ, we have

$$\xi^\varepsilon(\omega) X^\varepsilon(\omega) \in L^1(Q)\,, \qquad \int_Q \xi^\varepsilon(\omega) X^\varepsilon(\omega) \leq \liminf_{n\to\infty} \int_Q \beta_{\lambda_n}(X^\varepsilon_{\lambda_n}(\omega)) X^\varepsilon_{\lambda_n}(\omega)\,. \tag{3.53}$$

Finally, let us show that (3.49) holds: in the next passages, we will omit to write ω to simplify notations. From equation (3.15) evaluated at time T, recalling conditions (3.37), (3.39), (3.40), (3.42), (3.53) and (3.13), we get that

$$\begin{aligned}\limsup_{n\to\infty} \int_Q \gamma_{\lambda_n}(\nabla X^\varepsilon_{\lambda_n}) \cdot \nabla X^\varepsilon_{\lambda_n} &= \frac{1}{2}\|X_0\|^2_H + \lim_{n\to\infty} \int_Q \gamma_{\lambda_n}(\nabla X^\varepsilon_{\lambda_n}) \cdot \nabla W^\varepsilon_B \\ &+ \lim_{n\to\infty} \lambda_n \int_Q \nabla X^\varepsilon_{\lambda_n} \cdot \nabla W^\varepsilon_B + \lim_{n\to\infty} \int_Q \beta_{\lambda_n}(X^\varepsilon_{\lambda_n}) W^\varepsilon_B \\ &- \frac{1}{2} \liminf_{n\to\infty} \left\|X^\varepsilon_{\lambda_n}(T) - W^\varepsilon_B(T)\right\|^2_H - \lim_{n\to\infty} \lambda_n \left\|\nabla X^\varepsilon_{\lambda_n}\right\|^2_H - \liminf_{n\to\infty} \int_Q \beta_{\lambda_n}(X^\varepsilon_{\lambda_n}) X^\varepsilon_{\lambda_n} \\ \leq \frac{1}{2}\|X_0\|^2_H + \int_Q \eta^\varepsilon \cdot \nabla W^\varepsilon_B + \int_Q \xi^\varepsilon W^\varepsilon_B &- \frac{1}{2}\|X^\varepsilon(T) - W^\varepsilon_B(T)\|^2_H - \int_Q \xi^\varepsilon X^\varepsilon\,.\end{aligned}$$

At this point, thanks to conditions (3.48)–(3.51), we can prove that the following testing formula holds:

$$\frac{1}{2}\|X^\varepsilon(T) - W^\varepsilon_B(T)\|^2_H + \int_Q \eta^\varepsilon \cdot \nabla(X^\varepsilon - W^\varepsilon_B) + \int_Q \xi^\varepsilon (X^\varepsilon - W^\varepsilon_B) = \frac{1}{2}\|X_0\|^2_H\,. \tag{3.54}$$

Remark 3.4. The proof of (3.54) relies on sharp approximations of elliptic type and is very technical: hence, we omit it here in order not to make the treatment heavier. The reader can refer to Appendix A for a complete and rigorous proof of (3.54).

Hence, thanks to (3.54), the last set of inequalities can be read as

$$\limsup_{n\to\infty} \int_Q \gamma_{\lambda_n}(\nabla X^\varepsilon_{\lambda_n}) \cdot \nabla X^\varepsilon_{\lambda_n} \leq \int_Q \eta^\varepsilon \cdot \nabla X^\varepsilon\,,$$

from which, using the definition of γ_{λ_n} and condition (3.41) we deduce that

$$\begin{aligned}\limsup_{n\to\infty}\int_Q \gamma_{\lambda_n}(\nabla X^\varepsilon_{\lambda_n})\cdot J_{\lambda_n}\left(\nabla X^\varepsilon_{\lambda_n}\right) &= \limsup_{n\to\infty}\int_Q \left[\gamma_{\lambda_n}(\nabla X^\varepsilon_{\lambda_n})\cdot\nabla X^\varepsilon_{\lambda_n} - \lambda_n|\gamma_{\lambda_n}(\nabla X^\varepsilon_{\lambda_n})|^2\right] \\ &= \limsup_{n\to\infty}\int_Q \gamma_{\lambda_n}(\nabla X^\varepsilon_{\lambda_n})\cdot\nabla X^\varepsilon_{\lambda_n} - \lim_{n\to\infty}\lambda_n\int_Q\left|\gamma_{\lambda_n}(\nabla X^\varepsilon_{\lambda_n})\right|^2 \le \int_Q \eta^\varepsilon\cdot\nabla X^\varepsilon\,.\end{aligned}$$

This last inequality together with (3.38) and (3.39) implies condition (3.49) thanks to the usual tools of monotone analysis.

3.6 Measurability properties of the solutions

In this section, we show that the solution components X^ε, η^ε and ξ^ε constructed in the previous section have also some regularity with respect to ω. Moreover, we prove uniform estimates with respect to ε: to this purpose, we will use the results of Sections 3.3 and 3.4, as well as natural lower semicontinuity properties.

First of all, note that, a priori, X^ε, η^ε and ξ^ε are not even measurable processes, because of the way they have been build (the sequence λ_n could depend on ω as well). To show measurability, we need to prove uniqueness for problem (3.48)–(3.51). Hence, let $(X^\varepsilon_1,\eta^\varepsilon_1,\xi^\varepsilon_1)$ and $(X^\varepsilon_2,\eta^\varepsilon_2,\xi^\varepsilon_2)$ satisfy conditions (3.48)–(3.51): taking the difference of (3.48) and setting $Y^\varepsilon := X^\varepsilon_1 - X^\varepsilon_2$, $\zeta^\varepsilon := \eta^\varepsilon_1-\eta^\varepsilon_2$ and $\psi^\varepsilon := \xi^\varepsilon_1-\xi^\varepsilon_2$ we have

$$Y^\varepsilon(t) - \int_0^t \operatorname{div}\zeta^\varepsilon(s)\,ds + \int_0^t \psi^\varepsilon(s)\,ds = 0 \quad \text{for every } t\in[0,T]\,, \quad \mathbb{P}\text{-almost surely}\,.$$

Now, by convexity we have $j(Y^\varepsilon/2)+j^*(\psi^\varepsilon/2) \le \frac{1}{2}\left(j(X^\varepsilon_1)+j(X^\varepsilon_2)+j^*(\xi^\varepsilon_1)+j^*(\xi^\varepsilon_2)\right)$, where the right-hand side is in $L^1(Q)$: hence, using the same argument as in Appendix A with $X_0=0$ and $B=0$, we infer that

$$\frac{1}{2}\left\|Y^\varepsilon(t)\right\|_H^2 + \int_0^t\int_D \zeta^\varepsilon(s)\cdot\nabla Y^\varepsilon(s)\,ds + \int_0^t\int_D \psi^\varepsilon(s)Y^\varepsilon(s)\,ds = 0\,.$$

The monotonicity of γ and β implies that $Y^\varepsilon=0$. Moreover, in view of (2.25), γ is a continuous function. This implies that $\zeta^\varepsilon=0$ and the first integral expression becomes $\int_0^t\psi^\varepsilon(s)\,ds=0$ for every $t\in[0,T]$, so that also $\psi^\varepsilon=0$ and uniqueness is proved.

At this point, we are ready to prove that the sequence $\{\lambda_n\}_{n\in\mathbb{N}}$ constructed in the previous section can be chosen independent of ω: more precisely, we can prove that for any sequence $\{\lambda_n\}_{n\in\mathbb{N}}$ decreasing to 0, conditions (3.37)–(3.40) and (3.44)–(3.45) hold. Indeed, let $\{\lambda_n\}_{n\in\mathbb{N}}$ be any sequence decreasing to 0 and fix $\omega\in\Omega$: then, for every subsequence of $\{\lambda_n\}_{n\in\mathbb{N}}$ (which we still denote with the same symbol for sake of simplicity), the estimates (3.17)–(3.21) hold. Proceeding as in Section 3.5 and invoking the uniqueness, we can then extract a further sub-subsequence (depending on ω) along which the same weak convergences to X^ε, η^ε and ξ^ε hold. This implies that the convergences (3.37)–(3.40) and (3.44)–(3.45) are true for the original sequence $\{\lambda_n\}_{n\in\mathbb{N}}$, which does not depend on ω.

Now, let us prove some measurability properties of the processes X^ε, η^ε and ξ^ε. First of all, since $X^\varepsilon_{\lambda_n}\to X^\varepsilon$ in $L^2(0,T;H)$ $\mathbb{P}$-almost surely, it is clear that X^ε is predictable (since so are $X^\varepsilon_{\lambda_n}$ for every $n \in \mathbb{N}$). Secondly, let us focus on ξ^ε: we prove that $\beta_{\lambda_n}(X^\varepsilon_{\lambda_n}) \rightharpoonup \xi^\varepsilon$ in $L^1(\Omega \times (0,T) \times D)$. To this aim, for any $g \in L^\infty(Q)$, setting

$$F^\varepsilon_{\lambda_n} := \int_Q \beta_{\lambda_n}(X^\varepsilon_{\lambda_n}) g \,, \qquad F^\varepsilon := \int_Q \xi^\varepsilon g \,,$$

we know that $F^\varepsilon_{\lambda_n} \to F^\varepsilon$ $\mathbb{P}$-almost surely: let us show that $F^\varepsilon_{\lambda_n} \rightharpoonup F^\varepsilon$ in $L^1(\Omega)$. Indeed, for any $h \in L^\infty(\Omega)$, if we define

$$j_0^*(\cdot) := j^*\left(\cdot/M\right) \,, \quad M := \frac{1}{(1 \vee \|g\|_{L^\infty(Q)})(1 \vee \|h\|_{L^\infty(\Omega)})} \,,$$

by the Jensen inequality we have that

$$\begin{aligned} \mathbb{E}\left[j_0^*(F^\varepsilon_{\lambda_n} h)\right] &= \mathbb{E}\left[j_0^*\left(\int_Q \beta_{\lambda_n}(X^\varepsilon_{\lambda_n}) g h\right)\right] \\ &\leq C_{T,|D|} \mathbb{E} \int_Q j_0^*(\beta_{\lambda_n}(X^\varepsilon_{\lambda_n}) g h) \leq \int_{\Omega \times Q} j^*(|\beta_{\lambda_n}(X^\varepsilon_{\lambda_n})|) \,, \end{aligned}$$

where the last term is bounded uniformly in n by (3.31). Consequently, since j_0^* is still superlinear at infinity, by the de la Vallée-Poussin criterion, we deduce that $\{F^\varepsilon_{\lambda_n} h\}_{n\in\mathbb{N}}$ is uniformly integrable on Ω: taking also into account that $F^\varepsilon_{\lambda_n} h \to F^\varepsilon h$ $\mathbb{P}$-almost surely, Vitali's convergence theorem ensures that $F^\varepsilon_{\lambda_n} h \to F^\varepsilon h$ in $L^1(\Omega)$. Since this is true for any h and g, this implies that $\beta_{\lambda_n}(X^\varepsilon_{\lambda_n}) \rightharpoonup \xi^\varepsilon$ in $L^1(\Omega \times (0,T) \times D)$. By Mazur's Lemma there is a sequence made up of convex combinations of $\beta_{\lambda_n}(X^\varepsilon_{\lambda_n})$ which converge strongly ξ^ε in $L^1(Q)$, $\mathbb{P}$-almost surely. This ensures that ξ^ε is predictable (since so are $\beta_\lambda(X^\varepsilon_\lambda)$ for every n). Finally, using a similar argument, one can show also that η^ε is adapted.

It is now time to prove some uniform estimates with respect to ε. By (3.37)–(3.40), (3.44) and the estimates (3.24)–(3.28), using the lower semicontinuity of the norm, we have

$$\begin{aligned} \|X^\varepsilon(\omega)\|_{L^\infty(0,T;H)} &\leq \liminf_{n\to\infty} \left\|X^\varepsilon_{\lambda_n}(\omega)\right\|_{L^\infty(0,T;H)} \,, \\ \|\nabla X^\varepsilon(\omega)\|_{L^p(Q)} &\leq \liminf_{n\to\infty} \left\|J_{\lambda_n}\left(\nabla X^\varepsilon_{\lambda_n}(\omega)\right)\right\|_{L^p(Q)} \,, \\ \|\eta^\varepsilon(\omega)\|_{L^q(Q)} &\leq \liminf_{n\to\infty} \left\|\gamma_{\lambda_n}\left(\nabla X^\varepsilon_{\lambda_n}(\omega)\right)\right\|_{L^q(Q)} \,, \\ \|\xi^\varepsilon(\omega)\|_{L^1(Q)} &\leq \liminf_{n\to\infty} \left\|\beta_{\lambda_n}(X^\varepsilon_{\lambda_n}(\omega))\right\|_{L^1(Q)} \,. \end{aligned}$$

Taking expectations and using (3.24)–(3.28) and (3.32), the Fatou's lemma implies

$$\begin{aligned} \mathbb{E}\,\|X^\varepsilon\|^2_{L^\infty(0,T;H)} &\leq \mathbb{E}\left[\left(\liminf_{n\to\infty}\left\|X^\varepsilon_{\lambda_n}\right\|_{L^\infty(0,T;H)}\right)^2\right] \leq \liminf_{n\to\infty}\left\|X^\varepsilon_{\lambda_n}\right\|^2_{L^2(\Omega;L^\infty(0,T;H))} \leq N \,, \\ \mathbb{E}\,\|\nabla X^\varepsilon\|^p_{L^p(Q)} &\leq \mathbb{E}\left[\left(\liminf_{n\to\infty}\left\|J_{\lambda_n}\left(\nabla X^\varepsilon_{\lambda_n}\right)\right\|_{L^p(Q)}\right)^p\right] \leq \liminf_{n\to\infty}\left\|J_{\lambda_n}\left(\nabla X^\varepsilon_{\lambda_n}\right)\right\|^p_{L^p(\Omega\times Q)} \leq N \,, \\ \mathbb{E}\,\|\eta^\varepsilon\|^q_{L^q(Q)} &\leq \mathbb{E}\left[\left(\liminf_{n\to\infty}\left\|\gamma_{\lambda_n}\left(\nabla X^\varepsilon_{\lambda_n}\right)\right\|_{L^q(Q)}\right)^q\right] \leq \liminf_{n\to\infty}\left\|\gamma_{\lambda_n}\left(\nabla X^\varepsilon_{\lambda_n}\right)\right\|^q_{L^q(\Omega\times Q)} \leq N \,, \\ \mathbb{E}\,\|\xi^\varepsilon\|_{L^1(Q)} &\leq \mathbb{E}\left[\liminf_{n\to\infty}\left\|\beta_{\lambda_n}(X^\varepsilon_{\lambda_n})\right\|_{L^1(Q)}\right] \leq \liminf_{n\to\infty}\left\|\beta_{\lambda_n}(X^\varepsilon_{\lambda_n})\right\|_{L^1(\Omega\times Q)} \leq N \,, \end{aligned}$$

for a certain positive constant N independent of ε. Hence, we have also proved that

$$X^\varepsilon \in L^2\left(\Omega; L^\infty(0,T;H)\right) \cap L^p\left(\Omega \times (0,T); V\right) , \tag{3.55}$$

$$\eta^\varepsilon \in L^q\left(\Omega \times (0,T) \times D\right)^d , \quad \xi^\varepsilon \in L^1\left(\Omega \times (0,T) \times D\right) \tag{3.56}$$

and that the following estimates hold:

$$\|X^\varepsilon\|_{L^2(\Omega;L^\infty(0,T;H))\cap L^p(\Omega\times(0,T);V)} \leq N \quad \text{for every } \varepsilon \in (0,1) , \tag{3.57}$$

$$\|\eta^\varepsilon\|_{L^q(\Omega\times(0,T)\times D)} \leq N \quad \text{for every } \varepsilon \in (0,1) , \tag{3.58}$$

$$\|\xi^\varepsilon\|_{L^1(\Omega\times(0,T)\times D)} \leq N \quad \text{for every } \varepsilon \in (0,1) . \tag{3.59}$$

Moreover, since $\beta_{\lambda_n}(X^\varepsilon_{\lambda_n}) \rightharpoonup \xi^\varepsilon$ in $L^1(Q)$ as $n \to \infty$, $\mathbb{P}$-almost surely, by the weak lower semicontinuity of the convex integral we have

$$\int_Q j^*(\xi^\varepsilon) \leq \liminf_{n\to\infty} \int_Q j^*\left(\beta_{\lambda_n}(X^\varepsilon_{\lambda_n})\right) \quad \mathbb{P}\text{-almost surely} :$$

hence, thanks to the Fatou lemma and condition (3.31), we deduce that

$$\int_{\Omega\times Q} j^*(\xi^\varepsilon) \leq \liminf_{n\to\infty} \int_{\Omega\times Q} j^*\left(\beta_{\lambda_n}(X^\varepsilon_{\lambda_n})\right) \leq N ,$$

where N is independent of ε. Consequently, since j^* is even thanks to (2.11), we have that $\{j^*(\xi^\varepsilon)\}_{\varepsilon\in(0,1)}$ is bounded in $L^1(\Omega \times Q)$: hence, since j^* is superlinear at ∞, the classical results by de la Vallée-Poussin and the Dunford-Pettis theorem ensure that

$$\{\xi^\varepsilon\}_{\varepsilon\in(0,1)} \quad \text{is weakly relatively compact in } L^1(\Omega \times (0,T) \times D) . \tag{3.60}$$

Similarly, $R_{\lambda_n} X^\varepsilon_{\lambda_n} \to X^\varepsilon$ in $L^1(Q)$ and $j(R_\lambda X^\varepsilon_\lambda) \leq j(R_\lambda X^\varepsilon_\lambda) + j^*(\beta_\lambda(X^\varepsilon_\lambda)) = \beta_\lambda(X^\varepsilon_\lambda) X^\varepsilon_\lambda$: hence, the weak lower semicontinuity of the convex integrals, Fatou's lemma and condition (3.31) imply

$$\int_{\Omega\times Q} j(X^\varepsilon) \leq \liminf_{n\to\infty} \int_{\Omega\times Q} j\left(R_{\lambda_n} X^\varepsilon_{\lambda_n}\right) \leq \sup_{\varepsilon,\lambda\in(0,1)} \|\beta_\lambda(X^\varepsilon_\lambda) X^\varepsilon_\lambda\|_{L^1(\Omega\times Q)} \leq N .$$

Taking these remarks into account, we have also obtained that

$$\|j(X^\varepsilon)\|_{L^1(\Omega\times(0,T)\times D)} + \|j^*(\xi^\varepsilon)\|_{L^1(\Omega\times(0,T)\times D)} \leq N \quad \text{for every } \varepsilon \in (0,1) . \tag{3.61}$$

3.7 Passage to the limit as $\varepsilon \searrow 0$

In this section, we pass to the limit as $\varepsilon \searrow 0$ in the sub-prolem (3.48)–(3.51) and we recover global solutions to the original problem: to this end, the passage to the limit takes place also in probability, as we have already anticipated.

First of all, thanks to (3.57)–(3.59), we deduce that there exist

$$X \in L^\infty\left(0,T; L^2(\Omega;H)\right) \cap L^p\left(\Omega \times (0,T); V\right) , \tag{3.62}$$

$$\eta \in L^q\left(\Omega \times (0,T) \times D\right)^d , \qquad \xi \in L^1\left(\Omega \times (0,T) \times D\right) , \tag{3.63}$$

and a sequence $\{\varepsilon_n\}_{n\in\mathbb{N}}$ with $\varepsilon_n \searrow 0$ as $n\to\infty$ such that

$$X^{\varepsilon_n} \overset{*}{\rightharpoonup} X \quad \text{in } L^\infty\left(0,T;L^2(\Omega;H)\right) , \tag{3.64}$$

$$X^{\varepsilon_n} \rightharpoonup X \quad \text{in } L^p\left(\Omega\times(0,T);V\right)) , \tag{3.65}$$

$$\eta^{\varepsilon_n} \rightharpoonup \eta \quad \text{in } L^q\left(\Omega\times(0,T)\times D\right)^d , \tag{3.66}$$

$$\xi^{\varepsilon_n} \rightharpoonup \xi \quad \text{in } L^1\left(\Omega\times(0,T)\times D\right) . \tag{3.67}$$

Let us prove a strong convergence for X^ε: given $\varepsilon,\delta\in(0,1)$, consider equation (3.48) evaluated for ε and δ. Then, taking the difference we have

$$\begin{aligned}X^\varepsilon(t) - X^\delta(t) &- \int_0^t \operatorname{div}(\eta^\varepsilon(s)-\eta^\delta(s))\,ds + \int_0^t \left(\xi^\varepsilon(s)-\xi^\delta(s)\right)\,ds\\ &= \int_0^t \left(B^\varepsilon(s)-B^\delta(s)\right)dW_s \quad \text{in } V_0^* \quad \text{for every } t\in[0,T]\,, \quad \mathbb{P}\text{-a.s}\,.\end{aligned}$$

Now, notice that thanks to (2.11) and the convexity of j and j^*, we have

$$j\left(\frac{X^\varepsilon - X^\delta}{2}\right) + j^*\left(\frac{\xi^\varepsilon-\xi^\delta}{2}\right) \leq \frac{1}{2}\left(j(X^\varepsilon)+j(X^\delta)+j^*(\xi^\varepsilon)+j^*(\xi^\delta)\right) ,$$

where the term on the right hand side is in $L^1(\Omega\times(0,T)\times D)$ thanks to (3.61): hence, recalling also condition (3.52) we can apply Proposition B.1 with the choices $Y = X^\varepsilon - X^\delta$, $f=\eta^\varepsilon-\eta^\delta$, $g=\xi^\varepsilon-\xi^\delta$, $T=B^\varepsilon-B^\delta$ and $\alpha=1/2$ to infer that

$$\begin{aligned}\frac{1}{2}\left\|X^\varepsilon(t)-X^\delta(t)\right\|_H^2 &+ \int_0^t\int_D \left(\eta^\varepsilon(s)-\eta^\delta(s)\right)\cdot\left(\nabla X^\varepsilon(s)-\nabla X^\delta(s)\right)\,ds\\ &+ \int_0^t\int_D \left(\xi^\varepsilon(s)-\xi^\delta(s)\right)\left(X^\varepsilon(s)-X^\delta(s)\right)\,ds\\ = \frac{1}{2}\int_0^t \left\|B^\varepsilon(s)-B^\delta(s)\right\|^2_{\mathscr{L}_2(U,H)}\,ds &+ \int_0^t \left((X^\varepsilon-X^\delta)(s),(B^\varepsilon-B^\delta)(s)\,dW_s\right)\end{aligned}$$

for every $t\in[0,T]$, $\mathbb{P}$-almost surely. Now, proceeding exactly as in Section 3.3, we take the supremum in t and expecations, use the monotonicity of γ and β together with (3.49)–(3.50) and the Davis inequality, so that we have

$$\left\|X^\varepsilon - X^\delta\right\|^2_{L^2(\Omega;L^\infty(0,T;H))} \leq c\left\|B^\varepsilon-B^\delta\right\|^2_{L^2(\Omega\times(0,T);\mathscr{L}_2(U,H))}$$

for every $\varepsilon,\delta\in(0,1)$, for a positive constant c independent of ε: taking into account (3.2), this implies that the sequence $\{X^\varepsilon\}_{\varepsilon\in(0,1)}$ is Cauchy in $L^2(\Omega;L^\infty(0,T;H))$, so that by (3.64) we deduce

$$X\in L^2\left(\Omega;L^\infty(0,T;H)\right) \tag{3.68}$$

and

$$X^\varepsilon \to X \quad \text{in } L^2\left(\Omega;L^\infty(0,T;H)\right) , \quad \text{as } \varepsilon\searrow 0\,. \tag{3.69}$$

We are now ready to pass to the limit in equation (3.48): to this purpose, fix $w\in V_0$ (recall that $V_0\hookrightarrow L^\infty(D)\cap V$). Then, thanks to (3.69), (3.65)–(3.67) and (3.2), for every

$t \in [0,T]$ we have as $n \to \infty$ that

$$\mathbb{E}\left[\operatorname*{ess\ sup}_{t\in(0,T)}\left|\int_D X^{\varepsilon_n}(t)w - \int_D X(t)w\right|\right] \to 0\,,$$
$$\mathbb{E}\left[\int_0^t\int_D \eta^{\varepsilon_n}\cdot\nabla w\,ds\right] \to \mathbb{E}\left[\int_0^t\int_D \eta\cdot\nabla w\,ds\right]\,,$$
$$\mathbb{E}\left[\int_0^t\int_D \xi^{\varepsilon_n}(s)w\,ds\right] \to \mathbb{E}\left[\int_0^t\int_D \xi(s)w\,ds\right]\,,$$
$$\mathbb{E}\left[\int_0^t (w, B^{\varepsilon_n}(s)\,dW_s)\right] \to \mathbb{E}\left[\int_0^t (w, B(s)\,dW_s)\right]\,,$$

so that evaluating (3.48) with ε_n and letting $n \to \infty$, we deduce

$$X(t) - \int_0^t \operatorname{div}\eta(s)\,ds + \int_0^t \xi(s)\,ds = X_0 + \int_0^t B(s)\,dW_s \qquad \text{in } V_0^*\,,$$
$$\text{for almost every } t\in(0,T)\,, \quad \mathbb{P}\text{-almost surely}\,.$$

Since all the terms except the first have $\mathbb{P}$-almost surely continuous paths in V_0^*, we have *a posteriori* that $X \in C^0([0,T];V_0^*)$ $\mathbb{P}$-almost surely. Moreover, it is not difficult to check that the fact that $X^\varepsilon \in C^0_w([0,T];H)$ for every ε together with (3.69) readily implies

$$X \in C^0_w([0,T];H) \quad \mathbb{P}\text{-almost surely}\,, \tag{3.70}$$

so that the integral relation holds for every $t \in [0,T]$ and (2.26)–(2.29) are proved. Furthermore, for every $t\in[0,T]$ and $\mathbb{P}$-almost surely, all the terms in (2.29) except $\int_0^t \eta(s)\,ds$ are in $L^1(D)$ and all the terms except $\int_0^t \xi(s)\,ds$ are in V^*, so that by difference the integral relation holds in $L^1(D)\cap V^*$.

At this point, let us focus on (2.31) and (2.32). By (3.69), we may assume that $X^{\varepsilon_n} \to X$ almost everywhere in $\Omega\times Q$; moreover, by (3.50), (3.61) and (2.15) we have

$$\int_{\Omega\times Q} \xi^\varepsilon X^\varepsilon = \int_{\Omega\times Q} (j(X^\varepsilon) + j^*(\xi^\varepsilon)) \le N\,,$$

where $N > 0$ is independent of ε. Hence, $\{\xi^\varepsilon X^\varepsilon\}_{\varepsilon\in(0,1)}$ is bounded in $L^1(\Omega\times Q)$, and recalling also (3.67) we can apply the result contained in Lemma 3.3, with the choices $Y = \Omega\times Q$, $\mu = \mathbb{P}\otimes \mathrm{Leb}_Q$, $y_n = X^{\varepsilon_n}$ and $g_n = \xi^{\varepsilon_n}$, to infer that (2.31) holds. Moreover, thanks to conditions (3.69), (3.67) and (3.61), using the weak lower semicontinuity of the convex integrals we have that

$$\int_{\Omega\times Q} (j(X) + j^*(\xi)) \le \liminf_{n\to\infty}\int_{\Omega\times Q} (j(X^{\varepsilon_n}) + j^*(\xi^{\varepsilon_n})) \le N\,,$$

so that (2.32) is proved. Let us also point out that from the last inequality, thanks to (2.31), (3.50) and (2.9) we obtain

$$\int_{\Omega\times Q} \xi X \le \liminf_{n\to\infty}\int_{\Omega\times Q} \xi^{\varepsilon_n} X^{\varepsilon_n}\,. \tag{3.71}$$

The next thing that we need to prove is condition (2.30). To this end, thanks to (3.33)–(3.36), (3.43), (3.48)–(3.51) and (3.52), we can apply Proposition B.1 to infer that for every $t \in [0,T]$

$$\frac{1}{2}\|X^{\varepsilon_n}(t)\|^2_{L^2(\Omega;H)} + \int_0^t\int_{\Omega\times D}\eta^{\varepsilon_n}(s)\cdot\nabla X^{\varepsilon_n}(s)\,ds + \int_0^t\int_{\Omega\times D}\xi^{\varepsilon_n}(s)X^{\varepsilon_n}(s)\,ds = \frac{1}{2}\|X_0\|^2_{L^2(\Omega;H)} + \frac{1}{2}\int_0^t \|B^{\varepsilon_n}(s)\|^2_{L^2(\Omega;\mathscr{L}_2(U,H))}\,ds\,,$$

from which, thanks to (3.69), (3.71) and (3.3), we have $\mathbb{P}$-almost surely that

$$\begin{aligned}\limsup_{n\to\infty}\int_{\Omega\times Q}\eta^{\varepsilon_n}\cdot\nabla X^{\varepsilon_n} &= \frac{1}{2}\|X_0\|^2_{L^2(\Omega;H)} + \frac{1}{2}\lim_{n\to\infty}\|B^{\varepsilon_n}\|^2_{L^2(\Omega\times(0,T);\mathscr{L}_2(U,H))}\\ &\quad - \frac{1}{2}\liminf_{n\to\infty}\|X^{\varepsilon_n}(T)\|^2_{L^2(\Omega;H)} - \liminf_{n\to\infty}\int_{\Omega\times Q}\xi^{\varepsilon_n}X^{\varepsilon_n}\\ &\leq \frac{1}{2}\|X_0\|^2_{L^2(\Omega;H)} + \frac{1}{2}\|B\|^2_{L^2(\Omega\times(0,T);\mathscr{L}_2(U,H))} - \frac{1}{2}\|X(T)\|^2_{L^2(\Omega;H)} - \int_{\Omega\times Q}\xi X\,.\end{aligned}$$

Now, thanks to conditions (3.62)–(3.63), (3.70), (2.29) and (2.32), we can apply a second time Proposition B.1 with the choices $Y = X$, $f = \eta$, $g = \xi$ and $T = B$: hence, the right hand side of the last set of inequality is exactly $\int_{\Omega\times Q}\eta\cdot\nabla X$, so that we have

$$\limsup_{n\to\infty}\int_{\Omega\times Q}\eta^{\varepsilon_n}\cdot\nabla X^{\varepsilon_n} \leq \int_{\Omega\times Q}\eta\cdot\nabla X\,.$$

This condition together with (3.65)–(3.66) and (3.49) implies exactly (2.30).

Finally, let us show that X and ξ are predictable processes, and η is adapted. At the end of Section 3.6 we checked that X^ε and ξ^ε are predictable, and η^ε is adapted, for every $\varepsilon \in (0,1)$. Now, from (3.69) it immediately follows that also X is predictable. Moreover, by conditions (3.66)–(3.67) and Mazur's Lemma we can recover strong convergences for some suitable convex combinations of $\{\eta^{\varepsilon_n}\}$ and $\{\xi^{\varepsilon_n}\}$: since these are still adapted and predicable, respectively, we can easily infer that η is adapted and ξ is predictable. This completes the proof.

3.8 The further existence result

In this section we prove the last part of Theorem 2.2, in which condition (2.25) is not assumed anymore. The idea is to to pass to the limit in a different way, using only the estimates in expectations and avoiding the pathwise arguments.

For any $\lambda \in (0,1)$, consider the approximated problem

$$dX_\lambda - \operatorname{div}\gamma_\lambda(\nabla X_\lambda)\,dt - \lambda\Delta X_\lambda\,dt + \beta_\lambda(X_\lambda)\,dt \ni B\,dW_t\ :$$

the classical variational approach in the Gelfand triple $H^1_0(D) \hookrightarrow H \hookrightarrow H^{-1}(D)$ ensures the existence of the approximated solutions

$$X_\lambda \in L^2\left(\Omega; C^0([0,T];H)\right) \cap L^2\left(\Omega\times(0,T); H^1_0(D)\right)\,.$$

Using Itô's formula and proceeding as in Sections 3.3 and 3.4, it is not difficult to prove that there exist a positive constant N, independent of λ, such that

$$\|X_\lambda\|_{L^2(\Omega;L^\infty(0,T;H))} \leq N\,, \qquad \|J_\lambda(\nabla X_\lambda)\|_{L^p(\Omega\times(0,T)\times D)} \leq N\,,$$
$$\|\gamma_\lambda(\nabla X_\lambda)\|_{L^q(\Omega\times(0,T)\times D)} \leq N\,,$$
$$\{\beta_\lambda(X_\lambda)\}_{\lambda\in(0,1)} \quad \text{is weakly relatively compact in } L^1(\Omega\times(0,T)\times D)\,,$$
$$\|j(X_\lambda)\|_{L^1(\Omega\times(0,T)\times D)} + \|j^*(\beta_\lambda(X_\lambda))\|_{L^1(\Omega\times(0,T)\times D)} \leq N\,,$$
$$\lambda^{1/2}\,\|\nabla X_\lambda\|_{L^2(\Omega\times(0,T);H)} \leq N\,,$$
$$\lambda^{1/2}\,\|\gamma_\lambda(\nabla X_\lambda)\|_{L^2(\Omega\times(0,T)\times D)} \leq N\,.$$

We deduce that there exist

$$X \in L^\infty\left(0,T;L^2(\Omega;H)\right) \cap L^p\left(\Omega\times(0,T);V\right)\,,$$
$$\eta \in L^q\left(\Omega\times(0,T)\times D\right)^d\,, \qquad \xi \in L^1\left(\Omega\times(0,T)\times D\right)\,,$$

and a sequence $\{\lambda_n\}_{n\in\mathbb{N}}$ decreasing to 0 such that, as $n\to\infty$,

$$X_{\lambda_n} \overset{*}{\rightharpoonup} X \quad \text{in } L^\infty\left(0,T;L^2(\Omega;H)\right)\,, \qquad J_{\lambda_n}\left(\nabla X_{\lambda_n}\right) \rightharpoonup \nabla X \quad \text{in } L^p\left(\Omega\times(0,T)\times D\right)^d\,,$$
$$\gamma_{\lambda_n}(\nabla X_{\lambda_n}) \rightharpoonup \eta \quad \text{in } L^q\left(\Omega\times(0,T)\times D\right)^d\,, \qquad \beta_{\lambda_n}(X_{\lambda_n}) \rightharpoonup \xi \quad \text{in } L^1\left(\Omega\times(0,T)\times D\right)\,.$$

Fix $w\in L^\infty(\Omega;V_0)$: then, since the four last convergences imply that $X_{\lambda_n}(t)\rightharpoonup X(t)$ in $L^2(\Omega;H)$ for almost every $t\in(0,T)$, we have, as $n\to\infty$,

$$\int_{\Omega\times D} X_{\lambda_n}(t)w \to \int_{\Omega\times D} X(t)w\,,$$
$$\int_0^t\int_{\Omega\times D}\gamma_{\lambda_n}(\nabla X_{\lambda_n})\cdot\nabla w \to \int_0^t\int_{\Omega\times D}\eta\cdot\nabla w\,, \qquad \int_0^t\int_{\Omega\times D}\beta_{\lambda_n}(X_{\lambda_n})w \to \int_0^t\int_{\Omega\times D}\xi w$$

for almost every $t\in(0,T)$. Hence, letting $n\to\infty$, we get, for almost every $t\in(0,T)$,

$$X(t) - \int_0^t \operatorname{div}\eta(s)\,ds + \int_0^t \xi(s)\,ds = X_0 + \int_0^t B(s)\,dW_s \quad \text{in } V_0^*\,, \quad \mathbb{P}\text{-almost surely}\ :$$

since all the terms except the first are continuous with values in $L^1(\Omega;V_0^*)$, we infer also that $X\in C^0([0,T];L^1(\Omega;V_0^*))$ and the integral relation holds for every $t\in[0,T]$. Moreover, since we also have $X\in L^\infty(0,T;L^2(\Omega;H))$, by [26, Thm. 2.1] we can infer that $X\in C^0_w([0,T];L^2(\Omega;H))$.

Secondly, using the weak lower semicontinuity of the convex integrals and the estimates on $j(X_\lambda)$ and $j^*(\beta_\lambda(X_\lambda))$, it is immediate to check that $j(X)+j^*(\xi)\in L^1(\Omega\times Q)$. Furthermore, as we did at the end of Section 3.7, using Mazur's lemma, we deduce also that X and ξ are predictable, and η is adapted.

The last thing that we have to check is that $\eta\in\gamma(\nabla X)$ and $\xi\in\beta(X)$ a.e. in $\Omega\times Q$. To this aim, by the second part of Proposition B.1, using the notation $\eta_\lambda := \gamma_\lambda(\nabla X_\lambda)$, we have that, for every $t\in[0,T]$,

$$\frac{1}{2}\|X_\lambda(t)\|^2_{L^2(\Omega;H)} + \int_0^t\int_{\Omega\times D}\eta_\lambda(s)\cdot\nabla X_\lambda(s)\,ds + \int_0^t\int_{\Omega\times D}\beta_\lambda(X_\lambda)(s)X_\lambda(s)\,ds$$
$$= \frac{1}{2}\|X_0\|^2_{L^2(\Omega;H)} + \frac{1}{2}\int_0^t\|B(s)\|^2_{L^2(\Omega;\mathscr{L}_2(U,H))}\,ds$$

and

$$\frac{1}{2}\|X(t)\|^2_{L^2(\Omega;H)} + \int_0^t\int_{\Omega\times D}\eta(s)\cdot\nabla X(s)\,ds + \int_0^t\int_{\Omega\times D}\xi(s)X(s)\,ds$$
$$= \frac{1}{2}\|X_0\|^2_{L^2(\Omega;H)} + \frac{1}{2}\int_0^t\|B(s)\|^2_{L^2(\Omega;\mathscr{L}_2(U,H))}\,ds\,.$$

We deduce that

$$\begin{aligned}
&\limsup_{n\to\infty}\left[\int_{\Omega\times Q}\eta_{\lambda_n}\cdot\nabla X_{\lambda_n} + \int_{\Omega\times Q}\beta_{\lambda_n}(X_{\lambda_n})X_{\lambda_n}\right]\\
&\quad= \frac{1}{2}\|X_0\|^2_{L^2(\Omega;H)} + \frac{1}{2}\int_0^T\|B(s)\|^2_{L^2(\Omega;\mathscr{L}_2(U,H))}\,ds - \frac{1}{2}\liminf_{n\to\infty}\|X_{\lambda_n}(T)\|^2_{L^2(\Omega;H)}\\
&\quad\le \frac{1}{2}\|X_0\|^2_{L^2(\Omega;H)} + \frac{1}{2}\int_0^T\|B(s)\|^2_{L^2(\Omega;\mathscr{L}_2(U,H))}\,ds - \frac{1}{2}\|X(T)\|^2_{L^2(\Omega;H)}\\
&\quad= \int_{\Omega\times Q}\eta\cdot\nabla X + \int_{\Omega\times Q}\xi X\,.
\end{aligned}$$

Let us identify $\mathbb{R}^d\times\mathbb{R}$ with $\mathbb{R}^{d+1}$, indicate the generic element in $\mathbb{R}^{d+1}$ as a couple (x,y), where $x\in\mathbb{R}^d$ and $y\in\mathbb{R}$, and use the symbol $\bullet$ for the usual scalar product in $\mathbb{R}^{d+1}$. Consider the proper, convex and lower semicontinuous function $\Phi:\mathbb{R}^{d+1}\to[0,+\infty)$ given by $\Phi(x,y):=k(x)+j(y)$, $(x,y)\in\mathbb{R}^{d+1}$: then the subdifferential of Φ is the operator $\Xi:\mathbb{R}^{d+1}\to 2^{\mathbb{R}^{d+1}}$ given by $\Xi(x,y)=\{(u,v)\in\mathbb{R}^{d+1}: u\in\gamma(x), v\in\beta(y)\}$. Hence, recalling that $\beta_\lambda(X_\lambda)R_\lambda X_\lambda = \beta_\lambda(X_\lambda)X_\lambda - \lambda|\beta_\lambda(X_\lambda)|^2\le\beta_\lambda(X_\lambda)X_\lambda$ and similarly $\eta_\lambda\cdot J_\lambda(\nabla X_\lambda)=\eta_\lambda\cdot\nabla X_\lambda-\lambda|\eta_\lambda|^2$, we have proved that

$$\limsup_{n\to\infty}\int_{\Omega\times Q}(\eta_{\lambda_n},\beta_{\lambda_n}(X_{\lambda_n}))\bullet(J_{\lambda_n}(\nabla X_{\lambda_n}),R_{\lambda_n}X_{\lambda_n}) \le \int_{\Omega\times Q}(\eta,\xi)\bullet(\nabla X,X)\,,$$

allowing us to infer that $(\eta,\xi)\in\Xi(\nabla X,X)$, i.e. that $\eta\in\gamma(\nabla X)$ and $\xi\in\beta(X)$ a.e. in $\Omega\times Q$, thanks to the classical results of convex analysis.

4 Continuous dependence on the initial datum with additive noise

This section is devoted to the proof of the continuous dependence and uniqueness results contained in Theorem 2.5. The main tool that we use is the generalized Itô formula contained in Proposition B.1.

We start assuming (2.25): let (X_0^1,B_1), (X_0^2,B_2), (X_1,η_1,ξ_1), (X_2,η_2,ξ_2) be as in Theorem 2.5. Then, writing relation (2.29) for $(X_1,\eta_1,\xi_1,X_0^1,B_1)$ and $(X_2,\eta_2,\xi_2,X_0^2,B_2)$ and taking the difference, $\mathbb{P}$-almost surely we obtain

$$X_1(t)-X_2(t)-\int_0^t\operatorname{div}\left[\eta_1(s)-\eta_2(s)\right]\,ds+\int_0^t(\xi_1(s)-\xi_2(s))\,ds$$
$$=X_0^1-X_0^2+\int_0^t(B_1(s)-B_2(s))\,dW_s \qquad \text{for every } t\in[0,T]\,.$$

Now, we note that thanks to (2.32) and (2.11), for $i = 1, 2$ we have

$$j\left(\frac{X_1 - X_2}{2}\right) + j^*\left(\frac{\xi_1 - \xi_2}{2}\right) \leq \frac{1}{2}\left[j(X_1) + j(X_2) + j^*(\xi_1) + j(\xi_2)\right],$$

where the right hand side is in $L^1(\Omega \times (0,T) \times D)$: hence, we can apply Proposition B.1 with the choices $Y = X_1 - X_2$, $f = \eta_1 - \eta_2$, $g = \xi_1 - \xi_2$, $T = B_1 - B_2$ and $\alpha = 1/2$ in order to infer that for every $t \in [0,T]$

$$\begin{aligned}
&\frac{1}{2}\|X_1(t) - X_2(t)\|_H^2 + \int_0^t \int_D (\eta_1(s) - \eta_2(s)) \cdot (\nabla X_1(s) - \nabla X_2(s))\, ds \\
&\qquad + \int_0^t \int_D (\xi_1(s) - \xi_2(s))\,(X_1(s) - X_2(s))\, ds \\
&= \frac{\|X_0^1 - X_0^2\|_H^2}{2} + \int_0^t \frac{\|(B_1 - B_2)(s)\|_{\mathscr{L}_2(U,H)}^2}{2}\, ds + \int_0^t ((X_1 - X_2)(s), (B_1 - B_2)(s)\, dW_s)\,.
\end{aligned}$$

Hence, taking into account (2.30)–(2.31) and the monotonicity of γ and β, we obtain

$$\begin{aligned}
\|X_1(t) - X_2(t)\|_H^2 &\leq \left\|X_0^1 - X_0^2\right\|_H^2 + \int_0^t \|B_1(s) - B_2(s)\|_{\mathscr{L}_2(U,H)}^2\, ds \\
&\qquad + 2 \sup_{t\in[0,T]} \left|\int_0^t ((X_1 - X_2)(s), (B_1 - B_2)(s)\, dW_s)\right|\,;
\end{aligned}$$

moreover, proceeding exactly as in Section 3.3, taking the supremum in $t \in [0,T]$ in the last expression and then expectations, thanks to the Davis inequality and the Young inequality, we easily obtain

$$\begin{aligned}
\|X_1 - X_2\|_{L^2(\Omega;L^\infty(0,T;H))}^2 &\leq \left\|X_0^1 - X_0^2\right\|_{L^2(\Omega;H)}^2 + c\,\|B_1 - B_2\|_{L^2(\Omega\times(0,T);\mathscr{L}_2(U,H))}^2 \\
&\qquad + \frac{1}{2}\|X_1 - X_2\|_{L^2(\Omega;L^\infty(0,T;H))}^2
\end{aligned}$$

for a positive constant c, from which (2.37) follows. Finally, if $X_0^1 = X_0^2$ and $B_1 = B_2$, we immediately get $X_1 = X_2$: substituting in the difference of the respective equations (2.29) we have $\int_0^t (-\operatorname{div}(\eta_1(s) - \eta_2(s)) + (\xi_1(s) - \xi_2(s)))\, ds = 0$ for every t. Relying now on hypothesis (2.25) and proceeding as in Section 3.6, we easily get also $\eta_1 = \eta_2$ and $\xi_1 = \xi_2$.

Let us prove now the second part of Theorem 2.5, in which condition (2.25) is not assumed. By the second part of Theorem 2.2, we have that, for every $t \in [0,T]$,

$$\begin{aligned}
X_1(t) - X_2(t) &- \int_0^t \operatorname{div}\left[\eta_1(s) - \eta_2(s)\right]\, ds + \int_0^t (\xi_1(s) - \xi_2(s))\, ds \\
&= X_0^1 - X_0^2 + \int_0^t (B_1(s) - B_2(s))\, dW_s \qquad \mathbb{P}\text{-almost surely :}
\end{aligned}$$

hence, using the second part of Proposition B.1, we infer that, for every $t \in [0,T]$,

$$\begin{aligned}
&\frac{1}{2}\|X_1(t) - X_2(t)\|_{L^2(\Omega;H)}^2 + \int_0^t \int_{\Omega\times D} (\eta_1(s) - \eta_2(s)) \cdot (\nabla X_1(s) - \nabla X_2(s))\, ds \\
&\qquad + \int_0^t \int_{\Omega\times D} (\xi_1(s) - \xi_2(s))\,(X_1(s) - X_2(s))\, ds \\
&\qquad = \frac{1}{2}\left\|X_0^1 - X_0^2\right\|_{L^2(\Omega;H)}^2 + \frac{1}{2}\int_0^t \|(B_1 - B_2)(s)\|_{L^2(\Omega;\mathscr{L}_2(U,H))}^2\, ds\,,
\end{aligned}$$

which together with the monotonicity of γ and β implies (2.38). Finally, if $X_0^1 = X_0^2$ and $B_1 = B_2$, we have $X_1 = X_2$ and $\int_0^t (-\operatorname{div}(\eta_1(s) - \eta_2(s)) + (\xi_1(s) - \xi_2(s)))\, ds = 0$ for every t, as before, so that $-\operatorname{div}\eta_1 + \xi_1 = -\operatorname{div}\eta_2 + \xi_2$.

5 Well-posedness with multiplicative noise

In this section, we prove the main theorem of the work, which ensures that the original problem is well-posed also with multiplicative noise. Let us describe the approach that we will follow.

The main idea is to prove existence of solutions proceeding step-by-step: we introduce a parameter $\tau > 0$, we prove using contraction estimates that we are able to recover some solutions on each subinterval $[0,\tau], [\tau, 2\tau], \dots [n\tau, (n+1)\tau], \dots$ provided that τ is chosen sufficiently small, and finally we paste together each solution on the whole interval $[0,T]$. In this sense, the main point of the argument is to prove that such a value of τ can be chosen uniformly with respect to n, so that the procedure stops when we reach the final time T (in a finite number of steps).

5.1 Existence

In this section we prove the two existence results contained in Theorem 2.7. We start from the first one, i.e. assuming (2.25). First of all, for every $a, b \in [0,T]$ with $b > a$ and for any progressively measurable process $Y \in L^2(\Omega \times (0,T) \times D)$, condition (2.42) implies that $B(\cdot,\cdot,Y) \in L^2(\Omega \times (a,b); \mathscr{L}_2(U,H))$: hence, for every $X_a \in L^2(\Omega, \mathcal{F}_a, \mathbb{P}; H)$, thanks to Theorem 2.2 we know that there exist

$$X_{a,b} \in L^2\left(\Omega; L^\infty(a,b;H)\right) \cap L^p\left(\Omega \times (a,b); V\right)\,, \tag{5.1}$$

$$\eta_{a,b} \in L^q\left(\Omega \times (a,b) \times D\right)^d\,, \qquad \xi_{a,b} \in L^1\left(\Omega \times (a,b) \times D\right)\,, \tag{5.2}$$

such that $X_{a,b}$ is adapted with $\mathbb{P}$-almost surely weakly continuous paths in H and the following relations hold:

$$X_{a,b}(t) - \int_a^t \operatorname{div}\eta_{a,b}(s)\, ds + \int_a^t \xi_{a,b}(s)\, ds = X_a + \int_a^t B(s, Y(s))\, dW_s \quad \text{in } V_0^*\,, \qquad \text{for every } t \in [a,b]\,, \quad \mathbb{P}\text{-almost surely}\,, \tag{5.3}$$

$$\eta_{a,b} \in \gamma(\nabla X_{a,b}) \quad \text{a.e. in } \Omega \times (a,b) \times D\,, \tag{5.4}$$

$$\xi_{a,b} \in \beta(X_{a,b}) \quad \text{a.e. in } \Omega \times (a,b) \times D\,, \tag{5.5}$$

$$j(X_{a,b}) + j^*(\xi_{a,b}) \in L^1\left(\Omega \times (a,b) \times D\right)\,, \tag{5.6}$$

where $X_{a,b}$ is unique in the sense of Theorem 2.5. Now, we need the following lemma.

Lemma 5.1. *For every $\tau > 0$ and $n \in \mathbb{N}$ fixed, consider $X_{n\tau} \in L^2(\Omega, \mathcal{F}_{n\tau}, \mathbb{P}; H)$ and $Y_1, Y_2 \in L^2(\Omega \times (n\tau, (n+1)\tau) \times D)$ progressively measurable: then, if (X_1, η_1, ξ_1) and (X_2, η_2, ξ_2) are any respective solutions to (5.1)–(5.6) with $a = n\tau$, $b = (n+1)\tau$ and same initial value $X_a = X_{n\tau}$, we have the following estimate:*

$$\|X_1 - X_2\|_{L^2(\Omega \times (n\tau,(n+1)\tau) \times D)} \leq \sqrt{\tau L_B}\, \|Y_1 - Y_2\|_{L^2(\Omega \times (n\tau,(n+1)\tau) \times D)}\,. \tag{5.7}$$

Proof. Taking the difference of equations (5.3) evaluated with $i = 1, 2$ and recalling the generalized Itô formula (B.29), setting $X := X_1 - X_2$, $\eta := \eta_1 - \eta_2$ and $\xi := \xi_1 - \xi_2$, we easily get that for every $t \in [m\tau, (m+1)\tau]$

$$\frac{1}{2}\|X(t)\|^2_{L^2(\Omega\times D)} + \int_0^t\int_{\Omega\times D}\eta(s)\cdot\nabla X(s)\,ds + \int_0^t\int_{\Omega\times D}\xi(s)X(s)\,ds = \frac{1}{2}\|B(Y_1) - B(Y_2)\|^2_{L^2(\Omega\times(m\tau,(m+1)\tau);\mathscr{L}_2(U,H))}.$$

Hence, using the Lipschitz continuity of B and the monotonicity of γ and β we have

$$\frac{1}{2}\|X_1 - X_2\|^2_{L^\infty(m\tau,(m+1)\tau;L^2(\Omega\times D))} \le \frac{L_B}{2}\|Y_1 - Y_2\|^2_{L^2(\Omega\times(m\tau,(m+1)\tau)\times D)},$$

from which (5.7) follows. □

Now, let us build some solutions X, η and ξ in each sub-interval. To this purpose, we choose $\tau > 0$ such that the constant appearing in (5.7) is less than 1, for example

$$\tau := \frac{1}{2L_B}. \tag{5.8}$$

Firstly, we focus on $[0,\tau]$: taking into account the remarks that we have just made, it is well defined the function

$$\Phi_0 : L^2(\Omega\times(0,\tau)\times D) \to L^2(\Omega\times(0,\tau)\times D), \quad \Phi_0(Y) := X, \tag{5.9}$$

where X is the unique solution to (5.1)–(5.6) with the choices $a = 0$ and $b = \tau$, with X_0 given by (2.39). It is clear that X is a solution of problem (2.43) in $[0,\tau]$ if and only if it is a fixed point for Φ_0. Thanks to the estimate (5.7) and the choice (5.8), Φ_0 is a contraction: hence, it has a fixed point

$$X^{(0)} \in L^2(\Omega; L^\infty(0,\tau;H)) \cap L^p(\Omega\times(0,\tau);V),$$

with $\mathbb{P}$-almost surely weakly continuous paths in H, which solves (2.43) with certain

$$\eta^{(0)} \in L^q(\Omega\times(0,\tau)\times D)^d, \qquad \xi^{(0)} \in L^1(\Omega\times(0,\tau)\times D).$$

Secondly, let us focus on $[\tau, 2\tau]$, set $X_\tau := X^{(0)}(\tau)$ (which is in $L^2(\Omega,\mathcal{F}_\tau,\mathbb{P};H)$ since $X^{(0)}$ is adapted) and define the function

$$\Phi_1 : L^2(\Omega\times(\tau,2\tau)\times D) \to L^2(\Omega\times(\tau,2\tau)\times D), \quad \Phi_1(Y) := X, \tag{5.10}$$

where X is the solution to (5.1)–(5.6) with the choices $a = \tau$ and $b = 2\tau$. As we have already done, thanks to the estimate (5.7) and the choice (5.8), Φ_1 is a contraction: hence, it has a fixed point

$$X^{(1)} \in L^2(\Omega; L^\infty(\tau,2\tau;H)) \cap L^p(\Omega\times(\tau,2\tau);V),$$

with $\mathbb{P}$-almost surely weakly continuous paths in H, which is a solution of (2.43) with certain

$$\eta^{(1)} \in L^q(\Omega\times(\tau,2\tau)\times D)^d, \qquad \xi^{(1)} \in L^1(\Omega\times(\tau,2\tau)\times D).$$

Suppose by induction that we have built $(X^{(0)}, \eta^{(0)}, \xi^{(0)}), \ldots, (X^{(m-1)}, \eta^{(m-1)}, \xi^{(m-1)})$ and let us show how to obtain $(X^{(m)}, \eta^{(m)}, \xi^{(m)})$. We focus on the interval $[m\tau, (m+1)\tau]$, set $X_{m\tau} := X^{(m-1)}(m\tau)$ (which is in $L^2(\Omega, \mathcal{F}_{m\tau}, \mathbb{P}; H)$ since $X^{(m-1)}$ is adapted) and define the function

$$\Phi_m : L^2\left(\Omega \times (m\tau, (m+1)\tau) \times D\right) \to L^2\left(\Omega \times (m\tau, (m+1)\tau) \times D\right) , \tag{5.11}$$

which maps Y into X, where X is the solution to (5.1)–(5.6) with the choices $a = m\tau$ and $b = (m+1)\tau$. Now, Φ_m is a contraction thanks to (5.7) and (5.8), so it has a fixed point

$$X^{(m)} \in L^2\left(\Omega; L^\infty(m\tau, (m+1)\tau; H)\right) \cap L^p\left(\Omega \times (m\tau, (m+1)\tau); V\right) :$$

with $\mathbb{P}$-almost surely weakly continuous paths in H, which is a solution of (2.43) with certain

$$\eta^{(m)} \in L^q\left(\Omega \times (m\tau, (m+1)\tau) \times D\right)^d , \qquad \xi^{(m)} \in L^1\left(\Omega \times (m\tau, (m+1)\tau) \times D\right) .$$

In this way, we can define the triplet (X, η, ξ) by setting $(X, \eta, \xi) := (X^{(m)}, \eta^{(m)}, \xi^{(m)})$ in $\Omega \times [m\tau, (m+1)\tau) \times D$ for every $m \in \mathbb{N}$ until we reach T: bearing in mind how we have built $(X^{(m)}, \eta^{(m)}, \xi^{(m)})$, it is clear that X, η and ξ are well-defined and satisfy conditions (2.26)–(2.28), (2.30)–(2.32) and (2.43).

Finally, if we do not assume (2.25), it is clear that, using the same argument, the respective solutions constructed in this way are well-defined and satisfy conditions (2.33) and (2.44) instead of (2.26) and (2.43), respectively.

5.2 Continuous dependence on the initial datum

We present here the proof of the proof of the continuous dependence results contained in the last part of Theorem 2.7. Here, we repeat exactly the same argument of Section 4 with the choices $B_1 := B(\cdot, X_1)$ and $B_2 := (\cdot, X_2)$.

If (2.25) is assumed, for any given $\tau > 0$, the same computations on the interval $(0, \tau)$ get us to

$$\|X_1 - X_2\|^2_{L^2(\Omega; L^\infty(0,\tau;H))} \leq c \left\|X_0^1 - X_0^2\right\|^2_{L^2(\Omega;H)} + c \left\|B(X_1) - B(X_2)\right\|^2_{L^2(\Omega\times(0,\tau);\mathscr{L}_2(U,H))}$$

for a constant $c > 0$ independent of τ; using the Lipschitz continuity of B we obtain

$$\|X_1 - X_2\|^2_{L^2(\Omega; L^\infty(0,\tau;H))} \leq c \left\|X_0^1 - X_0^2\right\|^2_{L^2(\Omega;H)} + c\tau \left\|X_1 - X_2\right\|^2_{L^2(\Omega;L^\infty(0,\tau;H))} .$$

Hence, choosing for example $\tau = \frac{c}{2}$, we get the desired relation on the interval $[0, \tau]$. The idea is clearly to iterate the procedure on the following intervals $[\tau, 2\tau]$, $[2\tau, 3\tau]$, ... until we reach the final time T, so that (2.45) is proved. The important point that we have to check is that the choice of τ can be made uniformly with respect to each sub-interval, but this is not difficult: as a matter of fact, for any $n \geq 1$, performing the same computations on $[n\tau, (n+1)\tau]$ we obtain

$$\begin{aligned}\|X_1 - X_2\|^2_{L^2(\Omega; L^\infty(n\tau,(n+1)\tau;H))} &\leq c \left\|X_1(n\tau) - X_2(n\tau)\right\|^2_{L^2(\Omega;H)} \\ &\quad + c\tau \left\|X_1 - X_2\right\|^2_{L^2(\Omega;L^\infty(n\tau,(n+1)\tau;H))} ,\end{aligned}$$

for the same constant c, from which we deduce that the choice of τ is independent of n, and one can easily conclude by induction on n. As we did in Section 4, if $X_0^1 = X_0^1$, then by (2.45) we have $X_1 = X_2$, and hypothesis (2.25) also ensures $\eta_1 = \eta_2$ and $\xi_1 = \xi_2$.

Secondly, if (2.25) is not assumed, proceeding as in the final part of Section 4 we get for every $t \in [0,T]$ that

$$\|X_1(t) - X_2(t)\|^2_{L^2(\Omega;H)} \leq \left\|X_0^1 - X_0^2\right\|^2_{L^2(\Omega;H)} + \int_0^t \|(B(X_1) - B(X_2))(s)\|^2_{L^2(\Omega;\mathscr{L}_2(U,H))}\, ds\,,$$

from which (2.46) follows using the Lipschitz continuity of B and the Gronwall lemma. Finally, if $X_0^1 = X_0^1$, then by (2.46) $X_1 = X_2$ and consequently $-\operatorname{div}\eta_1 + \xi_1 = -\operatorname{div}\eta_2 + \xi_2$.

A An integration-by-parts formula

The aim of this Appendix is to give a complete proof of the generalized testing formula contained in equation (3.54): throughout the section, we assume to work with the notations and setting of Section 3.5. Here, $\varepsilon \in (0,1)$ and $\omega \in \Omega$ are fixed as usual.

The main point is that we cannot directly test equation (3.48) by $X^\varepsilon - W_B^\varepsilon$, as we did in Section 3.2, since the regularity of X^ε is not enough: more specifically, $\partial_t(X^\varepsilon - W_B^\varepsilon)$ is only intended in V_0^* and we would need that $X^\varepsilon - W_B^\varepsilon$ takes values in V_0, but this is not the case. However, by condition (3.53) and the regularities of X^ε, W_B^ε and η^ε, all the terms in (3.54) make sense: hence, the intuitive idea is that (3.54) holds at least in a formal way. To give a rigorous proof of it, a natural way could be to try to pass to the limit as $\lambda \searrow 0$ in (3.15): however, it is not necessarily true in our framework that equation (3.15) converges to (3.54) as $\lambda \searrow 0$, so this approach does not work. Hence, the idea is to see (3.54) as a limit problem as $\delta \searrow 0$, for another parameter δ, such that the approximations in δ have good smoothing properties and behave better that the approximations in λ. In this sense, a similar approach was presented in [5] and [21], where the approximations were built using suitable powers of the resolvent of the laplacian. However, in our case we have to approximate also elements in $W^{-1,q}(D)$ (namely, $-\operatorname{div}\eta^\varepsilon$) and the resolvent of the laplacian does not work since $-\Delta$ is not coercive on V: the idea is thus to identify another suitable space, in which (3.48) can be intended, and to define appropriate approximations on it. To this purpose, we need some preparatory work.

First of all, note that the operator $-\operatorname{div} : L^q(D)^d \to V^*$ is linear, continuous and satisfies $\|-\operatorname{div} u\|_{V^*} \leq \|u\|_{L^q(D)}$ for every $u \in L^q(D)^d$. Let us define the space

$$V^*_{div} := \left\{-\operatorname{div} u : u \in L^q(D)^d\right\} \subseteq V^*\,.$$

Secondly, we introduce the space $V^*_{div} \oplus L^1(D)$ as the subspace of V_0^* given by all the formal linear combinations of elements in V^*_{div} and $L^1(D)$. With this notations, we can note that equation (3.48) actually holds in $V^*_{div} \oplus L^1(D)$: in other words, for every $t \in [0,T]$, we have

$$\left(X^\varepsilon - W_B^\varepsilon\right)(t) + \int_0^t \left(-\operatorname{div}\eta^\varepsilon(s) + \xi^\varepsilon(s)\right)\, ds = X_0 \quad \text{in } V^*_{div} \oplus L^1(D)\,. \tag{A.12}$$

Hence, the idea is that it is sufficient to identify a way to approximate only elements in $V_{div}^* \oplus L^1(D)$, and not any element of V_0^*, which would be much more demanding.

To this end, for every $\delta \in (0,1)$, let $\mathcal{R}_\delta := (I - \delta\Delta)^{-1}$ be the resolvent of the Laplace operator. It is well-known that for every $r \in [1, +\infty)$, $\mathcal{R}_\delta : L^r(D) \to L^r(D)$ is a linear contraction converging to the identity as $\delta \searrow 0$ in the strong operator topology (the reader can refer to [4,6,12]). In this setting, we define the operator $\mathbf{R}_\delta : L^r(D)^d \to L^r(D)^d$ extending $\mathcal{R}_\delta$ component-by-component: consequently, we easily deduce that also $\mathbf{R}_\delta$ is a linear contraction on $L^r(D)^d$ converging to the identity as $\delta \searrow 0$. With this notations, we have the following result.

Lemma A.1. *For every $u \in L^q(D)^d$ such that $-\operatorname{div} u \in L^1(D)$ (in the distributional sense), we have*

$$-\operatorname{div} \mathbf{R}_\delta u = \mathcal{R}_\delta \left(-\operatorname{div} u\right) .$$

Moreover, for every $f \in H^1(D)$, we have

$$\nabla \mathcal{R}_\delta f = \mathbf{R}_\delta \nabla f \, .$$

Proof. Let us first assume that $u \in (C_c^\infty(D))^d$: then, using the definition of $\mathbf{R}_\delta$ and $\mathcal{R}_\delta$, integration by parts and the fact that $\mathcal{R}_\delta$ commutes with Δ, for every $\varphi \in C_c^\infty(D)$ we have

$$\begin{aligned}\int_D (-\operatorname{div} u)\, \varphi = \int_D u \cdot \nabla \varphi = \sum_{i=1}^d \int_D u_i \frac{\partial \varphi}{\partial x_i} = \sum_{i=1}^d \int_D \left(\mathcal{R}_\delta u_i - \delta \Delta \mathcal{R}_\delta u_i\right) \frac{\partial \varphi}{\partial x_i} \\ = \int_D \mathbf{R}_\delta u \cdot \nabla \varphi + \delta \int_D \Delta(\operatorname{div} \mathbf{R}_\delta u) \varphi = \int_D \left[-\operatorname{div} \mathbf{R}_\delta u - \delta \Delta(-\operatorname{div} \mathbf{R}_\delta u)\right] \varphi \, .\end{aligned}$$

Hence, by definition of the resolvent, we deduce that $-\operatorname{div} \mathbf{R}_\delta u = \mathcal{R}_\delta(-\operatorname{div} u)$ for every $u \in (C_c^\infty(D))^d$. At this point, if $u \in L^q(D)^d$ and $-\operatorname{div} u \in L^1(D)$, the first thesis follows by approximating u with a sequence $\{u_n\}_{n\in\mathbb{N}} \subseteq (C^\infty(D))^d$ such that $u_n \to u$ in $L^q(D)^d$ and $-\operatorname{div} u_n \to -\operatorname{div} u$ in $L^1(D)$. Finally, in a similar way, the second assertion is clearly true for every $f \in C^\infty(\overline{D})$: hence, given $f \in H^1(D)$, we can conclude by density approximating f with a sequence $\{f_n\}_{n\in\mathbb{N}} \subseteq C^\infty(\overline{D})$. □

Now, for every $\delta \in (0,1)$, we introduce the operator

$$\Lambda_\delta^1 : V_{div}^* \to V_{div}^*$$

in the following way: for any given $f \in V_{div}^*$, with $f = -\operatorname{div} u$ for a certain $u \in L^q(D)^d$, we set $\Lambda_\delta^1 f := -\operatorname{div} \mathbf{R}_\delta u$. Note that Λ_δ^1 is well-defined: indeed, if $f = -\operatorname{div} u_1 = -\operatorname{div} u_2$, we have $-\operatorname{div}(u_1 - u_2) = 0$ and by Lemma A.1 we deduce that $0 = \mathcal{R}_\delta(-\operatorname{div}(u_1 - u_2)) = -\operatorname{div}(\mathbf{R}_\delta(u_1 - u_2))$, so that $-\operatorname{div} \mathbf{R}_\delta u_1 = -\operatorname{div} \mathbf{R}_\delta u_2$. Secondly, we set

$$\Lambda_\delta^2 : L^1(D) \to L^1(D) \, , \quad \Lambda_\delta^2 := \mathcal{R}_\delta \, .$$

The first part of Lemma A.1 ensures that $\Lambda_\delta^1 = \Lambda_\delta^2$ on the intersection $V_{div}^* \cap L^1(D)$: hence, it is well-defined the operator

$$\Lambda_\delta := \Lambda_\delta^1 \oplus \Lambda_\delta^2 : V_{div}^* \oplus L^1(D) \to V_{div}^* \oplus L^1(D) \tag{A.13}$$

such that

$$\Lambda_\delta(-\operatorname{div} u) = -\operatorname{div} \mathbf{R}_\delta u\,, \quad \Lambda_\delta(f) = \mathcal{R}_\delta f \quad \text{for every } u \in L^q(D)^d\,,\ f \in L^1(D)\,, \tag{A.14}$$

which is automatically linear.

We are now ready to build the approximations. First of all, we choose $k \in \mathbb{N}$ as in (2.21), so that the k-th power $\mathcal{R}_\delta^k$ maps H into $V_0 \subseteq V \cap L^\infty(D)$. At this point, we define

$$X_\delta^\varepsilon := \mathcal{R}_\delta^k X^\varepsilon\,, \quad W_\delta^\varepsilon := \mathcal{R}_\delta^k W_B^\varepsilon\,, \quad \eta_\delta^\varepsilon := \mathbf{R}_\delta^k \eta^\varepsilon\,, \quad \xi_\delta^\varepsilon := \mathcal{R}_\delta^k \xi^\varepsilon\,, \quad X_0^\delta := \mathcal{R}_\delta^k X_0 \,: \tag{A.15}$$

then, taking into account the properties of $\mathcal{R}_\delta$ and $\mathbf{R}_\delta$ and the second part of Lemma A.1, thanks to conditions (3.35)–(3.36), (3.43) and (3.52) we have as $\delta \searrow 0$ that

$$X_\delta^\varepsilon(t) \to X^\varepsilon(t) \quad \text{in } H \quad \text{for every } t \in [0,T]\,, \quad X_\delta^\varepsilon \to X^\varepsilon \quad \text{in } L^p(0,T;V) \tag{A.16}$$

$$W_\delta^\varepsilon(t) \to W^\varepsilon(t) \quad \text{in } H \quad \text{for every } t \in [0,T]\,, \quad W_\delta^\varepsilon \to W_B^\varepsilon \quad \text{in } L^p(0,T;V)\,, \tag{A.17}$$

$$\eta_\delta^\varepsilon \to \eta^\varepsilon \quad \text{in } L^q(Q)^d\,, \qquad \xi_\delta^\varepsilon \to \xi^\varepsilon \quad \text{in } L^1(Q)\,, \tag{A.18}$$

$$X_0^\delta \to X_0 \quad \text{in } H\,. \tag{A.19}$$

Now, applying the operator Λ_δ^k to equation (A.12), we get for every $t \in [0,T]$ that

$$\left(X_\delta^\varepsilon - W_\delta^\varepsilon\right)(t) - \int_0^t \operatorname{div} \eta_\delta^\varepsilon(s)\,ds + \int_0^t \xi_\delta^\varepsilon(s)\,ds = X_0^\delta\,. \tag{A.20}$$

With our choice of k, it now makes sense to test by $X_\delta^\varepsilon - W_\delta^\varepsilon$: it easily follows that

$$\frac{1}{2}\left\|X_\delta^\varepsilon(T) - W_\delta^\varepsilon(T)\right\|_H^2 + \int_Q \nabla\eta_\delta^\varepsilon \cdot \nabla\left(X_\delta^\varepsilon - W_\delta^\varepsilon\right) + \int_Q \xi_\delta^\varepsilon\left(X_\delta^\varepsilon - W_\delta^\varepsilon\right) = \frac{1}{2}\left\|X_0^\delta\right\|_H^2\,, \tag{A.21}$$

from which, taking into account (A.16)–(A.19), we deduce that

$$\lim_{\delta\searrow 0}\int_Q \xi_\delta^\varepsilon\left(X_\delta^\varepsilon - W_\delta^\varepsilon\right) = \frac{1}{2}\left\|X_0\right\|_H^2 - \frac{1}{2}\left\|(X^\varepsilon - W_B^\varepsilon)(T)\right\|_H^2 - \int_Q \nabla\eta^\varepsilon \cdot \nabla\left(X^\varepsilon - W_B^\varepsilon\right). \tag{A.22}$$

In order to evaluate the limit in the previous expression, we take advantage of Vitali convergence theorem: to this purpose, thanks to (A.16)–(A.18), we can assume with no restriction that $\xi_\delta^\varepsilon \to \xi^\varepsilon$ and $X_\delta^\varepsilon - W_\delta^\varepsilon \to X^\varepsilon - W_B^\varepsilon$ almost everywhere in Q. Let us show that $\{\xi_\delta^\varepsilon(X_\delta^\varepsilon - W_\delta^\varepsilon)\}_{\delta\in(0,1)}$ is uniformly integrable in Q: by conditions (2.10)–(2.11) and thanks to the generalized Jensen inequality for the positive operator R_δ (see [14, 15] for references), we have

$$\begin{aligned}
\pm\xi_\delta^\varepsilon(X_\delta^\varepsilon - W_\delta^\varepsilon) &\le j\left(\pm(X_\delta^\varepsilon - W_\delta^\varepsilon)\right) + j^*(\xi_\delta^\varepsilon) = j\left(X_\delta^\varepsilon - W_\delta^\varepsilon\right) + j^*(\xi_\delta^\varepsilon) \\
&\le \mathcal{R}_\delta^k\left[j\left(X^\varepsilon - W_B^\varepsilon\right) + j^*(\xi^\varepsilon)\right] \quad \text{a.e. in } Q\,.
\end{aligned}$$

Now, since $j(X^\varepsilon - W_B^\varepsilon), j^*(\xi^\varepsilon) \in L^1(Q)$ thanks to (3.51) and (3.14), the right hand side of the previous expression converges in $L^1(Q)$ and consequently it is uniformly integrable in Q: we deduce that also $\{\xi_\delta^\varepsilon(X_\delta^\varepsilon - W_\delta^\varepsilon)\}_{\delta\in(0,1)}$ is uniformly integrable in Q. Hence, by Vitali convergence theorem, we infer that

$$\xi_\delta^\varepsilon\left(X_\delta^\varepsilon - W_\delta^\varepsilon\right) \to \xi^\varepsilon\left(X^\varepsilon - W_B^\varepsilon\right) \quad \text{in } L^1(Q) \quad \text{as } \delta \searrow 0\,,$$

so that passing to the limit in (A.22) we recover exactly (3.54).

B The generalized Itô formula

In this appendix, we prove a generalized Itô formula, which is widely used in Sections 3.7 and 4: we collect the general result in the following proposition. Throughout the section, we assume the general setting (2.1)–(2.22).

Proposition B.1. *Assume the following conditions:*

$$Y_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H) \,, \tag{B.23}$$

$$T \in L^2(\Omega \times (0,T); \mathscr{L}_2(U,H)) \quad \textit{progressively measurable}\,, \tag{B.24}$$

$$Y \in L^2(\Omega; L^\infty(0,T;H)) \cap L^p(\Omega \times (0,T); V) \,, \qquad Y \in C^0_w([0,T];H) \quad \mathbb{P}\text{-}a.s.\,, \tag{B.25}$$

$$f \in L^q(\Omega \times (0,T) \times D)^d \,, \qquad g \in L^1(\Omega \times (0,T) \times D) \,, \tag{B.26}$$

$$\exists\, \alpha > 0 \,: \quad j(\alpha Y) + j^*(\alpha g) \in L^1(\Omega \times (0,T) \times D) \,, \tag{B.27}$$

$$Y(t) - \int_0^t \operatorname{div} f(s)\,ds + \int_0^t g(s)\,ds = Y_0 + \int_0^t T(s)\,dW_s \quad \textit{in } V_0^* \tag{B.28}$$

for every $t \in [0,T]$, $\mathbb{P}$-almost surely. Then, the following Itô formula holds

$$\begin{aligned} \frac{1}{2}\|Y(t)\|_H^2 + \int_0^t\int_D f(s)\cdot\nabla Y(s)\,ds + \int_0^t\int_D g(s)Y(s)\,ds \\ = \frac{1}{2}\|Y_0\|_H^2 + \frac{1}{2}\int_0^t \|T(s)\|^2_{\mathscr{L}_2(U,H)}\,ds + \int_0^t (Y(s), T(s)\,dW_s) \end{aligned} \tag{B.29}$$

for every $t \in [0,T]$, $\mathbb{P}$-almost surely. Furthermore, if hypothesis (B.25) *is replaced by the weaker condition*

$$Y \in L^\infty\left(0,T; L^2(\Omega;H)\right) \cap L^p(\Omega \times (0,T); V) \cap C^0_w\left([0,T]; L^2(\Omega;H)\right) \,, \tag{B.30}$$

then instead of (B.29) *we have the following for every $t \in [0,T]$:*

$$\begin{aligned} \frac{1}{2}\|Y(t)\|^2_{L^2(\Omega;H)} + \int_0^t\int_{\Omega\times D} f(s)\cdot\nabla Y(s)\,ds + \int_0^t\int_{\Omega\times D} g(s)Y(s)\,ds \\ = \frac{1}{2}\|Y_0\|^2_{L^2(\Omega;H)} + \frac{1}{2}\int_0^t \|T(s)\|^2_{L^2(\Omega;\mathscr{L}_2(U,H))}\,ds\,. \end{aligned} \tag{B.31}$$

Proof. We proceed exactly in the same way as in Appendix A. If k is given by (2.21) and for every $\delta \in (0,1)$, $\mathcal{R}_\delta$ and $\mathbf{R}_\delta$ are as in Appendix A, we define

$$Y_\delta := \mathcal{R}_\delta^k Y\,, \quad T_\delta := \mathcal{R}_\delta^k T\,, \quad f_\delta := \mathbf{R}_\delta^k f\,, \quad g_\delta := \mathcal{R}_\delta^k g\,, \quad Y_0^\delta := \mathcal{R}_\delta^k Y_0 \;:$$

hence, thanks to (B.23)–(B.26) and Lemma A.1 we have as $\delta \searrow 0$

$$Y_\delta(t) \to Y(t) \quad \text{in } H \quad \text{for every } t \in [0,T]\,, \quad \mathbb{P}\text{-almost surely}\,, \tag{B.32}$$

$$Y_\delta \to Y \quad \text{in } L^p(\Omega \times (0,T); V) \tag{B.33}$$

$$T_\delta \to T \quad \text{in } L^2(\Omega \times (0,T); \mathscr{L}_2(U,H)) \,, \tag{B.34}$$

$$f_\delta \to f \quad \text{in } L^q(\Omega \times Q)^d \,, \qquad g_\delta \to g \quad \text{in } L^1(\Omega \times Q)\,, \tag{B.35}$$

$$Y_0^\delta \to Y_0 \quad \text{in } L^2(\Omega; H)\,. \tag{B.36}$$

Consequently, if we apply the operator Λ_δ^k to (B.28), taking definition (A.13)–(A.14) into account, we have $\mathbb{P}$-almost surely that

$$Y_\delta(t) - \int_0^t \operatorname{div} f_\delta(s)\,ds + \int_0^t g_\delta(s)\,ds = Y_0^\delta + \int_0^t T_\delta(s)\,dW_s \quad \text{in } H\,, \quad \text{for every } t \in [0,T]\,.$$

Now, with our choice of k, we can apply the classical Itô formula (see [24, Thm. 4.2.5] for example) to recover that $\mathbb{P}$-almost surely, for every $t \in [0,T]$,

$$\begin{aligned}\frac{1}{2}\|Y_\delta(t)\|_H^2 &+ \int_0^t\int_D f_\delta(s)\cdot\nabla Y_\delta(s)\,ds + \int_0^t\int_D g_\delta(s)Y_\delta(s)\,ds \\ &= \frac{1}{2}\left\|Y_0^\delta\right\|_H^2 + \frac{1}{2}\int_0^t \|T_\delta(s)\|_{L_2(U,H)}^2\,ds + \int_0^t (Y_\delta(s), T_\delta(s)\,dW_s)\,.\end{aligned} \tag{B.37}$$

Now, let us focus on the stochastic integral: we have

$$\begin{aligned}&\int_0^t (Y_\delta(s), T_\delta(s)\,dW_s) - \int_0^t (Y(s), T(s)\,dW_s) \\ &\qquad = \int_0^t (Y_\delta(s), (T_\delta - T)(s)\,dW_s) + \int_0^t ((Y_\delta - Y)(s), T(s)\,dW_s)\ ,\end{aligned}$$

where thanks to the Davis inequality and (B.33)–(B.34) we have (renominating the positive constant c)

$$\begin{aligned}\mathbb{E}\sup_{t\in[0,T]}\left|\int_0^t (Y_\delta(s), (T_\delta - T)(s)\,dW_s)\right| &\le c\mathbb{E}\left[\left(\int_0^T \|Y_\delta(s)\|_H^2\,\|(T_\delta - T)(s)\|_{\mathscr{L}_2(U,H)}^2\,ds\right)^{1/2}\right] \\ &\le c\,\|T_\delta - T\|_{L^2(\Omega\times(0,T);\mathscr{L}_2(U,H))} \to 0\end{aligned}$$

and, by the dominated convergence theorem, also

$$\mathbb{E}\sup_{t\in[0,T]}\left|\int_0^t ((Y_\delta - Y)(s), T_s\,dW_s)\right|^2 \le c\mathbb{E}\left[\left(\int_0^T \|(Y_\delta - Y)(s)\|_H^2\,\|T_s\|_{\mathscr{L}_2(U,H)}^2\,ds\right)^{1/2}\right] \to 0\,.$$

Hence, we deduce that $\int_0^\cdot (Y_\delta(s), T_\delta(s)\,dW_s) \to \int_0^\cdot (Y(s), T(s)\,dW_s)$ in $L^2(\Omega; L^\infty(0,T))$, so that consequently (at least for a subsequence)

$$\int_0^t (Y_\delta(s), T_\delta(s)\,dW_s) \to \int_0^t (Y(s), T(s)\,dW_s) \quad \text{for every } t \in [0,T]\,, \quad \mathbb{P}\text{-almost surely}\,.$$

Hence, letting $\delta \searrow 0$ and taking into account (B.32)–(B.36), $\mathbb{P}$-almost surely we have

$$\begin{aligned}\lim_{\delta\searrow 0}\int_{(0,t)\times D} g_\delta Y_\delta &= \frac{1}{2}\|Y_0\|_H^2 + \frac{1}{2}\int_0^t \|T(s)\|_{\mathscr{L}_2(U,H)}^2\,ds + \int_0^t (Y(s), T(s)\,dW_s) \\ &\quad - \frac{1}{2}\|Y(t)\|_H^2 - \int_{(0,t)\times D} f_\delta\cdot\nabla Y \quad \text{for every } t \in [0,T]\ :\end{aligned} \tag{B.38}$$

we evaluate the limit on the left hand side using Vitali's convergence theorem. To this purpose, by (B.32) and (B.35) we can assume with no restriction that $Y_\delta \to Y$ and $g_\delta \to g$

almost everywhere in $\Omega \times (0,t) \times D$; moreover, thanks to conditions (2.10)–(2.11) and the generalized Jensen inequality for positive operators (see [14, 15]), we have

$$\pm\alpha^2 g_\delta Y_\delta \leq j\left(\pm\alpha Y_\delta\right) + j^*\left(\alpha g_\delta\right) = j\left(\alpha Y_\delta\right) + j^*\left(\alpha g_\delta\right) \leq \mathcal{R}_\delta^k\left[j(\alpha Y) + j^*(\alpha g)\right] .$$

Thanks to (B.27) and the properties of $\mathcal{R}_\delta$, the term on the right hand side converges in $L^1(\Omega \times (0,t) \times D)$, hence it is uniformly integrable: consequently, we deduce that also $\{g_\delta Y_\delta\}_{\delta\in(0,1)}$ is uniformly integrable, and Vitali's convergence theorem implies that

$$g_\delta Y_\delta \to gY \quad \text{in } L^1\left(\Omega \times (0,t) \times D\right) , \quad \text{as } \delta \searrow 0 ,$$

so that passing to the limit in (B.38) we obtain (B.29).
To show (B.31), we proceed in a very similar way: note that since (B.25) is replaced by (B.30), then instead of (B.32) we have

$$Y_\delta(t) \to Y(t) \quad \text{in } L^2(\Omega; H) , \quad \text{for every } t \in [0,T] .$$

Once we have obtained (B.37) as before, we observe that the stochastic integral in (B.37) is a local martingale, so that there exists a sequence of increasing stopping times $\{\tau_n\}_{n\in\mathbb{N}}$ such that $\tau_n \nearrow \infty$ and the corresponding stopped processes are martingales: hence, stopping (B.37) at time τ_n, taking expectations and then letting $n \to \infty$, thanks to dominated convergence theorem we directly obtain for every $t \in [0,T]$

$$\frac{1}{2}\left\|Y_\delta(t)\right\|^2_{L^2(\Omega;H)} + \int_0^t\int_{\Omega\times D} f_\delta(s)\cdot\nabla Y_\delta(s)\,ds + \int_0^t\int_{\Omega\times D} g_\delta(s)Y_\delta(s)\,ds$$
$$= \frac{1}{2}\left\|Y_0^\delta\right\|^2_{L^2(\Omega;H)} + \frac{1}{2}\int_0^t \left\|T_\delta(s)\right\|^2_{L^2(\Omega;\mathscr{L}_2(U,H))}\,ds .$$

At this point, (B.31) follows as before letting $\delta \searrow 0$ in the previous equation. □

References

[1] V. Barbu. *Nonlinear semigroups and differential equations in Banach spaces*. Editura Academiei Republicii Socialiste România, Bucharest; Noordhoff International Publishing, Leiden, 1976. Translated from the Romanian.

[2] V. Barbu. *Analysis and control of nonlinear infinite-dimensional systems*. Academic Press Inc., Boston, MA, 1993.

[3] V. Barbu. Existence for semilinear parabolic stochastic equations. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.*, 21(4):397–403, 2010.

[4] V. Barbu. *Nonlinear differential equations of monotone types in Banach spaces*. Springer Monographs in Mathematics. Springer, New York, 2010.

[5] V. Barbu, G. Da Prato, and M. Röckner. Existence of strong solutions for stochastic porous media equation under general monotonicity conditions. *Ann. Probab.*, 37(2):428–452, 2009.

[6] P. Benilan, H. Brezis, and M. G. Crandall. A semilinear equation in $L^1(R^N)$. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 2(4):523–555, 1975.

[7] L. Boccardo. Some nonlinear Dirichlet problems in L^1 involving lower order terms in divergence form. In *Progress in elliptic and parabolic partial differential equations (Capri, 1994)*, volume 350 of *Pitman Res. Notes Math. Ser.*, pages 43–57. Longman, Harlow, 1996.

[8] L. Boccardo and G. Croce. *Elliptic partial differential equations*, volume 55 of *De Gruyter Studies in Mathematics*. De Gruyter, Berlin, 2014. Existence and regularity of distributional solutions.

[9] L. Boccardo and T. Gallouët. Strongly nonlinear elliptic equations having natural growth terms and L^1 data. *Nonlinear Anal.*, 19(6):573–579, 1992.

[10] H. Brézis. Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations. In *Contributions to nonlinear functional analysis (Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1971)*, pages 101–156. Academic Press, New York, 1971.

[11] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*. North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1973. North-Holland Mathematics Studies, No. 5. Notas de Matemática (50).

[12] H. Brézis and W. A. Strauss. Semi-linear second-order elliptic equations in L^1. *J. Math. Soc. Japan*, 25:565–590, 1973.

[13] G. Da Prato. *Kolmogorov equations for stochastic PDEs*. Advanced Courses in Mathematics. CRM Barcelona. Birkhäuser Verlag, Basel, 2004.

[14] A. Guessab and G. Schmeisser. Necessary and sufficient conditions for the validity of Jensen's inequality. *Arch. Math. (Basel)*, 100(6):561–570, 2013.

[15] M. Haase. Convexity inequalities for positive operators. *Positivity*, 11(1):57–68, 2007.

[16] H. Kardestuncer and D. H. Norrie, editors. Chapters 1–3 in *Finite element handbook*. McGraw-Hill Book Co., New York, 1987.

[17] N. V. Krylov and B. L. Rozovskiĭ. Stochastic evolution equations. In *Current problems in mathematics, Vol. 14 (Russian)*, pages 71–147, 256. Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Informatsii, Moscow, 1979.

[18] W. Liu. On the stochastic p-Laplace equation. *J. Math. Anal. Appl.*, 360(2):737–751, 2009.

[19] W. Liu. Well-posedness of stochastic partial differential equations with Lyapunov condition. *J. Differential Equations*, 255(3):572–592, 2013.

[20] C. Marinelli. On well-posedness of semilinear stochastic evolution equations on L_p spaces. *ArXiv e-prints*, Dec. 2015.

[21] C. Marinelli and L. Scarpa. A variational approach to dissipative SPDEs with singular drift. *ArXiv e-prints*, Apr. 2016.

[22] É. Pardoux. Sur des équations aux dérivées partielles stochastiques monotones. *C. R. Acad. Sci. Paris Sér. A-B*, 275:A101–A103, 1972.

[23] E. Pardoux. Équations aux dérivées partielles stochastiques de type monotone. In *Séminaire sur les Équations aux Dérivées Partielles (1974–1975), III, Exp. No. 2*, page 10. Collège de France, Paris, 1975.

[24] C. Prévôt and M. Röckner. *A concise course on stochastic partial differential equations*, volume 1905 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007.

[25] J. Simon. Compact sets in the space $L^p(0,T;B)$. *Ann. Mat. Pura Appl. (4)*, 146:65–96, 1987.

[26] W. A. Strauss. On continuity of functions with values in various Banach spaces. *Pacific J. Math.*, 19:543–551, 1966.