

DOUBLY NONLINEAR STOCHASTIC EVOLUTION EQUATIONS

LUCA SCARPA AND ULISSE STEFANELLI

ABSTRACT. We present an existence theory for martingale and strong solutions to doubly nonlinear evolution equations in a separable Hilbert space in the form

$$d(Au) + Bu \, dt \ni F(u) \, dt + G(u) \, dW$$

where both A and B are maximal monotone operators, possibly multivalued, F and G are Lipschitz-continuous, and W is a cylindrical Wiener process. Via regularization and passage-to-the-limit we show the existence of martingale solutions. The identification of the limit is obtained by a lower-semicontinuity argument based on a suitably generalized Itô's formula. If either A or B is linear and symmetric, existence and uniqueness of strong solutions follows. Eventually, several applications are discussed, including doubly nonlinear stochastic Stefan-type problems.

1. INTRODUCTION

This paper is concerned with the existence of martingale solutions to nonlinear stochastic partial differential equations (SPDEs) of the form

$$d\alpha(u) - \operatorname{div}(\beta_1(\nabla u)) \, dt + \beta_0(u) \, dt \ni f(u) \, dt + G(u) \, dW. \quad (1)$$

Here, u is real-valued and the graphs $\alpha, \beta_0 : \mathbb{R} \rightarrow 2^{\mathbb{R}}$ and $\beta_1 : \mathbb{R}^d \rightarrow 2^{\mathbb{R}^d}$ are maximal monotone, possibly multivalued, coercive, and linearly bounded. The function f is assumed to be Lipschitz-continuous and G is a Lipschitz-type operator, stochastically integrable with respect to W , a cylindrical Wiener process on the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Relations of the type of (1) arise in connection with nonlinear diffusion phenomena and, under different choices of the graph α , may related to different models in nonequilibrium thermodynamics. In particular, by letting the graph α include a vertical segment through the origin, equation (1) corresponds to a nonlinear-diffusion version of the two-phase Stefan problem with stochastic perturbation. More specifically, one can choose $\alpha = \operatorname{sign} + \tilde{\alpha}$, where $\tilde{\alpha} : \mathbb{R} \rightarrow \mathbb{R}$ is a nondecreasing continuous function and $\operatorname{sign} : \mathbb{R} \rightarrow 2^{\mathbb{R}}$ is the sign-graph, defined as $\operatorname{sign}(r) = r/|r|$ for $r \neq 0$ and $\operatorname{sign}(0) = [-1, 1]$. By allowing noncoercive graphs α , an option which is however not covered by our analysis, relation (1) arises in connection with the Hele-Shaw cell and filtration through porous media as well. The reader is referred to VISINTIN [54] for a discussion on the relevance of relation (1) in the frame of phase-transition modelling.

The focus of the paper is on a variational reformulation of the initial-boundary value problem for (1) in terms of the abstract stochastic doubly nonlinear equation

$$d(Au) + Bu \, dt \ni F(u) \, dt + G(u) \, dW, \quad (Au)(0) \ni v_0. \quad (2)$$

2010 *Mathematics Subject Classification.* 35K55, 35R60, 60H15.

Key words and phrases. Maximal monotone operators, doubly nonlinear stochastic equations, strong and martingale solutions, existence, generalized Itô's formula.

Here, the process u takes values in the real, separable Hilbert space V , which embeds compactly and densely into a second Hilbert space H . The operators $A = \partial\varphi : H \rightarrow H$ and $B : V \rightarrow V^*$ are maximal monotone, possibly multivalued, coercive and linearly bounded. The operator A is assumed to be cyclic monotone, with Gâteaux-differentiable inverse A^{-1} (see below), and the map $F : H \rightarrow H$ is Lipschitz-continuous. Eventually, W is a cylindrical Wiener process on a third, separable Hilbert space U and the time-dependent operator G takes values in the space of Hilbert-Schmidt operators from U to H . The term *doubly nonlinear* refers here to the fact that both operators A and B are nonlinear.

In the deterministic case $G = 0$, problem (2) is classical. Its analysis can be traced back at least to BARDOS & BREZIS [13] and RAVIART [44]. Other early contributions are from GRANGE & MIGNOT [26], BARBU [8], DiBENEDETTO & SHOWALTER [18], ALT & LUCKHAUS [7], and BERNIS [15]. For a collection of further developments, the reader can check [1, 4–6, 21, 25, 28–30, 34, 50–52], among many others.

In the stochastic case $G \neq 0$, problem (2) is well-studied for A linear and not degenerate. The reader is referred to the seminal contributions by PARDOUX [41, 42] and KRYLOV & ROZOVSKIĬ [33], and to the monographs [17, 43], for a general overview. In the context of variational approach, problem (2) with A linear has been proved by GESS [22] to admit strong solutions for B cyclic monotone and subhomogeneous. Well-posedness from a variational approach have been obtained also under no growth condition on the drift in [35, 37, 38] for semilinear equations, in [36, 39, 46] for equations in divergence form, and in [11, 40, 47] for porous-media, Allen-Cahn, and Cahn-Hilliard equations.

Alternatively to the variational approach, the analysis of stochastic evolution equations in the form (2) (in the case A linear) have been developed in several directions. First of all, in the classical work by BENSOUSSAN & RASCANU [14], existence of strong and martingale solutions in terms of stochastic variational inequalities is proved. Stochastic variational inequalities have then been used to formulate weaker concepts of solutions also for other types of equations as divergence-form equations [24] or fast-diffusion equations [23]. More recently, an operatorial approach to monotone equations with maximal monotone drift and linear multiplicative noise has been given by BARBU & RÖCKNER in [12]: here solutions are defined using a suitable rescaling argument and monotonicity techniques in spaces of stochastic processes.

The case of A nonlinear and B linear has been originally treated by BARBU & DA PRATO [10] (see also [9, 32]) in the framework of the two-phase stochastic Stefan problem. There, the authors study equation (1) with the choices $\beta_1 = I$, $\beta_0 = f = 0$, and $\alpha = \text{sign} + I$, with I being the identity on \mathbb{R} . Existence and uniqueness of strong solutions is obtained through a suitable change of variable, rewriting the equation in the dual space H^{-1} , and using regularization and passage to the limit techniques. Such reformulation hinges on the linearity of $\beta_1 = I$.

The only contribution tackling the genuinely doubly nonlinear case is SAPOUNTZOGLU, WITTBOLD, & ZIMMERMANN [45], where nonetheless A is assumed to be Lipschitz-continuous, B has the specific divergence form of (1), and $F = 0$. By contrast, these assumptions are dropped here. In particular, we stress that in our analysis A is allowed to be multivalued, as for $\alpha = \text{sign} + \tilde{\alpha}$ mentioned above.

Our setting is exactly that of DI BENEDETTO & SHOWALTER [18], whose findings we extend here to the stochastic case. In particular, note that doubly nonlinear equations in the form (2) cannot be treated using existing techniques. Indeed, the intuitive substitution $v = A(u)$ does not work, as the operator $B \circ A^{-1}$ is not well-defined on V^* due to the lack of coercivity of A

on V . For this reason, equations in the form (2) are usually referred to as *implicit*, in the sense that they *cannot* be rewritten in the form $dv + \tilde{B}v dt = G dW$, for any suitable choice of the operator \tilde{B} .

Our main result is the existence of martingale solutions to problem (2) (Theorem 2.8). These are obtained via a regularization and passage-to-the-limit procedure. The identification of the nonlinear term Au in the limit directly follows from a compactness argument, based on the linear boundedness of A in the intermediate space H . On the other hand, the limiting Bu is identified by a semicontinuity argument, which in turn hinges on the availability of an Itô formula for φ^* (Proposition 4.1), where $\partial\varphi^* = A^{-1}$. This however does not fall within the framework of PARDOUX [41] or DA PRATO & ZABCZYK [17] due to the nonlinearity of A . Note that well-known approximation techniques based on regularization through linear smoothing operators are ineffective in our framework, since they are in general not compatible with the nonlinearity A . Such difficulties are overcome by an *ad-hoc* regularization based on smoothing nonlinear elliptic operators: this procedure turns out to be effective for our purpose, and requires a specific and detailed asymptotic analysis. For the validity of Itô's formula, whose proof represents the technical core of the paper, one has to ask A^{-1} to be well-behaved. In particular, A^{-1} is here assumed to be Gâteaux differentiable and its differential $D(A^{-1})$ to be smooth enough. These smoothness assumptions are nonetheless fulfilled in the case of (1) whenever α^{-1} is smooth enough (see Section 7). Let us point out that such condition is satisfied also when α has the form $\alpha = \text{sign} + \tilde{\alpha}$, as in the doubly nonlinear stochastic Stefan problem. In particular, α needs not be Lipschitz-continuous nor single-valued.

Exactly as in the deterministic situation [18], in case A or B is linear, continuous, and symmetric, one can prove that martingale solutions are unique. It hence turns out that they are also strong in probability (Theorem 2.9). The conditions ensuring uniqueness are essentially sharp, in the sense that if any of them is not satisfied then nonuniqueness of solutions to problem (2) may occur, even if $F = G = 0$ and $V = H = \mathbb{R}$.

Let us now briefly summarize how the paper is structured. We fix the setting and state our main results in Section 2. A collection of preliminary observations on the approximation of the nonlinear operators is recorded in Section 3. Then, Section 4 is devoted to the proof of the above-mentioned Itô formula. This is used in Section 5 in order to prove the existence of martingale solutions, namely Theorem 2.8. Theorem 2.9 on existence and uniqueness of strong solutions in probability is then proved in Section 6. Eventually, we collect some comments on the application of our abstract existence results to SPDEs of the form (1) in Section 7.

2. SETTING AND STATEMENT OF THE MAIN RESULTS

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space endowed with a filtration $(\mathcal{F}_t)_{t \in [0, T]}$ which is saturated and right-continuous, where $T > 0$ is a fixed final time.

For any Banach space E we shall use the usual symbols $L^p(\Omega; E)$, $L^p(0, T; E)$ and $C^0([0, T]; E)$ for the spaces of p -Bochner integrable E -valued functions on Ω and $(0, T)$, and for the space of continuous functions $[0, T] \rightarrow E$, respectively. Furthermore, if E_1 and E_2 are Banach spaces, the symbols $\mathcal{L}(E_1, E_2)$, $\mathcal{L}_s(E_1, E_2)$ and $\mathcal{L}_w(E_1, E_2)$ denote the space of linear continuous operators from E_1 to E_2 endowed with the norm topology, strong operator topology, or weak operator topology, respectively. If E_1 and E_2 are Hilbert spaces, we shall also use $\mathcal{L}^1(E_1, E_2)$ and $\mathcal{L}^2(E_1, E_2)$ to indicate the spaces of trace-class and Hilbert-Schmidt operators from E_1 to E_2 , respectively.

Let W be a cylindrical Wiener process on a separable Hilbert space U . This amounts to saying that W is *formally* defined as the infinite sum

$$W(t) = \sum_{k=0}^{\infty} \beta_k(t) e_k, \quad t \in [0, T],$$

where $(e_k)_k$ is a complete orthonormal system in U and $(\beta_k)_k$ are real-valued independent Brownian motions.

Let V and H be separable Hilbert spaces such that $V \hookrightarrow H$ densely, continuously and compactly. By identifying H with its dual H^* , (V, H, V^*) turns out to be a classical Hilbert triplet. In particular,

$$V \xhookrightarrow{c} H \xhookrightarrow{c} V^*,$$

where all inclusions are dense, continuous, and compact. Norms, scalar products, and dualities will be denoted by the symbols $\|\cdot\|$, (\cdot, \cdot) , and $\langle \cdot, \cdot \rangle$, respectively, with a sub-script specifying the spaces in consideration. We shall denote by $R : V \rightarrow V^*$ the Riesz isomorphism of V and define the Hilbert space

$$V_0 := \{x \in V : Rx \in H\}, \quad \|x\|_{V_0}^2 := \|x\|_V^2 + \|Rx\|_H^2, \quad x \in V_0.$$

Note that the inclusion $V_0 \hookrightarrow V$ is compact and dense. Indeed, if $(x_n)_n \subset V_0$, $x \in V_0$ and $x_n \rightharpoonup x$ in V_0 , then by compactness of V in H we have $x_n \rightarrow x$ strongly in H ; since also $Rx_n \rightarrow Rx$ in H , we infer that

$$\|x_n\|_V^2 = \langle Rx_n, x_n \rangle_V = (Rx_n, x_n)_H \rightarrow (Rx, x)_H = \|x\|_V^2,$$

so that $x_n \rightarrow x$ in V strongly, and $V_0 \xhookrightarrow{c} V$ compactly. Moreover, it is also not difficult to check that V_0 is dense in V .

The following assumptions will be in order throughout the work:

(H1): $A = \partial\varphi$, where $\varphi : H \rightarrow [0, +\infty)$ is proper, convex, and lower-semicontinuous, $A(0) \ni 0$, and there exists $C_A > 0$ such that

$$\|y\|_H \leq C_A(1 + \|x\|_H) \quad \forall x \in H, \quad \forall y \in A(x).$$

For any $\varepsilon > 0$, the ε -Yosida approximation of A will be denoted by A^ε , namely $A^\varepsilon := (I - (I + \varepsilon A)^{-1})/\varepsilon$ where I is the identity on H (see [16]), and we assume that

$$(A^\varepsilon(x), Rx)_H \geq 0 \quad \forall x \in V_0, \quad \forall \varepsilon > 0.$$

(H2): A is strongly monotone on H , i.e. there exists $c_A > 0$ such that

$$(y_1 - y_2, x_1 - x_2)_H \geq c_A \|x_1 - x_2\|_H^2 \quad \forall x_i \in D(A), \quad \forall y_i \in A(x_i), \quad i = 1, 2.$$

Since $A(0) \ni 0$, this implies in particular that A is coercive on H , hence also surjective by maximal monotonicity, and that the inverse operator $A^{-1} : H \rightarrow H$ is well-defined and Lipschitz-continuous.

(H3): $A^{-1} : H \rightarrow H$ is Gâteaux-differentiable and there exists a Banach space $Y \hookrightarrow H$ continuously and densely such that

$$D(A^{-1}) \in C^0(H; \mathcal{L}_w(H, H) \cap \mathcal{L}(Y, H)), \\ [I + D(A^{-1})((I + A^{-1})^{-1}x)]^{-1} \in \mathcal{L}(V, Y) \quad \forall x \in V.$$

Note that such assumption implies that $(\varphi^*)|_V \in C^2(V)$, so that in particular $D(A^{-1})(y)$ is symmetric for all $y \in V$.

(H4): for any family $(x_\varepsilon)_{\varepsilon>0} \subset V$, $x \in V$, and $y \in A(x)$ such that $x_\varepsilon \rightharpoonup x$ in V and $A^\varepsilon(x_\varepsilon) \rightharpoonup y$ in H as $\varepsilon \searrow 0$, it holds

$$D(A^{-1})(A^\varepsilon(x_\varepsilon)) \rightarrow D(A^{-1})(y) \quad \text{in } \mathcal{L}_s(H, H).$$

This assumption is of a technical nature and has to be checked in each specific problem. Note nonetheless that it is satisfied in several relevant situations (see Section 7 for some concrete examples).

(H5): $B : V \rightarrow 2^{V^*}$ is maximal monotone and there exists $C_B, c_B > 0$ such that

$$\|y\|_{V^*} \leq C_B(1 + \|x\|_V) \quad \text{and} \quad \langle y, x \rangle_V \geq c_B \|x\|_V^2 \quad \forall x \in V, \quad \forall y \in B(x).$$

(H6): $F : [0, T] \times H \rightarrow H$ is measurable with $F(\cdot, 0) \in L^2(0, T; H)$ and there exists $L_F > 0$ such that, for almost every $t \in (0, T)$,

$$\|F(t, x_1) - F(t, x_2)\|_H \leq L_F \|x_1 - x_2\|_H \quad \forall x_1, x_2 \in H.$$

(H7): $G : [0, T] \times H \rightarrow \mathcal{L}^2(U, H)$ is measurable and there exists $L_G > 0$ such that, for every $t \in [0, T]$ and $x_1, x_2, x \in H$,

$$\begin{aligned} \|G(t, x_1) - G(t, x_2)\|_{\mathcal{L}^2(U, H)} &\leq L_G \|x_1 - x_2\|_H, \\ \|G(t, x)\|_{\mathcal{L}^2(U, H)} &\leq L_G(1 + \|x\|_H). \end{aligned}$$

(H8): $v_0 \in L^q(\Omega, \mathcal{F}_0; H)$, $\varphi^*(v_0) \in L^{q/2}(\Omega, \mathcal{F}_0)$, $u_0 := A^{-1}(v_0) \in L^q(\Omega, \mathcal{F}_0; V)$, for $q > 2$.

Let us comment now on assumptions (H1)–(H5), pointing out their major consequences and giving some sufficient conditions for these to hold. For explicit examples of operators A and B we refer to Section 7.

Remark 2.1 (Hypothesis (H1)). Assumption (H1) is very common in the context of doubly nonlinear evolution equations: see for example [18]. In particular, it entails that $A|_V : V \rightarrow 2^{V^*}$ is maximal monotone. Indeed, the monotonicity is trivial. As for the maximality, note that for any $\varepsilon > 0$ the Yosida approximation $A^\varepsilon : H \rightarrow H$ is Lipschitz-continuous, so that for every $y \in V^*$ there is a unique $x_\varepsilon \in V$ such that

$$Rx_\varepsilon + A^\varepsilon x_\varepsilon = y.$$

Testing by $x_\varepsilon \in V$, it follows that

$$\|x_\varepsilon\|_V^2 + (A^\varepsilon x_\varepsilon, x_\varepsilon)_H = \langle y, x_\varepsilon \rangle_V \leq \frac{1}{2} \|x_\varepsilon\|_V^2 + \frac{1}{2} \|y\|_{V^*}^2,$$

so that $(x_\varepsilon)_\varepsilon$ is bounded in V . Recalling (H1) we deduce that $(A_\varepsilon x_\varepsilon)_\varepsilon$ is bounded in H . Hence, there are $x \in V$ and $z \in H$ such that, as $\varepsilon \searrow 0$, $x_\varepsilon \rightharpoonup x$ in V , $x_\varepsilon \rightarrow x$ in H , and $A^\varepsilon x_\varepsilon \rightharpoonup z$ in H , which yield $z \in Ax$ by strong-weak closure of A in $H \times H$. Letting $\varepsilon \searrow 0$ we also deduce that $Rx + z = y$, from which we conclude that $A|_V : V \rightarrow 2^{V^*}$ is maximal. Since one readily has that $A|_V \subseteq \partial\varphi|_V : V \rightarrow 2^{V^*}$, by maximality of $A|_V$ it holds $A|_V = \partial\varphi|_V$.

Remark 2.2 (Hypothesis (H2)). Note that, although $A^{-1} : H \rightarrow H$ is Lipschitz-continuous by (H2), A can still be multivalued. A relevant class of strongly monotone operators A is given by those of the form $A = cI + \tilde{A}$, where \tilde{A} is maximal monotone on H and $c > 0$: see Section 7.

Remark 2.3 (Hypothesis (H3)). We are requiring that A^{-1} is Gâteaux-differentiable on the whole space H , with Gâteaux derivative $D(A^{-1})$ continuous from H to $\mathcal{L}_w(H, H)$, and Fréchet-differentiable with continuous derivative in Y . Of course, if $A^{-1} \in C^1(H, H)$ these conditions are easily satisfied. However, in many applications the operator A^{-1} is *not* Fréchet-differentiable in H , so it is important to require less stringent assumptions as in (H3). For example, if A is

the Nemitzsky operator associated to a maximal monotone graph α on \mathbb{R} , it is well-known that A^{-1} is always Gâteaux differentiable as soon as α^{-1} is $C^{1,1}$, and is Fréchet-differentiable if and only if the α^{-1} is an affine function, which is clearly a too restrictive condition. Further details are given in Section 7.

Remark 2.4 (Hypothesis (H4)). If A is the Nemitzsky operator associated to a maximal monotone graph α on \mathbb{R} , as in the framework of problem (1), it can be seen that (H4) is always satisfied when α is a continuous function (single-valued), not necessarily Lipschitz-continuous. Additionally, one could consider *multivalued* graphs α as well, see Section 7 for details.

Remark 2.5 (Hypothesis (H5)). In [45] the operator B is just required to be polynomially bounded, which calls for framing the problem in a Banach-space setting. We impose a linear bound on B instead, see (H5), which allows a formulation in Hilbert spaces.

Let us now state the concepts of strong and martingale solution for problem (2).

Definition 2.6 (Strong solution). *A strong solution to (2) is a triple (u, v, w) of progressively measurable processes with values in V , H , and V^* , respectively, such that*

$$\begin{aligned} u &\in L^0(\Omega; L^2(0, T; V)), \\ v &\in L^0(\Omega; L^2(0, T; H) \cap C^0([0, T]; V^*)), \\ w &\in L^0(\Omega; L^2(0, T; V^*)), \\ v &\in Au, \quad w \in Bu \quad \text{a.e. in } \Omega \times (0, T), \\ v(t) + \int_0^t w(s) ds &= v_0 + \int_0^t F(s, u(s)) ds + \int_0^t G(s, u(s)) dW(s) \quad \text{in } V^* \quad \forall t \in [0, T] \quad \mathbb{P}\text{-a.s.} \end{aligned}$$

Definition 2.7 (Martingale solution). *A martingale solution to (2) is a quintuplet*

$$((\hat{\Omega}, \hat{\mathcal{F}}, (\hat{\mathcal{F}}_t)_{t \in [0, T]}, \hat{\mathbb{P}}), \hat{W}, \hat{u}, \hat{v}, \hat{w}),$$

where $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$ is a probability space endowed with a filtration $(\hat{\mathcal{F}}_t)_{t \in [0, T]}$ which is saturated and right-continuous, \hat{W} is a $(\hat{\mathcal{F}}_t)_t$ -cylindrical Wiener process on U , and \hat{u} , \hat{v} , and \hat{w} are progressively measurable processes with values in V , H , and V^* , respectively, such that

$$\begin{aligned} \hat{u} &\in L^0(\hat{\Omega}; L^2(0, T; V)), \\ \hat{v} &\in L^0(\hat{\Omega}; L^2(0, T; H) \cap C^0([0, T]; V^*)), \\ \hat{w} &\in L^0(\hat{\Omega}; L^2(0, T; V^*)), \\ \hat{v} &\in A\hat{u}, \quad \hat{w} \in B\hat{u} \quad \text{a.e. in } \hat{\Omega} \times (0, T), \\ \hat{v}(t) + \int_0^t \hat{w}(s) ds &= \hat{v}(0) + \int_0^t F(s, \hat{u}(s)) ds + \int_0^t G(s, \hat{u}(s)) d\hat{W}(s) \quad \text{in } V^* \quad \forall t \in [0, T] \quad \hat{\mathbb{P}}\text{-a.s.} \end{aligned}$$

and $\hat{v}(0)$ has the same law of v_0 on V^* .

The main results of this work are the following.

Theorem 2.8 (Existence of martingale solutions). *Assume (H1)–(H8). Then problem (2) admits a martingale solution which additionally satisfies*

$$\begin{aligned} \hat{u} &\in L^q(\hat{\Omega}; L^\infty(0, T; H) \cap L^2(0, T; V)), \\ \hat{v} &\in L^q(\hat{\Omega}; L^\infty(0, T; H) \cap C^0([0, T]; V^*)), \\ \hat{w} &\in L^q(\hat{\Omega}; L^2(0, T; V^*)). \end{aligned}$$

Theorem 2.9 (Existence and uniqueness of strong solutions). *Assume (H1)–(H8) and that the initial datum $v_0 \in H$ is nonrandom. If either A or B is linear, continuous, and symmetric, then problem (2) admits a unique strong solution (u, v, w) with*

$$\begin{aligned} u &\in L^q(\Omega; L^\infty(0, T; H) \cap L^2(0, T; V)), \\ v &\in L^q(\Omega; L^\infty(0, T; H) \cap C^0([0, T]; V^*)), \\ w &\in L^q(\Omega; L^2(0, T; V^*)). \end{aligned}$$

Remark 2.10. Note that the conditions on A and B ensuring uniqueness are sharp, in the sense that they cannot be weakened, even for $H = \mathbb{R}$: see [18] for details.

3. PRELIMINARY RESULTS

We collect in this section some auxiliary results that will be used throughout the work.

Recall that, for all $\varepsilon \in (0, 1)$, A^ε is the ε -Yosida approximation of A , namely, $A^\varepsilon := (I - (I + \varepsilon A)^{-1})/\varepsilon$ where I is the identity on H . We can equivalently rewrite A^ε as

$$A^\varepsilon := (\varepsilon I + A^{-1})^{-1} : H \rightarrow H.$$

Note that the operator $\varepsilon I + A^{-1} : H \rightarrow H$ is maximal monotone, Lipschitz-continuous, and coercive on H , hence invertible. In order to prove that the two definitions coincide, fix $x \in H$, and define $y^\varepsilon := (\varepsilon I + A^{-1})^{-1}(x)$ so that $\varepsilon y^\varepsilon + A^{-1}(y^\varepsilon) = x$. Setting also $x^\varepsilon := A^{-1}(y^\varepsilon)$ we infer that $x^\varepsilon + \varepsilon A(x^\varepsilon) \ni x$, so that $x^\varepsilon = J_A^\varepsilon(x)$, where $J_A^\varepsilon := (I + \varepsilon A)^{-1} : H \rightarrow H$ is the resolvent of A . Eventually, we deduce that $y^\varepsilon = (x - x^\varepsilon)/\varepsilon$, as desired.

We collect some useful properties of A^ε in the following lemma.

Lemma 3.1 (Properties of A^ε). *Let $\varepsilon \in (0, c_A^{-1})$. Then the following properties hold:*

(P1): *for every $x_1, x_2, x \in H$ it holds*

$$\begin{aligned} (A^\varepsilon(x_1) - A^\varepsilon(x_2), x_1 - x_2)_H &\geq \frac{c_A}{2} \|x_1 - x_2\|_H^2, \\ \|A^\varepsilon(x_1) - A^\varepsilon(x_2)\|_H &\leq \frac{1}{\varepsilon} \|x_1 - x_2\|_H, \\ \|A^\varepsilon(x)\|_H &\leq C_A (1 + 2\|x\|_H). \end{aligned}$$

(P2): *there exists a convex function $\varphi^\varepsilon : H \rightarrow [0, +\infty)$ with $\varphi^\varepsilon \in C^1(H)$, $D\varphi^\varepsilon = A^\varepsilon$ and*

$$(\varphi^\varepsilon)^*(y) = \frac{\varepsilon}{2} \|y\|_H^2 + \varphi^*(y) \quad \forall y \in H.$$

(P3): *A^ε is Gâteaux-differentiable and $D(A^\varepsilon) \in C^0(H; \mathcal{L}(V, H))$.*

Proof. Ad (P1). Since $A^{-1} : H \rightarrow H$ is monotone and Lipschitz continuous, the operator $\varepsilon I + A^{-1}$ is monotone, coercive, and Lipschitz-continuous. Hence, $A^\varepsilon : H \rightarrow H$ is well-defined, monotone, and Lipschitz-continuous as well. Moreover, for every $x_1, x_2 \in H$, setting $y_1 :=$

$A^\varepsilon(x_1)$ and $y_2 := A^\varepsilon(x_2)$, by strong monotonicity of A we have

$$\begin{aligned} & (A^\varepsilon(x_1) - A^\varepsilon(x_2), x_1 - x_2)_H \\ &= (y_1 - y_2, \varepsilon(y_1 - y_2) + A^{-1}(y_1) - A^{-1}(y_2))_H \\ &= \varepsilon \|y_1 - y_2\|_H^2 + (y_1 - y_2, A^{-1}(y_1) - A^{-1}(y_2))_H \\ &\geq \varepsilon \|y_1 - y_2\|_H^2 + c_A \|A^{-1}(y_1) - A^{-1}(y_2)\|_H^2 \\ &= \varepsilon \|y_1 - y_2\|_H^2 + c_A \|x_1 - x_2 - \varepsilon(y_1 - y_2)\|_H^2. \end{aligned}$$

Noting that $\varepsilon c_A \leq 1$ by assumption, we also have

$$\begin{aligned} \frac{c_A}{2} \|x_1 - x_2\|_H^2 &\leq c_A \|x_1 - x_2 - \varepsilon(y_1 - y_2)\|_H^2 + c_A \varepsilon^2 \|y_1 - y_2\|_H^2 \\ &\leq c_A \|x_1 - x_2 - \varepsilon(y_1 - y_2)\|_H^2 + \varepsilon \|y_1 - y_2\|_H^2, \end{aligned}$$

so that the strong monotonicity of A^ε follows. The Lipschitz-continuity of A^ε is well-known (see e.g. [16]). Finally, if $x \in H$ and $y := A^\varepsilon(x)$, we have $\varepsilon y + A^{-1}(y) = x$, so that $y \in A(x - \varepsilon y)$: hence, by (H1) and the already proved Lipschitz-continuity we have

$$\|y\|_H \leq C_A (1 + \|x - \varepsilon y\|_H) \leq C_A (1 + \|x\|_H + \varepsilon \|y\|_H) \leq C_A (1 + 2 \|x\|_H),$$

from which the linear growth condition of A^ε follows, uniformly in ε .

Ad (P2). It is well-known that, by defining φ^ε as the Moreau-Yosida regularization of φ , namely

$$\varphi^\varepsilon(x) := \inf_{v \in H} \left(\frac{\|x - v\|_H^2}{2\varepsilon} + \varphi(v) \right),$$

we have $\partial\varphi^\varepsilon = A^\varepsilon$ and $\partial(\varphi^\varepsilon)^* = (A^\varepsilon)^{-1}$. Since A^ε is Lipschitz-continuous, we deduce that actually $\varphi^\varepsilon \in C^1(H)$ and $D\varphi^\varepsilon = A^\varepsilon$, as required.

Ad (P3). Let us show that A^ε is Gâteaux differentiable. By (H3), $(A^\varepsilon)^{-1} = \varepsilon I + A^{-1}$ is Gâteaux differentiable and, for all $y \in H$, $D((A^\varepsilon)^{-1})(y) = \varepsilon I + D(A^{-1})(y)$ is a linear isomorphism of H . Hence, this implies that A^ε is Gâteaux-differentiable and its differential is given by

$$D(A^\varepsilon)(x) = (\varepsilon I + D(A^{-1})(A^\varepsilon(x)))^{-1} \in \mathcal{L}(H, H), \quad x \in H.$$

Let us show that $D(A^\varepsilon) \in C^0(H; \mathcal{L}(V, H))$. Let $(x_n)_n \subset H$, $x \in H$ with $x_n \rightarrow x$ in H , and $z \in V$ be arbitrary. Setting

$$h_n := D(A^\varepsilon)(x_n)z, \quad h := D(A^\varepsilon)(x)z,$$

we have that

$$z = \varepsilon h_n + D(A^{-1})(A^\varepsilon(x_n))h_n, \quad z = \varepsilon h + D(A^{-1})(A^\varepsilon(x))h,$$

from which

$$\varepsilon(h_n - h) + D(A^{-1})(A^\varepsilon(x_n))(h_n - h) = (D(A^{-1})(A^\varepsilon(x)) - D(A^{-1})(A^\varepsilon(x_n)))h.$$

Hence, testing by $h_n - h$, using the monotonicity of A^{-1} we have

$$\begin{aligned} \varepsilon \|h_n - h\|_H^2 &\leq ((D(A^{-1})(A^\varepsilon(x)) - D(A^{-1})(A^\varepsilon(x_n)))h, h_n - h)_H \\ &\leq \frac{\varepsilon}{2} \|h_n - h\|_H^2 + \frac{1}{2\varepsilon} \|(D(A^{-1})(A^\varepsilon(x)) - D(A^{-1})(A^\varepsilon(x_n)))h\|_H^2, \end{aligned}$$

yielding

$$\|h_n - h\|_H \leq \frac{1}{\varepsilon} \|(D(A^{-1})(A^\varepsilon(x)) - D(A^{-1})(A^\varepsilon(x_n)))h\|_H.$$

Now, recalling (H3), we have that $D(A^\varepsilon)(x) = (\varepsilon I + D(A^{-1})(A^\varepsilon(x)))^{-1} \in \mathcal{L}(V, Y)$: hence, since $z \in V$, we infer that

$$h = D(A^\varepsilon)(x)z \in Y \quad \text{and} \quad \|h\|_Y \leq \|D(A^\varepsilon)(x)\|_{\mathcal{L}(V, Y)} \|z\|_V .$$

We deduce that

$$\|h_n - h\|_H \leq \frac{1}{\varepsilon} \|DA^\varepsilon(x)\|_{\mathcal{L}(V, Y)} \|(D(A^{-1})(A^\varepsilon(x)) - D(A^{-1})(A^\varepsilon(x_n)))\|_{\mathcal{L}(Y, H)} \|z\|_V ,$$

hence also, from the arbitrariness of $z \in V$,

$$\begin{aligned} & \|D(A^\varepsilon)(x_n) - D(A^\varepsilon)(x)\|_{\mathcal{L}(V, H)} \\ & \leq \frac{1}{\varepsilon} \|D(A^\varepsilon)(x)\|_{\mathcal{L}(V, Y)} \|(D(A^{-1})(A^\varepsilon(x)) - D(A^{-1})(A^\varepsilon(x_n)))\|_{\mathcal{L}(Y, H)} . \end{aligned}$$

By the Lipschitz-continuity of A^ε we have that $A^\varepsilon(x_n) \rightarrow A^\varepsilon(x)$ in H as $n \rightarrow \infty$, hence the right-hand side converges to 0 as $n \rightarrow \infty$ again by (H3). \square

For every $\varepsilon, \lambda > 0$, we define the operator

$$A_\lambda^\varepsilon := \lambda R + A_{|V}^\varepsilon : V \rightarrow V^* .$$

Since A_λ^ε is maximal monotone and coercive, its inverse $(A_\lambda^\varepsilon)^{-1} : V^* \rightarrow V$ is well-defined and Lipschitz-continuous. We prove some properties of A_λ^ε in the following lemma.

Lemma 3.2 (Properties of A_λ^ε). *Let $\varepsilon, \lambda > 0$ and define*

$$\varphi_\lambda^\varepsilon : V \rightarrow [0, +\infty), \quad \varphi_\lambda^\varepsilon(x) := \frac{\lambda}{2} \|x\|_V^2 + \varphi^\varepsilon(x), \quad x \in V .$$

Then the convex conjugate $(\varphi_\lambda^\varepsilon)^$ of $\varphi_\lambda^\varepsilon$, defined as*

$$(\varphi_\lambda^\varepsilon)^* : V^* \rightarrow [0, +\infty), \quad (\varphi_\lambda^\varepsilon)^*(y) := \sup_{x \in V} \{ \langle y, x \rangle_V - \varphi_\lambda^\varepsilon(x) \}, \quad y \in V^* ,$$

satisfies $(\varphi_\lambda^\varepsilon)^ \in C^2(V^*)$, $D(\varphi_\lambda^\varepsilon)^* = (A_\lambda^\varepsilon)^{-1}$, and $D(\varphi_\lambda^\varepsilon)^*$ and $D^2(\varphi_\lambda^\varepsilon)^*$ are locally bounded in V^* . Moreover, $(A_\lambda^\varepsilon)^{-1} : H \rightarrow V_0$ is Gâteaux-differentiable and the following characterization holds: for every $y \in H$, setting $x_\lambda^\varepsilon = (A_\lambda^\varepsilon)^{-1}(y)$, we have*

$$D((A_\lambda^\varepsilon)^{-1})(y) = [I + \lambda D((A^\varepsilon)^{-1})(A^\varepsilon(x_\lambda^\varepsilon)) \circ R]^{-1} \circ D((A^\varepsilon)^{-1})(A^\varepsilon(x_\lambda^\varepsilon)) \in \mathcal{L}(H, V_0) .$$

Proof. As λ and ε are fixed throughout, we shall simplify notation and drop them in this proof. By the classical results on the sum of subdifferentials [16, Cor. 2.1, p. 41], we have that $A_\lambda^\varepsilon = \lambda R + A_{|V}^\varepsilon = \partial\varphi_\lambda^\varepsilon$ so that

$$\partial(\varphi_\lambda^\varepsilon)^* = (\lambda R + A_{|V}^\varepsilon)^{-1} : V^* \rightarrow V$$

is Lipschitz-continuous. This implies that $(\varphi_\lambda^\varepsilon)^* \in C^1(V^*)$ and $D(\varphi_\lambda^\varepsilon)^* = (\lambda R + A_{|V}^\varepsilon)^{-1}$ is bounded on bounded subsets of V^* . Let us show now that $(\lambda R + A_{|V}^\varepsilon)^{-1} \in C^1(V^*, V)$. To this end, we first note that the strong monotonicity of A^ε and the definition of Gâteaux derivative readily imply that

$$(DA^\varepsilon(x)h, h)_H \geq \frac{c_A}{2} \|h\|_H^2 \quad \forall x, h \in H .$$

Let $y \in V^*$ be arbitrary and set $x := (\lambda R + A_{|V}^\varepsilon)^{-1}y \in V$. Since R and $A_{|V}^\varepsilon$ are Fréchet-differentiable in x by (P3), it is well-defined the operator

$$D(\lambda R + A_{|V}^\varepsilon)(x) = \lambda R + DA_{|V}^\varepsilon(x) \in \mathcal{L}(V, V^*) :$$

let us show that it is an isomorphism from V to V^* . It is clear that if $k \in V$ satisfies $(\lambda R + A_{|V}^\varepsilon(x))k = 0$, then

$$\lambda \|k\|_V^2 + \frac{c_A}{2} \|k\|_H^2 \leq 0,$$

from which $k = 0$, hence $\lambda R + DA_{|V}^\varepsilon(x)$ is injective. Moreover, since $k \mapsto \lambda Rk + DA_{|V}^\varepsilon(x)k$ is linear, continuous, monotone, and coercive on V , it is also surjective, hence an isomorphism from V to V^* . The theorem on differentiability of the inverse function yields then that $(\lambda R + A_{|V}^\varepsilon)^{-1}$ is Fréchet-differentiable in V^* and

$$D(\lambda R + A_{|V}^\varepsilon)^{-1}(y) = (\lambda R + DA_{|V}^\varepsilon((\lambda R + A_{|V}^\varepsilon)^{-1}y))^{-1} \quad \forall y \in V^*.$$

Let us show finally that $y \mapsto D(\lambda R + A_{|V}^\varepsilon)^{-1}(y)$ is continuous from V^* to $\mathcal{L}(V^*, V)$. Let $y \in V^*$, $(y_n)_n \subset V^*$ with $y_n \rightarrow y$ in V^* , and $h \in V^*$. Setting $k_n := D(\lambda R + A_{|V}^\varepsilon)^{-1}(y_n)h$ and $k := D(\lambda R + A_{|V}^\varepsilon)^{-1}(y)h$, we have

$$\lambda R(k_n - k) + DA_{|V}^\varepsilon((\lambda R + A_{|V}^\varepsilon)^{-1}y_n)k_n - DA_{|V}^\varepsilon((\lambda R + A_{|V}^\varepsilon)^{-1}y)k = h - h = 0,$$

from which

$$\lambda \|k_n - k\|_V^2 \leq \left\| DA_{|V}^\varepsilon((\lambda R + A_{|V}^\varepsilon)^{-1}y_n) - DA_{|V}^\varepsilon((\lambda R + A_{|V}^\varepsilon)^{-1}y) \right\|_{\mathcal{L}(V, V^*)} \|k_n - k\|_V.$$

The Young inequality yields

$$\lambda \|k_n - k\|_V^2 \leq \frac{1}{\lambda} \left\| DA_{|V}^\varepsilon((\lambda R + A_{|V}^\varepsilon)^{-1}y_n) - DA_{|V}^\varepsilon((\lambda R + A_{|V}^\varepsilon)^{-1}y) \right\|_{\mathcal{L}(V, V^*)}^2.$$

Since $(\lambda R + A_{|V}^\varepsilon)^{-1}y_n \rightarrow (\lambda R + A_{|V}^\varepsilon)^{-1}y$ in V and $A_{|V}^\varepsilon \in C^1(V, V^*)$ by (P3), we deduce that the right-hand side converges to 0 as $n \rightarrow \infty$, hence also

$$\left\| D(\lambda R + A_{|V}^\varepsilon)^{-1}(y_n) - D(\lambda R + A_{|V}^\varepsilon)^{-1}(y) \right\|_{\mathcal{L}(V^*, V)}^2 = \sup_{\|h\|_{V^*} \leq 1} \|k_n - k\|_V^2 \rightarrow 0$$

as $n \rightarrow \infty$, from which $(\lambda R + A_{|V}^\varepsilon)^{-1} \in C^1(V^*, V)$. We deduce that $(\varphi_\lambda^\varepsilon)^* \in C^2(V^*)$. Moreover, the fact that $(\lambda R + A_{|V}^\varepsilon)^{-1}$ is Lipschitz-continuous yields immediately that $D^2(\varphi_\lambda^\varepsilon)^*$ is bounded in V^* .

Let us prove the last part of the lemma. Note that, for every $y \in H$, setting $x_\lambda^\varepsilon := (A_\lambda^\varepsilon)^{-1}(v) \in V_0$, we have $\lambda R x_\lambda^\varepsilon + A^\varepsilon(x_\lambda^\varepsilon) = y$, from which $x_\lambda^\varepsilon = \frac{1}{\lambda} R^{-1}(y - A^\varepsilon(x_\lambda^\varepsilon))$, so that

$$(A_\lambda^\varepsilon)^{-1}(y) = \frac{1}{\lambda} R^{-1}(y - A^\varepsilon((A_\lambda^\varepsilon)^{-1}(y))) \quad \forall y \in H.$$

Since we already know that $y \mapsto (A_\lambda^\varepsilon)^{-1}(y) \in C^1(H; V)$, recalling that $A_{|V}^\varepsilon \in C^1(V; H)$ by (P3) and that $R^{-1} : H \rightarrow V_0$ is linear and continuous, we infer that the operator

$$y \mapsto (A_\lambda^\varepsilon)^{-1}(y) = \frac{1}{\lambda} R^{-1}(y - A^\varepsilon((A_\lambda^\varepsilon)^{-1}(y)))$$

is Fréchet-differentiable from H to V_0 . Furthermore, since $x_\lambda^\varepsilon = (A^\varepsilon)^{-1}(y - \lambda R x_\lambda^\varepsilon)$, we also have that

$$(A_\lambda^\varepsilon)^{-1}(y) = (A^\varepsilon)^{-1}(y - \lambda R (A_\lambda^\varepsilon)^{-1}(y)) \quad \forall y \in H.$$

Since we have just proved that $y \mapsto (A_\lambda^\varepsilon)^{-1}(y)$ is Fréchet-differentiable from H to V_0 , taking into account that $R : V_0 \rightarrow H$ is linear continuous and that $(A^\varepsilon)^{-1}$ is Lipschitz-continuous and Gâteaux-differentiable from H to H , we get

$$D((A_\lambda^\varepsilon)^{-1})(y) = D((A^\varepsilon)^{-1})(y - \lambda R x_\lambda^\varepsilon) \circ (I - \lambda R \circ D((A_\lambda^\varepsilon)^{-1})(y)),$$

from which

$$[I + \lambda D((A^\varepsilon)^{-1})(A^\varepsilon(x_\lambda^\varepsilon)) \circ R] \circ D((A_\lambda^\varepsilon)^{-1})(y) = D((A^\varepsilon)^{-1})(A^\varepsilon(x_\lambda^\varepsilon)).$$

Since $[I + \lambda D((A^\varepsilon)^{-1})(A^\varepsilon(x_\lambda^\varepsilon)) \circ R] \in \mathcal{L}(V_0, H)$ is a linear isomorphism, we conclude. \square

The next lemmata state some asymptotic properties of the operator A_λ^ε when $\lambda \searrow 0$ and ε is fixed.

Lemma 3.3. *Let $y \in H$ and $\varepsilon \in (0, c_A^{-1})$ be fixed. For any $\lambda > 0$ set $x_\lambda^\varepsilon := (A_\lambda^\varepsilon)^{-1}(y)$: then, as $\lambda \searrow 0$,*

$$\begin{aligned} x_\lambda^\varepsilon &\rightarrow (A^\varepsilon)^{-1}(y) \quad \text{in } H, \\ A^\varepsilon(x_\lambda^\varepsilon) &\rightarrow y \quad \text{in } H, \\ D((A_\lambda^\varepsilon)^{-1})(y) &\rightarrow D((A^\varepsilon)^{-1})(y) \quad \text{in } \mathcal{L}_w(H, H). \end{aligned}$$

Proof. Since $\lambda R x_\lambda^\varepsilon + A^\varepsilon(x_\lambda^\varepsilon) = y$, testing by x_λ^ε and using (P1) we get

$$\lambda \|x_\lambda^\varepsilon\|_V^2 + \frac{c_A}{2} \|x_\lambda^\varepsilon\|_H^2 \leq \|y\|_H \|x_\lambda^\varepsilon\|_H \leq \frac{c_A}{4} \|x_\lambda^\varepsilon\|_H^2 + \frac{1}{c_A} \|y\|_H^2,$$

from which $\lambda x_\lambda^\varepsilon \rightarrow 0$ in V . Moreover, testing by $A^\varepsilon(x_\lambda^\varepsilon)$ and using the monotonicity assumption in (H1) we get $\|A^\varepsilon(x_\lambda^\varepsilon)\|_H \leq \|y\|_H$ for every λ , from which the second convergence follows by the uniform convexity of H . The first convergence is then a consequence of the fact that $(A^\varepsilon)^{-1}$ is Lipschitz-continuous. As for the third one, let $h \in H$ be arbitrary, and set $k_\lambda^\varepsilon := D((A_\lambda^\varepsilon)^{-1})(y)h$ and $h_\lambda^\varepsilon := D((A^\varepsilon)^{-1})(A^\varepsilon(x_\lambda^\varepsilon))h$, so that by Lemma 3.2 we have

$$k_\lambda^\varepsilon + \lambda D((A^\varepsilon)^{-1})(A^\varepsilon(x_\lambda^\varepsilon)) R k_\lambda^\varepsilon = h_\lambda^\varepsilon.$$

Note that by definition of A^ε and (H3), we have

$$h_\lambda^\varepsilon = \varepsilon h + D(A^{-1})(A^\varepsilon(x_\lambda^\varepsilon))h \rightarrow \varepsilon h + D(A^{-1})(y)h = D((A^\varepsilon)^{-1})(y)h \quad \text{in } H,$$

so that in particular $(k_\lambda^\varepsilon)_\lambda$ is bounded in H . Hence, testing by $\lambda R k_\lambda^\varepsilon$ and employing the monotonicity of A^{-1} we get

$$\lambda \|k_\lambda^\varepsilon\|_V^2 + \varepsilon \lambda^2 \|R k_\lambda^\varepsilon\|_H^2 \leq \lambda \|h_\lambda^\varepsilon\|_H \|R k_\lambda^\varepsilon\|_H \leq \frac{\varepsilon \lambda^2}{2} \|R k_\lambda^\varepsilon\|_H^2 + \frac{1}{2\varepsilon} \|h_\lambda^\varepsilon\|_H^2.$$

Since $(h_\lambda^\varepsilon)_\lambda$ is bounded in H , we infer that $(\lambda k_\lambda^\varepsilon)_\lambda$ is bounded in V_0 and $(\lambda^{1/2} k_\lambda^\varepsilon)_\lambda$ is bounded in V : it follows that $\lambda k_\lambda^\varepsilon \rightarrow 0$ in V_0 , hence in particular that $\lambda R k_\lambda^\varepsilon \rightarrow 0$ in H . Since $D(A^{-1}) \in C^0(H; \mathcal{L}(V, H))$ and $D(A^{-1})(A^\varepsilon(x_\lambda^\varepsilon))$ is symmetric, for every $z \in V$ we have

$$(\lambda D((A^\varepsilon)^{-1})(A^\varepsilon(x_\lambda^\varepsilon)) R k_\lambda^\varepsilon, z)_H = \varepsilon (\lambda R k_\lambda^\varepsilon, z)_H + (\lambda R k_\lambda^\varepsilon, D(A^{-1})(A^\varepsilon(x_\lambda^\varepsilon))z)_H \rightarrow 0,$$

where we have used that $\lambda R k_\lambda^\varepsilon \rightarrow 0$ and $D((A^\varepsilon)^{-1})(A^\varepsilon(x_\lambda^\varepsilon))z \rightarrow D((A^\varepsilon)^{-1})(y)z$ in H . Consequently, we have that

$$\lambda D(A^{-1})(A^\varepsilon(x_\lambda^\varepsilon)) R k_\lambda^\varepsilon \xrightarrow{*} 0 \quad \text{in } V^*.$$

Since we also have that $(k_\lambda^\varepsilon)_\lambda$ is bounded in H , hence $k_\lambda^\varepsilon \rightarrow k^\varepsilon$ in H for a certain $k^\varepsilon \in H$, letting $\lambda \searrow 0$ we infer that $k^\varepsilon = D((A^\varepsilon)^{-1})(y)h$, and we conclude. \square

Lemma 3.4. *Let $\varepsilon \in (0, c_A^{-1})$ and $y \in H$ be fixed. For any $\lambda > 0$ let $y_\lambda \in H$, and define $x^\varepsilon := (A^\varepsilon)^{-1}(y)$ and $x_\lambda^\varepsilon := (A_\lambda^\varepsilon)^{-1}(y_\lambda)$. If $x_\lambda^\varepsilon \rightarrow x^\varepsilon$ in H as $\lambda \searrow 0$ and $(y_\lambda)_\lambda$ is bounded in H , then it holds that, as $\lambda \searrow 0$,*

$$\begin{aligned} A^\varepsilon(x_\lambda^\varepsilon) &\rightarrow y \quad \text{in } H, \\ y_\lambda &\rightarrow y \quad \text{in } H, \\ D((A_\lambda^\varepsilon)^{-1})(y_\lambda) &\rightarrow D((A^\varepsilon)^{-1})(y) \quad \text{in } \mathcal{L}_w(H, H). \end{aligned}$$

Proof. First of all, since A^ε is Lipschitz-continuous, we have $A^\varepsilon(x_\lambda^\varepsilon) \rightarrow A^\varepsilon(x^\varepsilon) = y$. Moreover, since $\lambda R x_\lambda^\varepsilon + A^\varepsilon(x_\lambda^\varepsilon) = y_\lambda$, testing by x_λ^ε and using (P1) we get

$$\lambda \|x_\lambda^\varepsilon\|_V^2 + \frac{c_A}{2} \|x_\lambda^\varepsilon\|_H^2 \leq \|y_\lambda\|_H \|x_\lambda^\varepsilon\|_H \leq \frac{c_A}{4} \|x_\lambda^\varepsilon\|_H^2 + \frac{1}{c_A} \|y_\lambda\|_H^2,$$

from which $\lambda x_\lambda^\varepsilon \rightarrow 0$ in V . Hence, by comparison in the equation and the boundedness of $(y_\lambda)_\lambda$ in H we infer that $y_\lambda \rightarrow y$ in H , and the second convergence is proved. Let us show the last one. Let $h \in H$ be arbitrary, and set $k_\lambda^\varepsilon := D((A_\lambda^\varepsilon)^{-1})(y_\lambda)h$ and $h_\lambda^\varepsilon := D((A^\varepsilon)^{-1})(A^\varepsilon(x_\lambda^\varepsilon))h$, so that by Lemma 3.2 we have

$$k_\lambda^\varepsilon + \lambda D((A^\varepsilon)^{-1})(A^\varepsilon(x_\lambda^\varepsilon))Rk_\lambda^\varepsilon = h_\lambda^\varepsilon,$$

where by definition of A^ε and (H3), we have

$$h_\lambda^\varepsilon = \varepsilon h + D(A^{-1})(A^\varepsilon(x_\lambda^\varepsilon))h \rightarrow \varepsilon h + D(A^{-1})(y)h = D((A^\varepsilon)^{-1})(y)h \quad \text{in } H,$$

so that in particular $(h_\lambda^\varepsilon)_\lambda$ is bounded in H . Arguing as in the proof of Lemma 3.3 we obtain that $k_\lambda^\varepsilon \rightarrow k^\varepsilon$ and $\lambda Rk_\lambda^\varepsilon \rightarrow 0$ in H for a certain $k^\varepsilon \in H$. Moreover, for every $z \in V$ we have $D((A^\varepsilon)^{-1})(A^\varepsilon(x_\lambda^\varepsilon))z \rightarrow D((A^\varepsilon)^{-1})(y)z$ in H thanks to (H3), so that by the symmetry of $D(A^{-1})(A^\varepsilon(x_\lambda^\varepsilon))$ we get again

$$(\lambda D((A^\varepsilon)^{-1})(A^\varepsilon(x_\lambda^\varepsilon))Rk_\lambda^\varepsilon, z)_H = \varepsilon (\lambda Rk_\lambda^\varepsilon, z)_H + (\lambda Rk_\lambda^\varepsilon, D(A^{-1})(A^\varepsilon(x_\lambda^\varepsilon))z)_H \rightarrow 0,$$

hence

$$\lambda D(A^{-1})(A^\varepsilon(x_\lambda^\varepsilon))Rk_\lambda^\varepsilon \xrightarrow{*} 0 \quad \text{in } V^*.$$

By comparison in the equation we infer then that $k^\varepsilon = D((A^\varepsilon)^{-1})(y)h$, and we conclude. \square

Finally, we prove a fundamental asymptotic property of A_λ^ε when $\varepsilon = \lambda$ converge jointly to 0. To this end, we introduce for brevity of notation the operator $\tilde{A}_\lambda := A_\lambda^\lambda$ for any $\lambda > 0$.

Lemma 3.5. *Let $y \in H$, $x := A^{-1}(y) \in H$, and for any $\lambda > 0$ set $x_\lambda := \tilde{A}_\lambda^{-1}(y) \in V$. Then, as $\lambda \searrow 0$,*

$$\begin{aligned} x_\lambda &\rightarrow x \quad \text{in } H, \\ A^\lambda(x_\lambda) &\rightarrow y \quad \text{in } H, \\ D(\tilde{A}_\lambda^{-1})(y) &\rightarrow D(A^{-1})(y) \quad \text{in } \mathcal{L}_w(H, H). \end{aligned}$$

Moreover, if $x \in V$ it holds that

$$x_\lambda \rightarrow x \quad \text{in } V.$$

Proof. Since $\lambda R x_\lambda + A^\lambda(x_\lambda) = y$, testing by x_λ and using (P1) we get

$$\lambda \|x_\lambda\|_V^2 + \frac{c_A}{2} \|x_\lambda\|_H^2 \leq \|y\|_H \|x_\lambda\|_H \leq \frac{c_A}{4} \|x_\lambda\|_H^2 + \frac{1}{c_A} \|y\|_H^2,$$

from which $\lambda x_\lambda \rightarrow 0$ in V , so that $A^\lambda(x_\lambda) \rightarrow y$ in V^* . Moreover, testing by $A^\lambda(x_\lambda)$ and using the monotonicity assumption in (H1) we get $\|A^\lambda(x_\lambda)\|_H \leq \|y\|_H$ for every λ , from which the second

convergence follows by the uniform convexity of H . The first convergence is then a consequence of the strong-weak closure of maximal monotone operators. Let us focus on the last one. Let $h_1, h_2 \in H$ be arbitrary and consider the double real sequence $(a_{\lambda\delta})_{\lambda,\delta}$ given by

$$a_{\lambda\delta} := \left(\frac{\tilde{A}_\lambda^{-1}(y + \delta h_1) - \tilde{A}_\lambda^{-1}(y)}{\delta}, h_2 \right)_H .$$

Setting $x_{\lambda\delta} := \tilde{A}_\lambda^{-1}(y + \delta h_1)$, we have

$$\lambda R(x_{\lambda\delta} - x_\lambda) + A^\lambda(x_{\lambda\delta}) - A^\lambda(x_\lambda) \ni \delta h_1 .$$

Testing by $x_{\lambda\delta} - x_\lambda$, using the strong monotonicity of A^λ and the Young inequality, it is a standard matter to check that

$$\lambda \left\| \frac{x_{\lambda\delta} - x_\lambda}{\delta} \right\|_V^2 + \frac{c_A}{2} \left\| \frac{x_{\lambda\delta} - x_\lambda}{\delta} \right\|_H^2 \leq \frac{1}{c_A} \|h_1\|_H^2 + \frac{c_A}{4} \left\| \frac{x_{\lambda\delta} - x_\lambda}{\delta} \right\|_H^2 .$$

Rearranging the terms we deduce that

$$|a_{\lambda\delta}| \leq \frac{2}{c_A} \|h_1\|_H \|h_2\|_H \quad \forall \lambda, \delta > 0 :$$

hence, the Bolzano-Weierstraß theorem for double sequences ensures that there is a double subsequence $(a_{\lambda_j \delta_k})_{j,k}$ and $a \in \mathbb{R}$ such that

$$\lim_{j,k \rightarrow \infty} a_{\lambda_j \delta_k} = a .$$

Now, note that the limits

$$\lim_{\lambda \searrow 0} a_{\lambda\delta} \quad \text{and} \quad \lim_{\delta \searrow 0} a_{\lambda\delta}$$

exist and are finite for every $\delta > 0$ and for every $\lambda > 0$, respectively. Indeed, by the results already proved we have

$$\lim_{\lambda \searrow 0} a_{\lambda\delta} = \left(\frac{A^{-1}(y + \delta h_1) - A^{-1}(y)}{\delta}, h_2 \right)_H \quad \forall \delta > 0$$

and, by definition of Gâteaux-differentiability of \tilde{A}_λ^{-1} ,

$$\lim_{\delta \searrow 0} a_{\lambda\delta} = \left(D(\tilde{A}_\lambda^{-1})(y)h_1, h_2 \right)_H \quad \forall \lambda > 0 .$$

This implies that the limit a can be computed using the iterated limit procedure in any order, i.e.

$$a = \lim_{j \rightarrow \infty} \lim_{k \rightarrow \infty} a_{\lambda_j \delta_k} = \lim_{k \rightarrow \infty} \lim_{j \rightarrow \infty} a_{\lambda_j \delta_k} .$$

In particular, the second equality implies that

$$a = \lim_{k \rightarrow \infty} \left(\frac{A^{-1}(y + \delta_k h_1) - A^{-1}(y)}{\delta_k}, h_2 \right)_H = (D(A^{-1})(y)h_1, h_2)_H .$$

Since this is true for any arbitrary subsequence of (λ, δ) , a well-known criterion of classical analysis implies that the convergence holds also along the original sequence, i.e.

$$\lim_{\lambda, \delta \searrow 0} a_{\lambda\delta} = (D(A^{-1})(y)h_1, h_2)_H .$$

Using the first iterated limit we have then

$$\lim_{\lambda \searrow 0} \left(D(\tilde{A}_\lambda^{-1})(y)h_1, h_2 \right)_H = (D(A^{-1})(y)h_1, h_2)_H ,$$

from which the thesis follows by arbitrariness of h_1 and h_2 .

Finally, let us show that if we also suppose $x \in V$, then $x_\lambda \rightarrow x$ in V . From the relation $\lambda R x_\lambda + A^\lambda(x_\lambda) - y = 0$, we test by x_λ and rearrange the terms in the following way:

$$\lambda \|x_\lambda\|_V^2 + (A^\lambda(x_\lambda) - y, x_\lambda - (A^\lambda)^{-1}(y))_H = (y - A^\lambda(x_\lambda), (A^\lambda)^{-1}(y))_H .$$

Noting that $y = A^\lambda((A^\lambda)^{-1}(y))$, the second term on the left-hand side is nonnegative; moreover, since $(A^\lambda)^{-1}(y) = \lambda y + A^{-1}(y) = \lambda y + x$, we get

$$\lambda \|x_\lambda\|_V^2 \leq (y - A^\lambda(x_\lambda), \lambda y + x)_H \leq \lambda \|y\|_H \|y - A^\lambda(x_\lambda)\|_H + (y - A^\lambda(x_\lambda), x)_H .$$

Recalling that $y - A^\lambda(x_\lambda) = \lambda R x_\lambda$ and using the Young inequality, we get

$$\lambda \|x_\lambda\|_V^2 \leq \lambda \|y\|_H \|y - A^\lambda(x_\lambda)\|_H + \lambda \|x_\lambda\|_V \|x\|_V \leq \lambda \|y\|_H \|y - A^\lambda(x_\lambda)\|_H + \frac{\lambda}{2} \|x_\lambda\|_V^2 + \frac{\lambda}{2} \|x\|_V^2$$

from which, dividing by λ and rearranging the terms we get

$$\|x_\lambda\|_V^2 \leq 2 \|y\|_H \|y - A^\lambda(x_\lambda)\|_H + \|x\|_V^2 .$$

Hence, recalling that $A^\lambda(x_\lambda) \rightarrow y$ in H , we infer that

$$\limsup_{\lambda \searrow 0} \|x_\lambda\|_V^2 \leq 2 \|y\|_H \lim_{\lambda \searrow 0} \|y - A^\lambda(x_\lambda)\|_H + \|x\|_V^2 = \|x\|_V^2 ,$$

from which we conclude by uniform convexity of V . \square

4. A GENERALIZED ITÔ'S FORMULA

In this section we prove a generalized Itô's formula that will be crucial in the proofs on the main results of the work. In particular, due to the weak assumptions on the derivatives of φ^* , we cannot rely directly on the classical frameworks by DA PRATO & ZABCZYK [17] or PARDOUX [41], as the second derivative $D^2\varphi^*$ is assumed to exist in the sense of Fréchet only in V and the process y below is not V -valued a priori. This gives rise to several nontrivial difficulties. Nevertheless, using the continuity of the Gâteaux derivative $D^2\varphi^*$ from H to $\mathcal{L}_w(H, H)$ and the fact that $y \in A(x)$ for a suitable V -valued process x , we are able to show that the Itô formula for φ^* can still be written in an appropriate sense.

Proposition 4.1 (Generalized Itô's formula). *Assume that x, y, w are progressively measurable processes with values in V, H , and V^* , respectively, such that*

$$\begin{aligned} x &\in L^0(\Omega; L^2(0, T; V)) , \\ y &\in L^0(\Omega; L^\infty(0, T; H) \cap C^0([0, T]; V^*)) , \\ w &\in L^0(\Omega; L^2(0, T; V^*)) , \\ y &\in A(x) \quad \text{a.e. in } \Omega \times (0, T) , \\ C &\in L^0(\Omega; L^2(0, T; \mathcal{L}^2(U, H))) , \\ y_0 &\in L^0(\Omega, \mathcal{F}_0; H) , \quad x_0 := A^{-1}(y_0) \in L^0(\Omega, \mathcal{F}_0; V) , \\ y(t) + \int_0^t w(s) ds &= y_0 + \int_0^t C(s) dW(s) \quad \text{in } V^* \quad \forall t \in [0, T] , \quad \mathbb{P}\text{-a.s.} \end{aligned}$$

Then, for every $t \in [0, T]$, \mathbb{P} -almost surely,

$$\begin{aligned} & \varphi^*(y(t)) + \int_0^t \langle w(s), x(s) \rangle_V ds \\ &= \varphi^*(y_0) + \int_0^t (x(s), C(s))_H dW(s) + \frac{1}{2} \int_0^t \text{Tr} (C^*(s)D(A^{-1})(y(s))C(s)) ds. \end{aligned}$$

Since $\partial\varphi^* = A^{-1} : H \rightarrow H$ is Lipschitz-continuous, we have in particular that $H = D(\varphi^*)$: since $y \in C_w([0, T]; H)$ and y_0 is H -valued, the first terms on the left and right hand side in Itô's formula are finite for every t . Moreover, since $x = A^{-1}(y)$ and A^{-1} is Lipschitz-continuous, we also have $x \in L^0(\Omega; L^\infty(0, T; H))$, hence $(x, C)_H \in L^2(0, T; \mathcal{L}^2(U, \mathbb{R}))$ and also the stochastic integral on the right-hand side is well-defined. Finally, the trace term is also well-defined since $D(A^{-1})$ is bounded by assumption (H3).

Proof of Proposition 4.1. For every $\lambda > 0$ by Lemma 3.2 we can apply the classical Itô's formula to the function $\tilde{\varphi}_\lambda^*$, where $\tilde{\varphi}_\lambda := \varphi_\lambda^\lambda$, getting

$$\begin{aligned} & \tilde{\varphi}_\lambda^*(y(t)) + \int_0^t \langle w(s), \tilde{A}_\lambda^{-1}y(s) \rangle_V ds \\ &= \tilde{\varphi}_\lambda^*(y_0) + \int_0^t (\tilde{A}_\lambda^{-1}y(s), C(s))_H dW(s) + \frac{1}{2} \int_0^t \text{Tr} (C^*(s)D(\tilde{A}_\lambda^{-1})(y(s))C(s)) ds \end{aligned}$$

for every $t \in [0, T]$, \mathbb{P} -almost surely. By Lemma 3.5 we have that

$$\begin{aligned} & \tilde{A}_\lambda^{-1}y \rightarrow x \quad \text{in } L^2(0, T; V), \\ & D(\tilde{A}_\lambda^{-1})(y(s)) \rightarrow D(A^{-1})(y(s)) \quad \text{in } \mathcal{L}_w(H, H) \quad \forall s \in [0, T]. \end{aligned}$$

Consequently, we have

$$\int_0^t \langle w(s), \tilde{A}_\lambda^{-1}y(s) \rangle_V ds \rightarrow \int_0^t \langle w(s), x(s) \rangle_V ds$$

and, by the dominated convergence theorem,

$$\int_0^t \text{Tr} (C^*(s)D(\tilde{A}_\lambda^{-1})(y(s))C(s)) ds \rightarrow \int_0^t \text{Tr} (C^*(s)D(A^{-1})(y(s))C(s)) ds.$$

Moreover, since $\partial\tilde{\varphi}_\lambda = \tilde{A}_\lambda$ and $\partial\varphi^\lambda = A^\lambda$, for every $t \in [0, T]$ we have

$$\begin{aligned} & \tilde{\varphi}_\lambda^*(y(t)) = (v(t), \tilde{A}_\lambda^{-1}y(t))_H - \tilde{\varphi}_\lambda(\tilde{A}_\lambda^{-1}y(t)) \\ &= (y(t), \tilde{A}_\lambda^{-1}y(t))_H - \frac{\lambda}{2} \left\| \tilde{A}_\lambda^{-1}y(t) \right\|_V^2 - \varphi^\lambda(\tilde{A}_\lambda^{-1}y(t)) \\ &= (y(t), \tilde{A}_\lambda^{-1}y(t))_H - \frac{\lambda}{2} \left\| \tilde{A}_\lambda^{-1}y(t) \right\|_V^2 \\ &\quad - (A^\lambda(\tilde{A}_\lambda^{-1}(y(t))), \tilde{A}_\lambda^{-1}(y(t)))_H + (\varphi^\lambda)^*(A^\lambda(\tilde{A}_\lambda^{-1}(y(t)))) \\ &= (y(t), \tilde{A}_\lambda^{-1}y(t))_H - \frac{\lambda}{2} \left\| \tilde{A}_\lambda^{-1}y(t) \right\|_V^2 - (A^\lambda(\tilde{A}_\lambda^{-1}(y(t))), \tilde{A}_\lambda^{-1}(y(t)))_H \\ &\quad + \frac{\lambda}{2} \left\| A^\lambda(\tilde{A}_\lambda^{-1}(y(t))) \right\|_H^2 + \varphi^*(A^\lambda(\tilde{A}_\lambda^{-1}(y(t)))) . \end{aligned}$$

Now, by Lemma 3.5 we have that

$$\tilde{A}_\lambda^{-1}(y(t)) \rightarrow x(t) \quad \text{in } H, \quad A^\lambda(\tilde{A}_\lambda^{-1}(y(t))) \rightarrow y(t) \quad \text{in } H, \quad \left\| \tilde{A}_\lambda^{-1}(y(t)) \right\|_V \leq M$$

for a constant M independent of λ . Hence, as $\lambda \searrow 0$ we get

$$\tilde{\varphi}_\lambda^*(y(t)) \rightarrow \varphi^*(y(t)).$$

Similarly, the same argument and the fact that $x_0 = A^{-1}(y_0) \in V$ yields

$$\tilde{\varphi}_\lambda^*(y_0) \rightarrow \varphi^*(y_0).$$

Finally, let us show the convergence of the stochastic integrals: note that

$$\left\| ((\tilde{A}^{-1}y - x), C)_H \right\|_{L^2(0, T; \mathcal{L}^2(U, \mathbb{R}))}^2 \leq \int_0^T \left\| \tilde{A}^{-1}y(s) - x(s) \right\|_H^2 \|C(s)\|_{\mathcal{L}^2(U, H)}^2 ds$$

where the integrand converges pointwise to 0 in $[0, T]$. Moreover, from the proof of Lemma 3.5 we infer that

$$\left\| \tilde{A}^{-1}y - x \right\|_H^2 \|C\|_{\mathcal{L}^2(U, H)}^2 \leq M \|C\|_{\mathcal{L}^2(U, H)}^2 (\|y\|_H^2 + \|x\|_H^2)$$

for a positive constant M only dependent on c_A . Since $x = A^{-1}(y)$ and A^{-1} is Lipschitz-continuous, recalling that $y \in L^\infty(0, T; H)$ we deduce that also $x \in L^\infty(0, T; H)$, hence the dominated convergence theorem yields

$$\left\| ((\tilde{A}^{-1}y - x), C)_H \right\|_{L^2(0, T; \mathcal{L}^2(U, \mathbb{R}))}^2 \rightarrow 0$$

\mathbb{P} -almost surely. We deduce that

$$\int_0^t \left(\tilde{A}_\lambda^{-1}y(s), C(s) \right)_H dW(s) \rightarrow \int_0^t (x(s), C(s))_H dW(s)$$

in probability. Hence, the thesis follows letting $\lambda \searrow 0$. \square

The following result follows using exactly the same proof with A^ε instead of A .

Proposition 4.2 (Generalized Itô's formula for A^ε). *Let $\varepsilon \in (0, c_A^{-1})$. Assume that $x^\varepsilon, y^\varepsilon, w^\varepsilon$ are progressively measurable processes with values in V, H , and V^* , respectively, such that*

$$\begin{aligned} x^\varepsilon &\in L^0(\Omega; L^2(0, T; V)), \\ y^\varepsilon &\in L^0(\Omega; L^\infty(0, T; H) \cap C^0([0, T]; V^*)), \\ w^\varepsilon &\in L^0(\Omega; L^2(0, T; V^*)), \\ y^\varepsilon &= A^\varepsilon(x^\varepsilon) \quad \text{a.e. in } \Omega \times (0, T), \\ C &\in L^0(\Omega; L^2(0, T; \mathcal{L}^2(U, H))), \\ y_0^\varepsilon &\in L^0(\Omega, \mathcal{F}_0; H), \quad x_0^\varepsilon := (A^\varepsilon)^{-1}(y_0^\varepsilon) \in L^2(\Omega, \mathcal{F}_0; V), \\ y^\varepsilon(t) + \int_0^t w^\varepsilon(s) ds &= y_0^\varepsilon + \int_0^t C(s) dW(s) \quad \text{in } V^* \quad \forall t \in [0, T], \quad \mathbb{P}\text{-a.s.} \end{aligned}$$

Then, for every $t \in [0, T]$, \mathbb{P} -almost surely,

$$\begin{aligned} &(\varphi^\varepsilon)^*(y^\varepsilon(t)) + \int_0^t \langle w^\varepsilon(s), x^\varepsilon(s) \rangle_V ds \\ &= (\varphi^\varepsilon)^*(y_0^\varepsilon) + \int_0^t (x^\varepsilon(s), C(s))_H dW(s) + \frac{1}{2} \int_0^t \text{Tr}(C^*(s)D((A^\varepsilon)^{-1})(y^\varepsilon(s))C(s)) ds. \end{aligned}$$

5. EXISTENCE OF MARTINGALE SOLUTIONS: PROOF OF THEOREM 2.8

We introduce two separate approximations on the problem, depending on two different parameters $\lambda, \varepsilon > 0$, prove uniform estimates on the regularized solutions (Subsection 5.1), and pass to the limit as $\lambda \searrow 0$ first (Subsection 5.2) and then as $\varepsilon \searrow 0$ (Subsection 5.3).

5.1. The approximated problem. Let $\varepsilon \in (0, c_A^{-1})$ be fixed. For every $\lambda > 0$, we consider the approximated problem

$$d(A_\lambda^\varepsilon u_\lambda^\varepsilon) + B_\lambda u_\lambda^\varepsilon dt = F(t, u_\lambda^\varepsilon) dt + G(t, u_\lambda^\varepsilon) dW, \quad A_\lambda^\varepsilon u_\lambda^\varepsilon(0) = v_0^\varepsilon,$$

where $B_\lambda : V \rightarrow V^*$ is the Yosida approximation of B and $v_0^\varepsilon := (I + \varepsilon R)^{-1}(v_0)$. We recall that since V is a Hilbert space then R is linear and B_λ is Lipschitz-continuous. Setting $v_\lambda^\varepsilon := A_\lambda^\varepsilon u_\lambda^\varepsilon$, the approximated problem can be written in terms of v_λ^ε as an evolution equation on V^* of the form

$$dv_\lambda^\varepsilon + B_\lambda(A_\lambda^\varepsilon)^{-1}v_\lambda^\varepsilon dt = F(t, (A_\lambda^\varepsilon)^{-1}v_\lambda^\varepsilon) dt + G(t, (A_\lambda^\varepsilon)^{-1}v_\lambda^\varepsilon) dW, \quad v_\lambda^\varepsilon(0) = v_0^\varepsilon.$$

Since $(A_\lambda^\varepsilon)^{-1} : V^* \rightarrow V$ is Lipschitz-continuous, the operator $B_\lambda \circ (A_\lambda^\varepsilon)^{-1} : V^* \rightarrow V^*$ is composition of Lipschitz-continuous operators, and there is a unique strong solution

$$v_\lambda^\varepsilon \in L^q(\Omega; C^0([0, T]; V^*))$$

such that, setting $u_\lambda^\varepsilon := (A_\lambda^\varepsilon)^{-1}v_\lambda^\varepsilon \in L^q(\Omega; C^0([0, T]; V))$, we have

$$v_\lambda^\varepsilon(t) + \int_0^t B_\lambda u_\lambda^\varepsilon(s) ds = v_0 + \int_0^t F(s, u_\lambda^\varepsilon(s)) ds + \int_0^t G(s, u_\lambda^\varepsilon(s)) dW(s) \quad \text{in } V^*$$

for every $t \in [0, T]$, \mathbb{P} -almost surely. We now prove a priori estimates, independent of λ and ε .

Lemma 5.1 (A priori estimates). *Let $\eta \in (0, 1/2)$. Then there exists a positive constant $M > 0$ such that, for any $\lambda > 0$ and $\varepsilon \in (0, c_A^{-1})$,*

$$\begin{aligned} & \lambda^{1/2} \|u_\lambda^\varepsilon\|_{L^q(\Omega; C^0([0, T]; V))} + \varepsilon^{1/2} \|A^\varepsilon(u_\lambda^\varepsilon)\|_{L^q(\Omega; C^0([0, T]; H))} \leq M, \\ & \|u_\lambda^\varepsilon\|_{L^q(\Omega; C^0([0, T]; H))} + \|A^\varepsilon(u_\lambda^\varepsilon)\|_{L^q(\Omega; C^0([0, T]; H))} + \|\varphi^*(A^\varepsilon(u_\lambda^\varepsilon))\|_{L^{q/2}(\Omega; L^\infty(0, T))} \leq M, \\ & \|J_\lambda^B u_\lambda^\varepsilon\|_{L^q(\Omega; L^2(0, T; V))} + \|B_\lambda u_\lambda^\varepsilon\|_{L^q(\Omega; L^2(0, T; V^*))} \leq M, \\ & \|G(\cdot, u_\lambda^\varepsilon)\|_{L^q(\Omega; C^0([0, T]; \mathcal{L}^2(U, H))} \leq M, \\ & \|G(\cdot, u_\lambda^\varepsilon) \cdot W\|_{L^q(\Omega; W^{\eta, q}(0, T; H))} \leq M, \\ & \|v_\lambda^\varepsilon - G(\cdot, u_\lambda^\varepsilon) \cdot W\|_{L^q(\Omega; H^1(0, T; V^*))} \leq M, \\ & \lambda \|R u_\lambda^\varepsilon\|_{L^q(\Omega; W^{\eta, q}(0, T; V^*))} + \|A^\varepsilon(u_\lambda^\varepsilon)\|_{L^q(\Omega; W^{\eta, q}(0, T; V^*))} \leq M, \end{aligned}$$

where $J_\lambda^B := (R + \lambda B)^{-1} \circ R : V \rightarrow V$ is the resolvent of B .

Proof. By Lemma 3.2, we can apply Itô's formula to $(\varphi_\lambda^\varepsilon)^*(v_\lambda^\varepsilon)$ in V^* , getting

$$\begin{aligned} & (\varphi_\lambda^\varepsilon)^*(v_\lambda^\varepsilon(t)) + \int_0^t \langle B_\lambda u_\lambda^\varepsilon(s), u_\lambda^\varepsilon(s) \rangle_V ds = (\varphi_\lambda^\varepsilon)^*(v_0) + \int_0^t (F(s, u_\lambda^\varepsilon(s)), u_\lambda^\varepsilon(s))_H ds \\ & + \int_0^t (u_\lambda^\varepsilon(s), G(s, u_\lambda^\varepsilon(s)))_H dW(s) + \frac{1}{2} \int_0^t \text{Tr} (G^*(s, u_\lambda^\varepsilon(s))(\lambda R + DA^\varepsilon(u_\lambda^\varepsilon(s)))^{-1}G(s, u_\lambda^\varepsilon(s))) ds \end{aligned}$$

for every $t \in [0, T]$, \mathbb{P} -almost surely. Let us analyse the different terms separately. First of all, by definition of convex conjugate and the fact that $\lambda R u_\lambda^\varepsilon + A^\varepsilon(u_\lambda^\varepsilon) = v_\lambda^\varepsilon$ we have

$$\begin{aligned} (\varphi_\lambda^\varepsilon)^*(v_\lambda^\varepsilon) &= (\varphi_\lambda^\varepsilon)^*(\lambda R u_\lambda^\varepsilon + A^\varepsilon(u_\lambda^\varepsilon)) = \langle \lambda R u_\lambda^\varepsilon + A^\varepsilon(u_\lambda^\varepsilon), u_\lambda^\varepsilon \rangle_V - \varphi_\lambda^\varepsilon(u_\lambda^\varepsilon) \\ &= \lambda \|u_\lambda^\varepsilon\|_V^2 + (A^\varepsilon(u_\lambda^\varepsilon), u_\lambda^\varepsilon)_H - \frac{\lambda}{2} \|u_\lambda^\varepsilon\|_V^2 - \varphi^\varepsilon(u_\lambda^\varepsilon) \\ &= \frac{\lambda}{2} \|u_\lambda^\varepsilon\|_V^2 + (\varphi^\varepsilon)^*(A^\varepsilon(u_\lambda^\varepsilon)) = \frac{\lambda}{2} \|u_\lambda^\varepsilon\|_V^2 + \frac{\varepsilon}{2} \|A^\varepsilon(u_\lambda^\varepsilon)\|_H^2 + \varphi^*(A^\varepsilon(u_\lambda^\varepsilon)). \end{aligned}$$

Since φ has at most quadratic growth in H , there are constants $M', M'' > 0$, independent of λ and ε , such that

$$\varphi^*(A^\varepsilon(u_\lambda^\varepsilon)) \geq M' \|A^\varepsilon(u_\lambda^\varepsilon)\|_H^2 - M''.$$

Noting further that $u_\lambda^\varepsilon = \varepsilon A^\varepsilon(u_\lambda^\varepsilon) + A^{-1}(A^\varepsilon(u_\lambda^\varepsilon))$, we infer by the Lipschitz-continuity of A^{-1} that

$$\varphi^*(A^\varepsilon(u_\lambda^\varepsilon)) \geq 2M' c_A^{-1} \|u_\lambda^\varepsilon\|_H^2 - M''.$$

Secondly, setting $u_{0\lambda}^\varepsilon := (\lambda R + A_{|V}^\varepsilon)^{-1} v_0^\varepsilon$, a similar argument yields

$$(\varphi_\lambda^\varepsilon)^*(v_0) = \frac{\lambda}{2} \|u_{0\lambda}^\varepsilon\|_V^2 + (\varphi^\varepsilon)^*(A^\varepsilon(u_{0\lambda}^\varepsilon)).$$

Since $\lambda R u_{0\lambda}^\varepsilon + A^\varepsilon(u_{0\lambda}^\varepsilon) = v_0^\varepsilon$, testing by $u_{0\lambda}^\varepsilon$ and using the Young inequality we obtain

$$\lambda \|u_{0\lambda}^\varepsilon\|_V^2 + \varphi^\varepsilon(u_{0\lambda}^\varepsilon) + (\varphi^\varepsilon)^*(A^\varepsilon(u_{0\lambda}^\varepsilon)) = (v_0, u_{0\lambda}^\varepsilon)_H \leq \varphi^\varepsilon(u_{0\lambda}^\varepsilon) + (\varphi^\varepsilon)^*(v_0^\varepsilon),$$

where, by definition of φ^ε ,

$$(\varphi^\varepsilon)^*(v_0^\varepsilon) = \frac{\varepsilon}{2} \|v_0^\varepsilon\|_H^2 + \varphi^*(v_0^\varepsilon) \leq M(1 + \|v_0^\varepsilon\|_H^2) \leq M(1 + \|v_0\|_H^2).$$

Hence,

$$(\varphi_\lambda^\varepsilon)^*(v_0^\varepsilon) = \frac{\lambda}{2} \|u_{0\lambda}^\varepsilon\|_V^2 + (\varphi^\varepsilon)^*(A^\varepsilon(u_{0\lambda}^\varepsilon)) \leq (\varphi^\varepsilon)^*(v_0) \leq M(1 + \|v_0\|_H^2) \in L^{q/2}(\Omega).$$

Moreover, hypothesis (H6) and the Young inequality immediately yield

$$\int_0^t (F(s, u_\lambda^\varepsilon(s)), u_\lambda^\varepsilon(s))_H ds \leq \frac{1}{4} \|F(\cdot, 0)\|_{L^2(0, T; H)}^2 + (L_F^2 + 1) \int_0^t \|u_\lambda^\varepsilon(s)\|_H^2 ds,$$

while by definition of J_λ^B and coercivity of B we have

$$\begin{aligned} \langle B_\lambda u_\lambda^\varepsilon, u_\lambda^\varepsilon \rangle_V &= \langle B_\lambda u_\lambda^\varepsilon, J_\lambda^B u_\lambda^\varepsilon \rangle_V + \langle B_\lambda u_\lambda^\varepsilon, u_\lambda^\varepsilon - J_\lambda^B u_\lambda^\varepsilon \rangle_V \\ &= \langle B_\lambda u_\lambda^\varepsilon, J_\lambda^B u_\lambda^\varepsilon \rangle_V + \lambda \langle B_\lambda u_\lambda^\varepsilon, R^{-1} B_\lambda u_\lambda^\varepsilon \rangle_V \geq c_B \|J_\lambda^B u_\lambda^\varepsilon\|_V^2 + \lambda \|B_\lambda u_\lambda^\varepsilon\|_{V^*}^2. \end{aligned}$$

Furthermore, let $h \in H$ be arbitrary and set $k_\lambda^\varepsilon := (\lambda R + DA^\varepsilon(u_\lambda^\varepsilon))^{-1} h \in V$, so that

$$\lambda R k_\lambda^\varepsilon + DA^\varepsilon(u_\lambda^\varepsilon) k_\lambda^\varepsilon = h.$$

Testing by k_λ^ε and using the strong monotonicity of A^ε we have

$$\lambda \|k_\lambda^\varepsilon\|_V^2 + \frac{c_A}{2} \|k_\lambda^\varepsilon\|_H^2 \leq (h, k_\lambda^\varepsilon)_H \leq \frac{c_A}{4} \|k_\lambda^\varepsilon\|_H^2 + \frac{1}{c_A} \|h\|_H^2.$$

We infer that the following uniform estimate holds:

$$\|(\lambda R + DA^\varepsilon(u_\lambda^\varepsilon))^{-1}\|_{\mathcal{L}(H, H)}^2 \leq \frac{4}{c_A^2}. \quad (3)$$

Hence, the trace term on the right-hand side of Itô's formula can be estimated by

$$\text{Tr} \left(G^*(\cdot, u_\lambda^\varepsilon) (\lambda R + DA_{|V}^\varepsilon(u_\lambda^\varepsilon))^{-1} G(\cdot, u_\lambda^\varepsilon) \right) \leq \frac{2}{c_A} \|G(\cdot, u_\lambda^\varepsilon)\|_{\mathcal{L}^2(U, H)}^2 \leq \frac{4L_G^2}{c_A} \left(1 + \|u_\lambda^\varepsilon\|_H^2 \right).$$

Finally, by the Burkholder-Davis-Gundy and Young inequalities we have, for every $\delta > 0$,

$$\begin{aligned} \mathbb{E} \sup_{t \in [0, T]} \left| \int_0^t (u_\lambda^\varepsilon(s), G(s, u_\lambda^\varepsilon(s)))_H dW(s) \right|^{q/2} &\leq c \mathbb{E} \left(\int_0^T \|u_\lambda^\varepsilon(s)\|_H^2 \|G(s, u_\lambda^\varepsilon(s))\|_{\mathcal{L}^2(U, H)}^2 ds \right)^{q/4} \\ &\leq \mathbb{E} \left(\|u_\lambda^\varepsilon\|_{L^\infty(0, T; H)}^{q/2} \|G(\cdot, u_\lambda^\varepsilon)\|_{L^2(0, T; \mathcal{L}^2(U, H))}^{q/2} \right) \\ &\leq \delta \mathbb{E} \|u_\lambda^\varepsilon\|_{L^\infty(0, T; H)}^q + \frac{1}{4\delta} \|G(\cdot, u_\lambda^\varepsilon)\|_{L^2(0, T; \mathcal{L}^2(U, H))}^q \\ &\leq \delta \mathbb{E} \|u_\lambda^\varepsilon\|_{L^\infty(0, T; H)}^q + \frac{L_G^2}{\delta} \left(1 + \|u_\lambda^\varepsilon\|_{L^2(0, T; H)}^q \right), \end{aligned}$$

where c is a positive constant, depending on data only. Hence, recalling that A and B have linear growth, taking supremum in time, power $q/2$ and expectations in Itô's formula, choosing δ small enough, rearranging the terms, and employing the Gronwall lemma yield the first three desired estimates. Moreover, the fourth and fifth estimates easily follows from (H7) and the properties of the stochastic integral.

Let us show now the last two estimates: since $(B_\lambda u_\lambda^\varepsilon)_{\lambda, \varepsilon}$ is uniformly bounded in the space $L^q(\Omega; L^2(0, T; V^*))$, we have

$$\left\| \int_0^\cdot B_\lambda(u_\lambda^\varepsilon(s)) ds \right\|_{L^q(\Omega; H^1(0, T; V^*))} \leq M$$

for a positive constant M independent of λ and ε . Now, since $\eta \in (0, 1/2)$ and $q > 2$, we have that $1 - 1/2 > \eta - 1/q$, so the Sobolev embeddings imply that $H^1(0, T; V^*) \hookrightarrow W^{\eta, q}(0, T; V^*)$ continuously, hence by comparison in the equation we have

$$\|v_\lambda^\varepsilon\|_{L^q(\Omega; W^{\eta, q}(0, T; V^*))} \leq M.$$

Now, recalling that $v_\lambda^\varepsilon = \lambda R u_\lambda^\varepsilon + A^\varepsilon(u_\lambda^\varepsilon)$, we have

$$\mathbb{E} \int_0^T \int_0^T \frac{\|\lambda R(u_\lambda^\varepsilon(s) - u_\lambda^\varepsilon(r)) + A^\varepsilon(u_\lambda^\varepsilon)(s) - A^\varepsilon(u_\lambda^\varepsilon)(r)\|_{V^*}^q}{|s - r|^{1+\eta q}} ds dr \leq M.$$

Since for almost every $s, r \in (0, T)$ we have, by monotonicity of A^ε ,

$$\begin{aligned} \lambda^2 \|R(u_\lambda^\varepsilon(s) - u_\lambda^\varepsilon(r))\|_{V^*}^2 &= \langle \lambda R(u_\lambda^\varepsilon(s) - u_\lambda^\varepsilon(r)), \lambda(u_\lambda^\varepsilon(s) - u_\lambda^\varepsilon(r)) \rangle_{V^*} \\ &\leq \langle \lambda R(u_\lambda^\varepsilon(s) - u_\lambda^\varepsilon(r)) + A^\varepsilon(u_\lambda^\varepsilon)(s) - A^\varepsilon(u_\lambda^\varepsilon)(r), \lambda(u_\lambda^\varepsilon(s) - u_\lambda^\varepsilon(r)) \rangle_{V^*} \\ &\leq \frac{1}{2} \|\lambda R(u_\lambda^\varepsilon(s) - u_\lambda^\varepsilon(r)) + A^\varepsilon(u_\lambda^\varepsilon)(s) - A^\varepsilon(u_\lambda^\varepsilon)(r)\|_{V^*}^2 + \frac{\lambda^2}{2} \|R(u_\lambda^\varepsilon(s) - u_\lambda^\varepsilon(r))\|_{V^*}^2, \end{aligned}$$

we deduce that $(\lambda R u_\lambda^\varepsilon)_\lambda$ is uniformly bounded in the space $L^q(\Omega; W^{\eta, q}(0, T; V^*))$, hence also $A^\varepsilon(u_\lambda^\varepsilon) = v_\lambda^\varepsilon - \lambda R u_\lambda^\varepsilon$ by difference. \square

5.2. Passage to the limit as $\lambda \searrow 0$. In this section we perform the passage to the limit as $\lambda \searrow 0$, while $\varepsilon \in (0, c_A^{-1})$ is fixed. We shall divide the passage to the limit in several steps.

Stochastic compactness. Fix $\eta \in (1/q, 1/2)$, which is possible since $q > 2$. First of all, recalling Lemma 5.1, we have that the families $(A^\varepsilon(u_\lambda^\varepsilon))_\lambda$ and $(G(\cdot, u_\lambda^\varepsilon) \cdot W)_\lambda$ are uniformly bounded in the space

$$L^q(\Omega; W^{\eta, q}(0, T; V^*) \cap C^0([0, T]; H))$$

and the family $(W)_\lambda$ is constant in $C^0([0, T]; U)$. Since $H \xrightarrow{c} V^*$ compactly and $\eta q > 1$, by the classical compactness results by Aubin-Lions and Simon (see [49, Cor. 5, p. 86]) we have the

compact inclusion

$$W^{n,q}(0, T; H) \cap C^0([0, T]; H) \xhookrightarrow{c} C^0([0, T]; V^*).$$

This ensures by a standard argument based on the Markov inequality that the family of laws of $(A^\varepsilon(u_\lambda^\varepsilon))_\lambda$ and $(G(\cdot, u_\lambda^\varepsilon) \cdot W)_\lambda$ are tight on the space $C^0([0, T]; V^*)$. Secondly, it is clear that the laws of the constant sequences $(W)_\lambda$ and $(v_0^\varepsilon)_\lambda$ are tight on the spaces $C^0([0, T]; U)$ and H , respectively.

In particular, so is the family of laws of $(A^\varepsilon(u_\lambda^\varepsilon), G(\cdot, u_\lambda^\varepsilon) \cdot W, W, v_0^\varepsilon)_\lambda$ on the product space $C^0([0, T]; V^*) \times C^0([0, T]; V^*) \times C^0([0, T]; U) \times H$. By Skorokhod's theorem (see [31, Thm. 2.7]) there exist a probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$, a family $(\phi_\lambda)_\lambda$ of measurable mappings $\phi_\lambda : (\hat{\Omega}, \hat{\mathcal{F}}) \rightarrow (\Omega, \mathcal{F})$ such that

$$\mathbb{P} = \hat{\mathbb{P}} \circ \phi_\lambda^{-1} \quad \forall \lambda > 0,$$

and measurable random variables $\hat{v}^\varepsilon, \hat{I}^\varepsilon : (\hat{\Omega}, \hat{\mathcal{F}}) \rightarrow C^0([0, T]; V^*)$, $\hat{W}^\varepsilon : (\hat{\Omega}, \hat{\mathcal{F}}) \rightarrow C^0([0, T]; U)$ and $\hat{v}_0^\varepsilon : (\hat{\Omega}, \hat{\mathcal{F}}) \rightarrow H$ such that, setting $\hat{u}_\lambda^\varepsilon := u_\lambda^\varepsilon \circ \phi_\lambda$, as $\lambda \searrow 0$,

$$\begin{aligned} A^\varepsilon(\hat{u}_\lambda^\varepsilon) &\rightarrow \hat{v}^\varepsilon && \text{in } C^0([0, T]; V^*), && \hat{\mathbb{P}}\text{-a.s.} \\ \hat{I}_\lambda^\varepsilon := (G(\cdot, u_\lambda^\varepsilon) \cdot W) \circ \phi_\lambda &\rightarrow \hat{I}^\varepsilon && \text{in } C^0([0, T]; V^*), && \hat{\mathbb{P}}\text{-a.s.} \\ \hat{W}_\lambda &:= W \circ \phi_\lambda \rightarrow \hat{W}^\varepsilon && \text{in } C^0([0, T]; U), && \hat{\mathbb{P}}\text{-a.s.} \\ \hat{v}_{0,\lambda} &:= v_0^\varepsilon \circ \phi_\lambda \rightarrow \hat{v}_0^\varepsilon && \text{in } H, && \hat{\mathbb{P}}\text{-a.s.} \end{aligned}$$

Setting also $\hat{v}_\lambda^\varepsilon := v_\lambda^\varepsilon \circ \phi_\lambda$, since $\mathbb{P} = \hat{\mathbb{P}} \circ \phi_\lambda^{-1}$ the uniform estimates given by Lemma 5.1 are preserved on the space $\hat{\Omega}$ for $(\hat{u}_\lambda^\varepsilon)_\lambda$ and $(\hat{v}_\lambda^\varepsilon)_\lambda$. Consequently, there exist also two measurable random variables $\hat{u}^\varepsilon : (\hat{\Omega}, \hat{\mathcal{F}}) \rightarrow L^2(0, T; V)$ and $\hat{w}^\varepsilon : (\hat{\Omega}, \hat{\mathcal{F}}) \rightarrow L^2(0, T; V^*)$ such that, as $\lambda \searrow 0$,

$$\begin{aligned} \lambda \hat{u}_\lambda^\varepsilon &\rightarrow 0 && \text{in } L^q(\hat{\Omega}; C^0([0, T]; V)), \\ J_\lambda^B \hat{u}_\lambda^\varepsilon &\rightarrow \hat{u}^\varepsilon && \text{in } L^q(\hat{\Omega}; L^2(0, T; V)), \\ B_\lambda \hat{u}_\lambda^\varepsilon &\rightarrow \hat{w}^\varepsilon && \text{in } L^q(\hat{\Omega}; L^2(0, T; V^*)), \\ \hat{v}_\lambda^\varepsilon, A^\varepsilon(\hat{u}_\lambda^\varepsilon) &\rightarrow \hat{v}^\varepsilon && \text{in } L^p(\hat{\Omega}; C^0([0, T]; V^*)) \quad \forall p \in [1, q) \\ A^\varepsilon(\hat{u}_\lambda^\varepsilon) &\rightarrow \hat{v}^\varepsilon && \text{in } L^q(\hat{\Omega}; L^2(0, T; H)), \\ \hat{I}_\lambda^\varepsilon &\xrightarrow{*} \hat{I}^\varepsilon && \text{in } L^q(\hat{\Omega}; W^{1,\eta}(0, T; H)). \end{aligned}$$

Moreover, noting that $\hat{u}_\lambda^\varepsilon - J_\lambda \hat{u}_\lambda^\varepsilon = \lambda R^{-1} B_\lambda \hat{u}_\lambda^\varepsilon$, it is immediate to deduce also that

$$\hat{u}_\lambda^\varepsilon \rightarrow \hat{u}^\varepsilon \quad \text{in } L^q(\hat{\Omega}; L^2(0, T; V)).$$

Also, by lower semicontinuity of φ^* and Lemma 5.1 we deduce that $\varphi^*(\hat{v}^\varepsilon) \in L^{q/2}(\hat{\Omega}; L^\infty(0, T))$. Since (H1) implies that φ^* is coercive on H this immediately yields

$$\hat{v}^\varepsilon \in L^q(\hat{\Omega}; L^\infty(0, T; H)).$$

Moreover, since $A_{|V}^\varepsilon : V \rightarrow 2V^*$ is maximal monotone by assumption (H1), by strong-weak closure we immediately infer that

$$\hat{v}^\varepsilon = A^\varepsilon(\hat{u}^\varepsilon) \quad \text{a.e. in } \hat{\Omega} \times (0, T).$$

Now, noting that $\lambda R \hat{u}_\lambda^\varepsilon + A^\varepsilon(\hat{u}_\lambda^\varepsilon) = \hat{v}_\lambda^\varepsilon$, it is clear that

$$\lambda R \hat{u}_\lambda^\varepsilon + A^\varepsilon(\hat{u}_\lambda^\varepsilon) - \hat{v}^\varepsilon = \hat{v}_\lambda^\varepsilon - \hat{v}^\varepsilon,$$

hence testing by $\hat{u}_\lambda^\varepsilon$ and using that $\hat{v}^\varepsilon = A^\varepsilon(\hat{u}^\varepsilon)$ yields

$$\lambda \|\hat{u}_\lambda^\varepsilon\|_V^2 + (A^\varepsilon(\hat{u}_\lambda^\varepsilon) - \hat{v}^\varepsilon, \hat{u}_\lambda^\varepsilon)_H = (\hat{v}_\lambda^\varepsilon - \hat{v}^\varepsilon, \hat{u}_\lambda^\varepsilon)_H.$$

Rearranging the terms and employing the strong monotonicity of A^ε yields

$$\begin{aligned} \lambda \|\hat{u}_\lambda^\varepsilon\|_V^2 + \frac{cA}{2} \|\hat{u}_\lambda^\varepsilon - \hat{u}^\varepsilon\|_H^2 &\leq (\hat{v}_\lambda^\varepsilon - \hat{v}^\varepsilon, \hat{u}_\lambda^\varepsilon)_H - (A^\varepsilon(\hat{u}_\lambda^\varepsilon) - \hat{v}^\varepsilon, \hat{u}^\varepsilon)_H \\ &\leq \|A^\varepsilon(\hat{u}_\lambda^\varepsilon) - \hat{v}^\varepsilon\|_{V^*} \|\hat{u}_\lambda^\varepsilon\|_V + \|A^\varepsilon(\hat{u}_\lambda^\varepsilon) - \hat{v}^\varepsilon\|_{V^*} \|\hat{u}^\varepsilon\|_V. \end{aligned}$$

Since $(\hat{u}_\lambda^\varepsilon)_\lambda$ is bounded in $L^q(\hat{\Omega}; L^2(0, T; V))$ and $A^\varepsilon(\hat{u}_\lambda^\varepsilon) \rightarrow \hat{v}^\varepsilon$ and $\hat{v}_\lambda^\varepsilon \rightarrow \hat{v}^\varepsilon$ in $L^p(\hat{\Omega}; L^2(0, T; V^*))$ for $p \in [1, q)$, we deduce also that

$$\hat{u}_\lambda^\varepsilon \rightarrow \hat{u}^\varepsilon \quad \text{in } L^p(\hat{\Omega}; L^2(0, T; H)) \quad \forall p \in [1, q),$$

which implies by the Lipschitz-continuity of F and G that

$$\begin{aligned} F(\cdot, \hat{u}_\lambda^\varepsilon) &\rightarrow F(\cdot, \hat{u}^\varepsilon) \quad \text{in } L^p(\hat{\Omega}; L^2(0, T; H)) \quad \forall p \in [1, q), \\ G(\cdot, \hat{u}_\lambda^\varepsilon) &\rightarrow G(\cdot, \hat{u}^\varepsilon) \quad \text{in } L^p(\hat{\Omega}; L^2(0, T; \mathcal{L}^2(U, H))) \quad \forall p \in [1, q). \end{aligned}$$

Identification of the stochastic integral. By definition of $\hat{I}_\lambda^\varepsilon : (\hat{\Omega}, \hat{\mathcal{F}}) \rightarrow C^0([0, T]; V^*)$ we have that

$$\hat{I}_\lambda^\varepsilon = \hat{v}_\lambda^\varepsilon + \int_0^\cdot B_\lambda \hat{u}_\lambda^\varepsilon(s) ds - \hat{v}_{0, \lambda} - \int_0^\cdot F(s, \hat{u}_\lambda^\varepsilon(s)) ds.$$

By introducing the filtration $(\hat{\mathcal{F}}_{\lambda, t}^\varepsilon)_{t \in [0, T]}$ as

$$\hat{\mathcal{F}}_{\lambda, t}^\varepsilon := \sigma\{\hat{u}_\lambda^\varepsilon(s), \hat{I}_\lambda^\varepsilon(s), \hat{W}_\lambda(s) : s \leq t\}, \quad t \in [0, T],$$

one can show that $\hat{I}_\lambda^\varepsilon$ is a square integrable V^* -valued martingale with respect to the filtration $(\hat{\mathcal{F}}_{\lambda, t}^\varepsilon)_{t \in [0, T]}$ with quadratic variation process given by

$$\langle \hat{I}_\lambda^\varepsilon \rangle = \int_0^\cdot \|G(s, \hat{u}_\lambda^\varepsilon(s))\|_{\mathcal{L}^2(U, V^*)}^2 ds.$$

Indeed, for any $s, t \in [0, T]$ with $s \leq t$ and for any real bounded continuous function g on $C^0([0, T]; V^*)$, recalling that u_λ^ε and $\hat{u}_\lambda^\varepsilon$ have the same law, it is possible to see (for further details we refer here to [20] and [17, § 8.4]) that

$$\hat{\mathbb{E}} \left(\left\langle \hat{I}_\lambda^\varepsilon(t) - \hat{I}_\lambda^\varepsilon(s), z \right\rangle_V g(\hat{u}_{\lambda|_{[0, s]}}^\varepsilon) \right) = 0 \quad \forall z \in V$$

and

$$\begin{aligned} \hat{\mathbb{E}} \left[\left(\left\langle \hat{I}_\lambda^\varepsilon(t), z_1 \right\rangle_V \left\langle \hat{I}_\lambda^\varepsilon(t), z_2 \right\rangle_V - \left\langle \hat{I}_\lambda^\varepsilon(s), z_1 \right\rangle_V \left\langle \hat{I}_\lambda^\varepsilon(s), z_2 \right\rangle_V \right. \right. \\ \left. \left. - \int_s^t (G(r, \hat{u}_\lambda^\varepsilon(r))^* z_1, G(r, \hat{u}_\lambda^\varepsilon(r))^* z_2)_U dr \right) g(\hat{u}_{\lambda|_{[0, s]}}^\varepsilon) \right] = 0 \quad \forall z_1, z_2 \in V. \end{aligned}$$

Hence, thanks to a classical representation theorem for martingales (see [17, Thm. 8.2]), there exists a further probability space that we can identify with no restrictions with $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$, such that (possibly enlarging the filtration $(\hat{\mathcal{F}}_{\lambda, t}^\varepsilon)_{t \in [0, T]}$)

$$\hat{I}_\lambda^\varepsilon(t) = \int_0^t G(s, \hat{u}_\lambda^\varepsilon(s)) d\hat{W}_\lambda(s), \quad t \in [0, T].$$

Now, passing to the limit as $\lambda \searrow 0$ in the approximated equation we infer that

$$\hat{v}^\varepsilon + \int_0^\cdot \hat{w}^\varepsilon(s) ds = \hat{v}_0^\varepsilon + \int_0^\cdot F(s, \hat{u}^\varepsilon(s)) ds + \hat{I}^\varepsilon,$$

while the boundedness and continuity of g together with the convergences obtained above imply that

$$\hat{\mathbb{E}} \left(\left\langle \hat{I}^\varepsilon(t) - \hat{I}^\varepsilon(s), z \right\rangle_V g(\hat{u}_{|[0,s]^\varepsilon}^\varepsilon) \right) = 0 \quad \forall z \in V$$

and

$$\begin{aligned} \hat{\mathbb{E}} \left[\left(\left\langle \hat{I}^\varepsilon(t), z_1 \right\rangle_V \left\langle \hat{I}^\varepsilon(t), z_2 \right\rangle_V - \left\langle \hat{I}^\varepsilon(s), z_1 \right\rangle_V \left\langle \hat{I}^\varepsilon(s), z_2 \right\rangle_V \right. \right. \\ \left. \left. - \int_s^t (G(r, \hat{u}^\varepsilon(r))^* z_1, G(r, \hat{u}^\varepsilon(r))^* z_2)_U dr \right) g(\hat{u}_{|[0,s]^\varepsilon}^\varepsilon) \right] = 0 \quad \forall z_1, z_2 \in V. \end{aligned}$$

Using now the strong convergences of \hat{W}_λ to \hat{W}^ε , $G(\cdot, \hat{u}_\lambda^\varepsilon)$ to $G(\cdot, \hat{u}^\varepsilon)$, and $\hat{I}_\lambda^\varepsilon$ to \hat{I}^ε , following the approach contained in [45, § 4.5] (and the references therein) one can show that, by possibly enlarging the probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$, there exists a complete right-continuous filtration $(\hat{\mathcal{F}}_t^\varepsilon)_{t \in [0, T]}$ such that \hat{W}^ε is a cylindrical Wiener process adapted to $(\hat{\mathcal{F}}_t^\varepsilon)_{t \in [0, T]}$ and

$$\hat{I}^\varepsilon(t) = \int_0^t G(s, \hat{u}^\varepsilon(s)) d\hat{W}^\varepsilon(s), \quad t \in [0, T].$$

Consequently, we deduce that

$$\hat{v}^\varepsilon(t) + \int_0^t \hat{w}^\varepsilon(s) ds = \hat{v}_0^\varepsilon + \int_0^t F(s, \hat{u}^\varepsilon(s)) ds + \int_0^t G(s, \hat{u}^\varepsilon(s)) d\hat{W}^\varepsilon(s) \quad \forall t \in [0, T], \quad \hat{\mathbb{P}}\text{-a.s.}$$

Identification of the nonlinearities. We have already proved that $\hat{v}^\varepsilon = A^\varepsilon(\hat{u}^\varepsilon)$. Let us show now that $\hat{w}^\varepsilon \in B\hat{u}^\varepsilon$ almost everywhere in $\hat{\Omega} \times (0, T)$. To this end, recalling the proof of Lemma 5.1 we have that

$$\begin{aligned} \hat{\mathbb{E}}(\varphi_\lambda^\varepsilon)^*(\hat{v}_\lambda^\varepsilon(T)) + \hat{\mathbb{E}} \int_0^T \langle B_\lambda \hat{u}_\lambda^\varepsilon(s), \hat{u}_\lambda^\varepsilon(s) \rangle_V ds \\ = \hat{\mathbb{E}}(\varphi_\lambda^\varepsilon)^*(\hat{v}_0^\varepsilon) + \hat{\mathbb{E}} \int_0^T (F(s, \hat{u}_\lambda^\varepsilon(s)), \hat{u}_\lambda^\varepsilon(s))_H ds \\ + \frac{1}{2} \hat{\mathbb{E}} \int_0^T \text{Tr} (G^*(s, \hat{u}_\lambda^\varepsilon(s))(\lambda R + DA^\varepsilon(\hat{u}_\lambda^\varepsilon(s)))^{-1} G(s, \hat{u}_\lambda^\varepsilon(s))) ds, \end{aligned}$$

with obvious meaning of the symbol $\hat{\mathbb{E}}$. Now, note that

$$(\varphi_\lambda^\varepsilon)^*(\hat{v}_\lambda^\varepsilon(T)) = \frac{\lambda}{2} \|\hat{u}_\lambda^\varepsilon(T)\|_V^2 + (\varphi^\varepsilon)^*(A^\varepsilon(\hat{u}_\lambda^\varepsilon(T))) \geq (\varphi^\varepsilon)^*(A^\varepsilon(\hat{u}_\lambda^\varepsilon(T))),$$

while from the proof of Lemma 5.1 we also know that

$$(\varphi_\lambda^\varepsilon)^*(\hat{v}_0^\varepsilon) = \frac{\lambda}{2} \|\hat{u}_{0\lambda}^\varepsilon\|_V^2 + (\varphi^\varepsilon)^*(A^\varepsilon(\hat{u}_{0\lambda}^\varepsilon)) \leq (\varphi^\varepsilon)^*(\hat{v}_0^\varepsilon).$$

Moreover, since $A^\varepsilon(\hat{u}_\lambda^\varepsilon(T)) \rightharpoonup \hat{v}^\varepsilon(T)$ in $L^2(\Omega; H)$, by weak lower semicontinuity of $(\varphi^\varepsilon)^*$ we have that

$$\begin{aligned} & \limsup_{\lambda \searrow 0} \hat{\mathbb{E}} \int_0^T \langle B_\lambda \hat{u}_\lambda^\varepsilon(s), \hat{u}_\lambda^\varepsilon(s) \rangle_V ds \\ & \leq \hat{\mathbb{E}}(\varphi^\varepsilon)^*(\hat{v}_0^\varepsilon) - \liminf_{\lambda \searrow 0} \hat{\mathbb{E}}(\varphi^\varepsilon)^*(A^\varepsilon(\hat{u}_\lambda^\varepsilon(T))) + \lim_{\lambda \searrow 0} \hat{\mathbb{E}} \int_0^T (F(s, \hat{u}_\lambda^\varepsilon(s)), \hat{u}_\lambda^\varepsilon(s))_H ds \\ & \quad + \frac{1}{2} \limsup_{\lambda \searrow 0} \hat{\mathbb{E}} \int_0^T \text{Tr} (G^*(s, \hat{u}_\lambda^\varepsilon(s))(\lambda R + DA^\varepsilon(\hat{u}_\lambda^\varepsilon(s)))^{-1} G(s, \hat{u}_\lambda^\varepsilon(s))) ds \\ & \leq \hat{\mathbb{E}}(\varphi^\varepsilon)^*(\hat{v}_0^\varepsilon) - \hat{\mathbb{E}}(\varphi^\varepsilon)^*(\hat{v}^\varepsilon(T)) + \hat{\mathbb{E}} \int_0^T (F(s, \hat{u}^\varepsilon(s)), \hat{u}^\varepsilon(s))_H ds \\ & \quad + \frac{1}{2} \limsup_{\lambda \searrow 0} \hat{\mathbb{E}} \int_0^T \text{Tr} (G^*(s, \hat{u}_\lambda^\varepsilon(s))(\lambda R + DA^\varepsilon(\hat{u}_\lambda^\varepsilon(s)))^{-1} G(s, \hat{u}_\lambda^\varepsilon(s))) ds. \end{aligned}$$

Now, by Lemma 3.4 and the strong-weak convergence, we deduce that

$$\text{Tr} (G^*(\cdot, \hat{u}_\lambda^\varepsilon)(\lambda R + DA^\varepsilon(\hat{u}_\lambda^\varepsilon))^{-1} G(\cdot, \hat{u}_\lambda^\varepsilon)) \rightarrow \text{Tr} (G^*(\cdot, \hat{u}^\varepsilon) D((A^\varepsilon)^{-1})(\hat{v}^\varepsilon)) G(\cdot, \hat{u}^\varepsilon)$$

almost everywhere in $\hat{\Omega} \times (0, T)$. Moreover, we have already proved that

$$|\text{Tr} (G^*(\cdot, \hat{u}^\varepsilon) D((A^\varepsilon)^{-1})(\hat{v}^\varepsilon)) G(\cdot, \hat{u}^\varepsilon)| \leq c \left(1 + \|\hat{u}_\lambda^\varepsilon\|^2\right),$$

where the right hand side is bounded in $L^{q/2}(\Omega \times (0, T))$ and c is a positive constant depending on data. Since $q/2 > 1$, the right-hand side is uniformly integrable in $\Omega \times (0, T)$, hence so is the left-hand side, and Vitali's convergence theorem yields

$$\begin{aligned} & \limsup_{\lambda \searrow 0} \hat{\mathbb{E}} \int_0^T \langle B_\lambda \hat{u}_\lambda^\varepsilon(s), \hat{u}_\lambda^\varepsilon(s) \rangle_V ds \leq \hat{\mathbb{E}}(\varphi^\varepsilon)^*(\hat{v}_0^\varepsilon) - \hat{\mathbb{E}}(\varphi^\varepsilon)^*(\hat{v}^\varepsilon(T)) \\ & \quad + \hat{\mathbb{E}} \int_0^T (F(s, \hat{u}^\varepsilon(s)), \hat{u}^\varepsilon(s))_H ds + \frac{1}{2} \hat{\mathbb{E}} \int_0^T \text{Tr} (G^*(s, \hat{u}^\varepsilon(s)) D((A^\varepsilon)^{-1})(\hat{v}^\varepsilon(s)) G(s, \hat{u}^\varepsilon(s))) ds. \end{aligned}$$

Finally, by Proposition 4.2 it is immediate to see that this implies

$$\limsup_{\lambda \searrow 0} \hat{\mathbb{E}} \int_0^T \langle B_\lambda \hat{u}_\lambda^\varepsilon(s), \hat{u}_\lambda^\varepsilon(s) \rangle_V ds \leq \hat{\mathbb{E}} \int_0^T \langle \hat{w}^\varepsilon(s), \hat{u}^\varepsilon(s) \rangle_V ds.$$

Hence, we infer that $\hat{w}^\varepsilon \in B\hat{u}^\varepsilon$ a.e. in $\hat{\Omega} \times (0, T)$ by [16, Prop. 2.5, p. 27].

5.3. The passage to the limit as $\varepsilon \searrow 0$. In this last section we perform the passage to the limit as $\varepsilon \searrow 0$.

First of all, by Lemma 5.1, the convergences proved in the previous section and the weak lower semicontinuity of the norms, there exists a positive constant M , independent of ε , such

that

$$\begin{aligned}
& \varepsilon^{1/2} \|\hat{v}^\varepsilon\|_{L^q(\hat{\Omega}; L^\infty(0, T; H))} + \|\hat{v}^\varepsilon\|_{L^q(\Omega; W^{\eta, q}(0, T; V^*))} \leq M, \\
& \|\hat{u}^\varepsilon\|_{L^q(\Omega; L^\infty(0, T; H))} + \|\hat{v}^\varepsilon\|_{L^q(\Omega; L^\infty(0, T; H))} + \|\varphi^*(\hat{v}^\varepsilon)\|_{L^{q/2}(\hat{\Omega}; L^\infty(0, T))} \leq M, \\
& \|\hat{u}^\varepsilon\|_{L^q(\hat{\Omega}; L^2(0, T; V))} + \|\hat{w}^\varepsilon\|_{L^q(\Omega; L^2(0, T; V^*))} \leq M, \\
& \|G(\cdot, \hat{u}^\varepsilon)\|_{L^q(\Omega; C^0([0, T]; \mathcal{L}^2(U, H)))} + \left\| G(\cdot, \hat{u}^\varepsilon) \cdot \hat{W}^\varepsilon \right\|_{L^q(\Omega; W^{\eta, q}(0, T; H))} \leq M, \\
& \left\| \hat{v}^\varepsilon - G(\cdot, \hat{u}^\varepsilon) \cdot \hat{W}^\varepsilon \right\|_{L^q(\Omega; H^1(0, T; V^*))} \leq M,
\end{aligned}$$

Proceeding now as in the passage to the limit as $\lambda \searrow 0$ in the previous section, using Skorokhod's theorem and the usual representation theorems for martingales, we infer that there exist a further filtered probability space and a cylindrical Wiener process on it, which we shall assume with no restriction to coincide with $(\hat{\Omega}, \hat{\mathcal{F}}, (\hat{\mathcal{F}})_{t \in [0, T]}, \hat{\mathbb{P}})$, and \hat{W} , respectively, such that

$$\begin{aligned}
\hat{v}^\varepsilon &\rightarrow \hat{v} && \text{in } C^0([0, T]; V^*), \quad \mathbb{P}\text{-a.s.}, \\
\hat{W}^\varepsilon &\rightarrow \hat{W} && \text{in } C^0([0, T]; U), \quad \mathbb{P}\text{-a.s.}, \\
G(\cdot, \hat{u}^\varepsilon) \cdot \hat{W}^\varepsilon &\rightarrow \hat{I} && \text{in } C^0([0, T]; V^*), \quad \mathbb{P}\text{-a.s.}, \\
\hat{v}_0^\varepsilon &\rightarrow \hat{v}_0 && \text{in } H, \quad \mathbb{P}\text{-a.s.}
\end{aligned}$$

and

$$\begin{aligned}
\varepsilon \hat{u}^\varepsilon &\rightarrow 0 && \text{in } L^q(\hat{\Omega}; L^\infty(0, T; H)), \\
\hat{u}^\varepsilon &\rightarrow \hat{u} && \text{in } L^q(\hat{\Omega}; L^2(0, T; V)), \\
\hat{w}^\varepsilon &\rightarrow \hat{w} && \text{in } L^q(\hat{\Omega}; L^2(0, T; V^*)), \\
\hat{v}^\varepsilon &\rightarrow \hat{v} && \text{in } L^p(\Omega; L^2(0, T; V^*)) \quad \forall p \in [1, q), \\
\hat{v}^\varepsilon &\rightarrow \hat{v} && \text{in } L^q(\hat{\Omega}; L^2(0, T; H)),
\end{aligned}$$

Hence, the strong-weak closure of A readily implies that $\hat{v} \in A(\hat{u})$ almost everywhere in $\Omega \times (0, T)$. Furthermore, since $\hat{v}^\varepsilon = A^\varepsilon(\hat{u}^\varepsilon)$, we have that $\hat{u}^\varepsilon = \varepsilon \hat{v}^\varepsilon + A^{-1}(\hat{v}^\varepsilon)$, from which

$$\hat{v}^\varepsilon \in A(\hat{u}^\varepsilon - \varepsilon \hat{v}^\varepsilon) \quad \text{and} \quad \hat{v} \in A(\hat{u}).$$

Consequently, the strong monotonicity of A yields

$$\begin{aligned}
c_A \|\hat{u}^\varepsilon - \hat{u} - \varepsilon \hat{v}^\varepsilon\|_H^2 &\leq (\hat{v}^\varepsilon - \hat{v}, \hat{u}^\varepsilon - \hat{u} - \varepsilon \hat{v}^\varepsilon)_H \\
&\leq \varepsilon \|\hat{v}^\varepsilon - \hat{v}\|_H \|\hat{v}^\varepsilon\|_H + \|\hat{v}^\varepsilon - \hat{v}\|_{V^*} \|\hat{u}^\varepsilon - \hat{u}\|_V,
\end{aligned}$$

so that, integrating in time and recalling that $\varepsilon \in (0, c_A^{-1})$ we have

$$\begin{aligned}
\frac{c_A}{2} \|\hat{u}^\varepsilon - \hat{u}\|_{L^2(0, T; H)}^2 &\leq c_A \|\hat{u}^\varepsilon - \hat{u} - \varepsilon \hat{v}^\varepsilon\|_{L^2(0, T; H)}^2 + c_A \varepsilon^2 \|\hat{v}^\varepsilon\|_{L^2(0, T; H)}^2 \\
&\leq \varepsilon \|\hat{v}^\varepsilon - \hat{v}\|_{L^2(0, T; H)} \|\hat{v}^\varepsilon\|_{L^2(0, T; H)} + \|\hat{v}^\varepsilon - \hat{v}\|_{L^2(0, T; V^*)} \|\hat{u}^\varepsilon - \hat{u}\|_{L^2(0, T; V)} + \varepsilon \|\hat{v}^\varepsilon\|_{L^2(0, T; H)}.
\end{aligned}$$

For every $p \in [1, q)$, taking power $p/2$ at both sides and expectations, it follows from the Hölder inequality that

$$\begin{aligned}
\|\hat{u}^\varepsilon - \hat{u}\|_{L^p(\hat{\Omega}; L^2(0, T; H))}^p &\leq \left(\frac{2}{c_A} \right)^{p/2} \left(\varepsilon^{p/2} \|\hat{v}^\varepsilon - \hat{v}\|_{L^p(\hat{\Omega}; L^2(0, T; H))}^{p/2} \|\hat{v}^\varepsilon\|_{L^p(\hat{\Omega}; L^2(0, T; H))}^{p/2} \right. \\
&\quad \left. + \|\hat{v}^\varepsilon - \hat{v}\|_{L^p(\hat{\Omega}; L^2(0, T; V^*))}^{p/2} \|\hat{u}^\varepsilon - \hat{u}\|_{L^p(\hat{\Omega}; L^2(0, T; V))}^{p/2} + \varepsilon^{p/2} \|\hat{v}^\varepsilon\|_{L^p(\hat{\Omega}; L^2(0, T; H))}^{p/2} \right),
\end{aligned}$$

from which, taking into account the already proved estimates,

$$\|\hat{u}^\varepsilon - \hat{u}\|_{L^p(\hat{\Omega}; L^2(0, T; H))}^p \leq M_p \left(\varepsilon^{p/2} + \|\hat{v}^\varepsilon - \hat{v}\|_{L^p(\hat{\Omega}; L^2(0, T; V^*))}^{p/2} \right),$$

where M_p is a positive constant independent of ε . Thanks to the strong convergence of $(\hat{v}^\varepsilon)_\varepsilon$, we get

$$\hat{u}^\varepsilon \rightarrow \hat{u} \quad \text{in } L^p(\hat{\Omega}; L^2(0, T; H)) \quad \forall p \in [1, q),$$

from which

$$\begin{aligned} F(\cdot, \hat{u}^\varepsilon) &\rightarrow F(\cdot, \hat{u}) && \text{in } L^p(\hat{\Omega}; L^2(0, T; H)) \quad \forall p \in [1, q), \\ G(\cdot, \hat{u}^\varepsilon) &\rightarrow G(\cdot, \hat{u}) && \text{in } L^p(\hat{\Omega}; L^2(0, T; \mathcal{L}^2(U, H))) \quad \forall p \in [1, q), \end{aligned}$$

and arguing again as in the previous section, we have that

$$\hat{I} = G(\cdot, \hat{u}) \cdot \hat{W} \quad \text{in } C^0([0, T]; V^*), \quad \mathbb{P}\text{-a.s.}$$

In order to conclude, we only need to prove that $\hat{w} \in B(\hat{u})$ almost everywhere in $\Omega \times (0, T)$. To this end, we recall the Itô formula for A^ε from Proposition 4.2 and by definition of A^ε and have

$$\begin{aligned} &\frac{\varepsilon}{2} \hat{\mathbb{E}} \|v^\varepsilon(T)\|_H^2 + \hat{\mathbb{E}} \varphi^*(\hat{v}^\varepsilon(T)) + \hat{\mathbb{E}} \int_0^T \langle \hat{w}^\varepsilon(s), \hat{u}^\varepsilon(s) \rangle_V ds \\ &= \frac{\varepsilon}{2} \hat{\mathbb{E}} \|\hat{v}_0^\varepsilon\|_H^2 + \hat{\mathbb{E}} \varphi^*(\hat{v}_0^\varepsilon) + \hat{\mathbb{E}} \int_0^T (F(s, \hat{u}^\varepsilon(s)), \hat{u}^\varepsilon(s))_H ds \\ &+ \frac{\varepsilon}{2} \hat{\mathbb{E}} \int_0^T \|G(s, \hat{u}^\varepsilon(s))\|_{\mathcal{L}^2(U, H)}^2 ds + \frac{1}{2} \hat{\mathbb{E}} \int_0^T \text{Tr} (G^*(s, \hat{u}^\varepsilon(s))(D(A^{-1})(\hat{v}^\varepsilon(s)))G(s, \hat{u}^\varepsilon(s))) ds, \end{aligned}$$

which by lower semicontinuity, assumption (H4), and the Dominated Convergence Theorem, implies

$$\begin{aligned} \limsup_{\varepsilon \searrow 0} \hat{\mathbb{E}} \int_0^T \langle \hat{w}^\varepsilon(s), \hat{u}^\varepsilon(s) \rangle_V ds &\leq \hat{\mathbb{E}} \varphi^*(\hat{v}_0) - \hat{\mathbb{E}} \varphi^*(\hat{v}(T)) \\ &+ \mathbb{E} \int_0^T (F(s, \hat{u}(s)), \hat{u}(s))_H ds + \frac{1}{2} \hat{\mathbb{E}} \int_0^T \text{Tr} (G^*(s, \hat{u}(s))(D(A^{-1})(\hat{v}(s)))G(s, \hat{u}(s))) ds. \end{aligned}$$

We now use the Itô formula for A from Proposition 4.1 in order to check that

$$\limsup_{\varepsilon \searrow 0} \hat{\mathbb{E}} \int_0^T \langle \hat{w}^\varepsilon(s), \hat{u}^\varepsilon(s) \rangle_V ds \leq \hat{\mathbb{E}} \int_0^T \langle \hat{w}(s), \hat{u}(s) \rangle_V ds.$$

Owing to the latter, we conclude that $\hat{w} \in B(\hat{u})$ almost everywhere in $\Omega \times (0, T)$ by [16, Prop. 2.5, p. 27].

6. UNIQUENESS AND EXISTENCE OF STRONG SOLUTIONS: PROOF OF THEOREM 2.9

We begin by showing uniqueness of martingale solutions on the same probability space. Let (u_1, v_1, w_1) and (u_2, v_2, w_2) be two martingale solutions to the problem (2) on the same probability space. Then we have

$$d(v_1 - v_2) + (w_1 - w_2) dt = (F(\cdot, u_1) - F(\cdot, u_2)) dt + (G(\cdot, u_1) - G(\cdot, u_2)) dW, \quad (v_1 - v_2)(0) = 0.$$

If A is linear, continuous, and symmetric we have that

$$dA(u_1 - u_2) + (w_1 - w_2) dt = (F(\cdot, u_1) - F(\cdot, u_2)) dt + (G(\cdot, u_1) - G(\cdot, u_2)) dW,$$

so that Itô's formula and the Burkholder-Davis-Gundy inequality yield, for every $r \in [0, T]$,

$$\begin{aligned} & \frac{1}{2} \mathbb{E} \sup_{t \in [0, r]} (A(u_1 - u_2))(t), (u_1 - u_2)(t)_H + \int_0^r \langle (w_1 - w_2)(s), (u_1 - u_2)(s) \rangle ds \\ & \leq \mathbb{E} \int_0^r (F(s, u_1(s)) - F(s, u_2(s)), (u_1 - u_2)(s))_H ds \\ & \quad + c \mathbb{E} \left(\int_0^r \|G(s, u_1(s)) - G(s, u_2(s))\|_{\mathcal{L}^2(U, H)}^2 \|u_1 - u_2(s)\|_H^2 ds \right)^{1/2} \\ & \quad + \frac{1}{2} \int_0^r \text{Tr} (G^*(s, (u_1 - u_2)(s)) D(A^{-1})(A(u_1 - u_2)(s)) G(s, (u_1 - u_2)(s))) ds \end{aligned}$$

where c is a positive constant depending on data. Using now the Lipschitz-continuity of F and G , the boundedness of $D(A^{-1})$, and the Young inequality, we deduce that for every $\delta > 0$

$$\begin{aligned} & \mathbb{E} \sup_{t \in [0, r]} (A(u_1 - u_2))(t), (u_1 - u_2)(t)_H + \int_0^r \langle (w_1 - w_2)(s), (u_1 - u_2)(s) \rangle ds \\ & \leq \delta \mathbb{E} \sup_{t \in [0, r]} \|u_1 - u_2(t)\|_H^2 + C_\delta \mathbb{E} \int_0^r \|u_1 - u_2(s)\|_H^2 ds. \end{aligned}$$

The monotonicity of B , the strong monotonicity and linearity of A and the Gronwall Lemma imply that we can choose δ sufficiently small such that $\|u_1 - u_2(t)\|_H = 0$ for every $t \in [0, T]$. It follows that $u_1(t) = u_2(t)$ and $v_1(t) = v_2(t)$ for every $t \in [0, T]$, hence also $w_1 = w_2$ by comparison in the equation.

If B is linear continuous and symmetric, we have

$$\begin{aligned} (v_1 - v_2)(t) + B \int_0^t (u_1 - u_2)(s) ds &= \int_0^t (F(s, u_1(s)) - F(s, u_2(s))) ds \\ & \quad + \int_0^t (G(s, u_1(s)) - G(s, u_2(s))) dW(s) \end{aligned}$$

for every $t \in [0, T]$. Testing by $(u_1 - u_2)(t)$, further integrating in time, and using the Young inequality, yield, for all $r \in [0, T]$,

$$\begin{aligned} & \mathbb{E} \int_0^r ((v_1 - v_2)(t), (u_1 - u_2)(t))_H dt + \frac{1}{2} \mathbb{E} \left\langle B \int_0^r (u_1 - u_2)(s) ds, \int_0^r (u_1 - u_2)(s) ds \right\rangle_V \\ & \leq \mathbb{E} \int_0^r \left\| \int_0^t F(s, u_1(s)) - F(s, u_2(s)) ds \right\|_H \|u_1 - u_2(t)\|_H dt \\ & \quad + \mathbb{E} \int_0^r \left\| \int_0^t (G(s, u_1(s)) - G(s, u_2(s))) dW(s) \right\|_H \|u_1 - u_2(t)\|_H dt \\ & \leq \frac{c_A}{2} \int_0^r \|u_1 - u_2(t)\|_H^2 dt + \frac{1}{c_A} \mathbb{E} \int_0^r \left\| \int_0^t F(s, u_1(s)) - F(s, u_2(s)) ds \right\|_H^2 dt \\ & \quad + \frac{1}{c_A} \mathbb{E} \int_0^r \left\| \int_0^t (G(s, u_1(s)) - G(s, u_2(s))) dW(s) \right\|_H^2 dt. \end{aligned}$$

Hence, the strong monotonicity of A , the monotonicity of B , and the Lipschitz-continuity of F and G imply that

$$\frac{c_A}{2} \int_0^r \|u_1 - u_2(t)\|_H^2 dt \leq \frac{L_F^2 + L_G^2}{c_A} \int_0^r \int_0^t \|u_1 - u_2(s)\|_H^2 ds dt,$$

so that $u_1 = u_2$ on $[0, T]$ by the Gronwall Lemma, hence also $w_1 = w_2$ by linearity. By comparison in the equation, also using the fact that $(v_1 - v_2)(0) = 0$ it follows that $v_1 = v_2$.

Finally, a classical argument shows that uniqueness of martingale solutions on the same probability space yields also existence (hence uniqueness) of a strong solution. Indeed, this follows by a direct application of the following lemma, due to GYÖNGY & KRYLOV [27, Lem. 1.1].

Lemma 6.1. *Let \mathcal{X} be a Polish space and $(Z_n)_n$ be a sequence of \mathcal{X} -valued random variables. Then $(Z_n)_n$ converges in probability if and only if for any pair of subsequences $(Z_{n_k})_k$ and $(Z_{n_j})_j$, there exists a joint sub-subsequence $(Z_{n_{k_\ell}}, Z_{n_{j_\ell}})_{\ell}$ converging in law to a probability measure ν on $\mathcal{X} \times \mathcal{X}$ such that $\nu(\{(z_1, z_2) \in \mathcal{X} \times \mathcal{X} : z_1 = z_2\}) = 1$.*

Going back to the proofs of Theorem 2.8, it is not difficult to check that the Skorokhod theorem and the uniqueness of the limit problem yield exactly the condition of the lemma above: see for example [53, § 5]. Hence, one can recover strong convergences of the approximating sequences $(v_\lambda^\varepsilon)_\lambda$ and $(v^\varepsilon)_\varepsilon$ in $C^0([0, T]; V^*)$ in probability also on the original probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The conclusion follows then by the same arguments on the space $(\Omega, \mathcal{F}, \mathbb{P})$, instead of $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$.

7. APPLICATIONS

In this last section we present applications of our abstract existence theory to nonlinear SPDE problems. In particular, we will provide an existence result for martingale solutions for relation (1), when complemented with initial and boundary conditions.

7.1. Doubly-nonlinear SPDEs with multivalued graphs. In the following, let $\mathcal{O} \subset \mathbb{R}^d$ ($d \in \mathbb{N}$) be a nonempty, open, bounded, and connected domain, with Lipschitz-continuous boundary Γ . For definiteness we shall let

$$V = H_0^1(\mathcal{O}), \quad H = L^2(\mathcal{O}), \quad V_0 = H^2(\mathcal{O}) \cap H_0^1(\mathcal{O}).$$

which corresponds to consider homogeneous Dirichlet conditions. Note however that other classes of boundary conditions, including Neumann, Robin, and mixed, also of nonlinear type, may be considered as well, at the expense of minor notational modifications. Moreover, we let

$$Y = L^r(\mathcal{O}), \quad \text{with } \begin{cases} 2 < r \leq \frac{2d}{d-2} & \text{if } d \geq 3, \\ 2 < r < +\infty & \text{if } d = 1, 2. \end{cases}$$

This choice ensures that $V \subset Y \subset H$ densely and continuously.

The function $f : (0, T) \times \mathcal{O} \times \mathbb{R} \rightarrow \mathbb{R}$ is assumed to be of Carathéodory type with $f(\cdot, \cdot, u) \in L^2((0, T) \times \mathcal{O})$ for all $u \in \mathbb{R}$ and Lipschitz continuous in the variable u , uniformly in $(0, T) \times \mathcal{O}$. In particular, by defining the Nemytzky operator $F(t, u)(x) = f(t, x, u(x))$, for a.e. $x \in \mathcal{O}$, we have that (H6) follows. Eventually, we assume to be given another separable Hilbert space U and require that the operator-valued function $G : (0, T) \times L^2(\mathcal{O}) \rightarrow \mathcal{L}^2(U, L^2(\mathcal{O}))$ satisfies assumption (H7).

We now turn to the specification of classes of operators A which can be treated in our framework. Let $\hat{\alpha} : \mathbb{R} \rightarrow \mathbb{R}$ be convex, define $\alpha = \partial \hat{\alpha} : \mathbb{R} \rightarrow 2^{\mathbb{R}}$ and assume that $\alpha^{-1} \in C^1(\mathbb{R})$,

$0 \in \alpha(0)$, and there exists $c_\alpha, C_\alpha > 0$ such that

$$\begin{aligned} (v_1 - v_2)(u_1 - u_2) &\geq c_\alpha |u_1 - u_2|^2 \quad \forall u_i \in \mathbb{R}, v_i \in \alpha(u_i), \\ |v| &\leq C_\alpha(1 + |u|) \quad \forall u \in \mathbb{R}, v \in \alpha(u). \end{aligned}$$

Let us show that these positions entail the structural assumptions (H1)-(H4).

First of all, from the linear growth of α one has that $\hat{\alpha}(u) \leq C(1 + |u|^2)$ for some positive $C > 0$. By defining $\varphi : H \rightarrow [0, \infty]$ as

$$\varphi(u) = \int_{\mathcal{O}} \hat{\alpha}(u(x)) dx$$

one has φ is convex, everywhere defined (hence proper), lower semicontinuous, and $A = \partial\varphi : H \rightarrow 2^H$ is maximal monotone with $0 \in A(0)$. The strong monotonicity and the sublinearity A follow from those of α with the choices $C_A = \sqrt{2}C_\alpha \min\{|\mathcal{O}|^{1/2}, 1\}$ and $c_A = c_\alpha$. In particular, we have that $D(A) = H \supset V$. Secondly, the ε -Yosida regularization A^ε is given by $u \in H \mapsto \alpha^\varepsilon(u) = (u - (\text{Id} + \varepsilon\alpha)^{-1}(u))/\varepsilon \in H$. Note that $A^\varepsilon(V) \subset V$ as α^ε is $(1/\varepsilon)$ -Lipschitz continuous. In particular, we have that

$$(A^\varepsilon(u), R(u))_H = \langle R(u), A^\varepsilon(u) \rangle = \int_{\mathcal{O}} \nabla u \cdot \nabla(\alpha^\varepsilon(u)) dx = \int_{\mathcal{O}} (\alpha^\varepsilon)'(u) |\nabla u|^2 dx \geq 0 \quad \forall u \in V_0,$$

so that assumptions (H1)–(H2) are satisfied.

Let us check (H3). The operator $A^{-1} : H \rightarrow H$ is Lipschitz-continuous and $A^{-1}(v) = \alpha^{-1}(v)$, $v \in H$, a.e. in \mathcal{O} . As α^{-1} is C^1 , A^{-1} is differentiable and its differential $D(A^{-1}) : H \rightarrow \mathcal{L}(H, H)$ is given by

$$D(A^{-1})(v)h = \gamma(v)h \quad \text{a.e. in } \mathcal{O}, \quad \text{for } \gamma := (\alpha^{-1})' \in C(\mathbb{R}) \cap L^\infty(\mathbb{R}).$$

Let now $v_n \rightarrow v$ in H . Then $\gamma(v_n) \rightarrow \gamma(v)$ in $L^r(\mathcal{O})$ for all $r < \infty$ and $\gamma(v_n) \xrightarrow{*} \gamma(v)$ in $L^\infty(\mathcal{O})$. We then have that

$$(D(A^{-1})(v_n)h_1, h_2)_H = \int_{\mathcal{O}} \gamma(v_n)h_1h_2 dx \rightarrow \int_{\mathcal{O}} \gamma(v)h_1h_2 dx = (D(A^{-1})(v)h_1, h_2)_H,$$

so that $D(A^{-1}) \in C^0(H; \mathcal{L}_w(H, H))$. Since $V \subset Y = L^r(\mathcal{O}) \subset H$, one can compute that

$$\begin{aligned} \|D(A^{-1})(v_n) - D(A^{-1})(v)\|_{\mathcal{L}(Y, H)} &= \sup_{\|h\|_Y=1} \|D(A^{-1})(v_n)h - D(A^{-1})(v)h\|_H \\ &\leq \|\gamma(v_n) - \gamma(v)\|_{L^{2r/(r-2)}(\mathcal{O})} \rightarrow 0. \end{aligned}$$

We have hence checked that $D(A^{-1}) \in C^0(H; \mathcal{L}(Y, H))$ as well.

Let now $x, v \in V$ be given and define $y = (I + D(A^{-1}))((I + A^{-1})^{-1}x)^{-1}v$, namely

$$y + D(A^{-1})((I + A^{-1})^{-1}x)^{-1}y = v.$$

Test this equation on $|y|^{r-2}y$ and integrate on \mathcal{O} in order to get

$$\|y\|_Y^r + \int_{\mathcal{O}} \gamma((I + A^{-1})^{-1}x)|y|^r dx \leq \int_{\mathcal{O}} |v| |y|^{r-1} dx \leq \frac{1}{r} \|v\|_Y^r + \frac{r}{r-1} \|y\|_Y^r$$

where we used the Young inequality. As $V \subset Y$ we conclude that $\|y\|_Y \leq M\|v\|_V$ for $M > 0$, as required.

We are hence left with checking (H4). Let then $u_\varepsilon \rightharpoonup u$ in V and $A^\varepsilon(u_\varepsilon) \rightharpoonup v$ in H with $v \in A(u)$: we have to show that

$$\gamma'(\alpha^\varepsilon(u_\varepsilon))h \rightarrow \gamma'(v)h \quad \text{in } H \quad \forall h \in H. \quad (4)$$

Assume at first that α is continuous, hence single-valued. Then $A^\varepsilon(u_\varepsilon) = A(J_\varepsilon^\alpha u_\varepsilon) = \alpha(J_\varepsilon^\alpha u_\varepsilon) \rightarrow \alpha(u) = v$ a.e. in \mathcal{O} . In particular, for all $h \in H$ we have that

$$D(A^{-1})(A^\varepsilon(u_\varepsilon))h = \gamma'(\alpha(J_\varepsilon^\alpha u_\varepsilon))h \rightarrow \gamma'(v)h = D(A^{-1})(v)h \quad \text{in } H$$

by dominated convergence and (4) follows. Note that we need no Lipschitz-continuity here. In fact, the Lipschitz-continuous case has been indeed discussed in [45].

The argument can however be extended to include the case of non-single-valued graphs. For the sake of definiteness, let $\alpha = s + \tilde{\alpha}$, where $\tilde{\alpha}$ is monotone and continuous and s is the *sign* graph $s(x) = x/|x|$ for $x \neq 0$ and $s(0) = [-1, 1]$. Let us start by checking that $\gamma'(\alpha^\varepsilon(u_\varepsilon)) \rightarrow \gamma'(v)$ a.e. in \mathcal{O} . In order to prove this, we use the notation 1_A for the indicator function of the measurable set $A \subset \mathcal{O}$, rewrite

$$\gamma'(\alpha^\varepsilon(u_\varepsilon)) = \gamma'(\alpha^\varepsilon(u_\varepsilon))1_{\{u>0\}} + \gamma'(\alpha^\varepsilon(u_\varepsilon))1_{\{u<0\}} + \gamma'(\alpha^\varepsilon(u_\varepsilon))1_{\{u=0\}},$$

and discuss each term of this sum separately. As $J_\varepsilon^\alpha u_\varepsilon \rightarrow u$ a.e. in \mathcal{O} , we have that

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \gamma'(\alpha^\varepsilon(u_\varepsilon))1_{\{u>0\}} &= \lim_{\varepsilon \rightarrow 0} \gamma'(\alpha(J_\varepsilon^\alpha u_\varepsilon))1_{\{u>0\}} = \lim_{\varepsilon \rightarrow 0} \gamma'(1 + \tilde{\alpha}(J_\varepsilon^\alpha u_\varepsilon))1_{\{u>0\}} \\ &= \gamma'(1 + \tilde{\alpha}(u))1_{\{u>0\}} = \gamma'(v)1_{\{u>0\}} \quad \text{a.e. in } \mathcal{O}. \end{aligned}$$

Analogously, $\gamma'(\alpha^\varepsilon(u_\varepsilon))1_{\{u<0\}} \rightarrow \gamma'(v)1_{\{u<0\}}$ a.e. in \mathcal{O} . We now show that the remaining term on the set $\{u = 0\}$ is infinitesimal, namely $\gamma'(\alpha^\varepsilon(u_\varepsilon))1_{\{u=0\}} \rightarrow 0$ a.e. in \mathcal{O} . Indeed, we can further decompose it as

$$\gamma'(\alpha^\varepsilon(u_\varepsilon))1_{\{u=0\}} = \gamma'(\alpha^\varepsilon(u_\varepsilon))1_{\{u=0, |u_\varepsilon| \leq \varepsilon\}} + \gamma'(\alpha^\varepsilon(u_\varepsilon))1_{\{u=0, u_\varepsilon > \varepsilon\}} + \gamma'(\alpha^\varepsilon(u_\varepsilon))1_{\{u=0, u_\varepsilon < -\varepsilon\}}.$$

Since $\alpha^\varepsilon(r) = r/\varepsilon$ for $|r| \leq \varepsilon$ and $\gamma' = 0$ on $[-1, 1]$, the first term in the above right-hand side vanishes. As for the remaining terms we argue as follows

$$\begin{aligned} &\lim_{\varepsilon \rightarrow 0} (\gamma'(\alpha^\varepsilon(u_\varepsilon))1_{\{u=0, u_\varepsilon > \varepsilon\}} + \gamma'(\alpha^\varepsilon(u_\varepsilon))1_{\{u=0, u_\varepsilon < -\varepsilon\}}) \\ &= \lim_{\varepsilon \rightarrow 0} (\gamma'(1 + \tilde{\alpha}(J_\varepsilon^\alpha u_\varepsilon))1_{\{u=0, u_\varepsilon > \varepsilon\}} + \gamma'(-1 + \tilde{\alpha}(J_\varepsilon^\alpha u_\varepsilon))1_{\{u=0, u_\varepsilon < -\varepsilon\}}) \\ &\leq \lim_{\varepsilon \rightarrow 0} (\gamma'(1 + \tilde{\alpha}(J_\varepsilon^\alpha u_\varepsilon)) + \gamma'(-1 + \tilde{\alpha}(J_\varepsilon^\alpha u_\varepsilon)))1_{\{u=0\}} \\ &= \gamma'(1 + \tilde{\alpha}(0)) + \gamma'(-1 + \tilde{\alpha}(0)) = \gamma'(1) + \gamma'(-1) = 0. \end{aligned}$$

We have hence proved that

$$\gamma'(\alpha^\varepsilon(u_\varepsilon)) \rightarrow \left\{ \begin{array}{ll} \gamma'(v) & \text{on } \{u \neq 0\} \\ 0 & \text{on } \{u = 0\} \end{array} \right\} \stackrel{v \in \alpha(u)}{\equiv} \gamma'(v) \quad \text{a.e. in } \mathcal{O}.$$

Since γ' is bounded, dominated convergence entails (4), so that (H4) follows.

As for initial conditions, we require $v_0 \in L^q(\Omega, \mathcal{F}_0; L^2(\mathcal{O}))$ with $\hat{\alpha}^*(v_0) \in L^{q/2}(\Omega, \mathcal{F}_0; L^1(\mathcal{O}))$, and $u_0 := \alpha^{-1}(v_0) \in L^q(\Omega, \mathcal{F}_0; H_0^1(\mathcal{O}))$, which is noting but (H8).

Let us now present a class of operators B fitting our frame. Assume to be given $\beta_1 : \mathbb{R}^d \rightarrow 2^{\mathbb{R}^d}$ and $\beta_0 : \mathbb{R} \rightarrow 2^{\mathbb{R}}$ maximal, monotone, and linearly bounded. Note that β_1 is not required to be cyclic monotone. Moreover, we assume β_1 to be coercive, namely

$$\exists c_\beta > 0 : \quad c_\beta |\xi|^2 \leq \eta \cdot \xi \quad \forall \xi \in \mathbb{R}^d, \eta \in \beta_1(\xi).$$

We define $B : H_0^1(\mathcal{O}) \rightarrow 2^{H^{-1}(\mathcal{O})}$ by letting $w \in B(u)$ iff there exist $\xi \in L^2(\mathcal{O}, \mathbb{R}^d)$ and $b \in L^2(\mathcal{O})$ with $\xi \in \beta_1(\nabla u)$ and $b \in \beta_0(u)$ a.e. such that

$$\langle w, v \rangle = \int_{\mathcal{O}} \xi \cdot \nabla v \, dx + \int_{\mathcal{O}} b v \, dx \quad \forall v \in H_0^1(\mathcal{O}).$$

It is a standard matter to check that B is actually defined in all of $H_0^1(\mathcal{O})$, is maximal monotone, linearly bounded, and coercive. In particular, (H5) holds.

Along with these positions, the abstract relation (2) corresponds to the variational formulation of (1) under homogeneous Dirichlet boundary conditions, namely

$$d(\alpha(u)) - \operatorname{div} \beta_1(\nabla u) dt + \beta_0(u) dt \ni f(u) dt + G(u) dW \quad \text{in } H^{-1}(\mathcal{O}), \text{ a.e. in } \Omega \times (0, T) \quad (5)$$

$$u = 0 \quad \text{on } \partial\mathcal{O}, \text{ a.e. in } \Omega \times (0, T), \quad (6)$$

$$\alpha(u)(0) = v_0 \quad \text{in } H^{-1}(\mathcal{O}), \text{ a.e. in } \Omega. \quad (7)$$

Under the above assumptions, a direct application Theorem 2.8 entails the existence of a martingale solution to (5)-(7).

Theorem 7.1 (Existence of a martingale solution to (5)-(7)). *There exists a quintuplet*

$$((\hat{\Omega}, \hat{\mathcal{F}}, (\hat{\mathcal{F}}_t)_{t \in [0, T]}, \hat{\mathbb{P}}), \hat{W}, \hat{u}, \hat{v}, \hat{w}),$$

where $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$ is a probability space endowed with a filtration $(\hat{\mathcal{F}}_t)_{t \in [0, T]}$ which is saturated and right-continuous, \hat{W} is a cylindrical Wiener process, and \hat{u} , \hat{v} , and $\hat{\xi}$ are progressively measurable processes with values in $H_0^1(\mathcal{O})$, $L^2(\mathcal{O})$, and $L^2(\mathcal{O}; \mathbb{R}^d)$, respectively, such that

$$\begin{aligned} \hat{u} &\in L^q(\hat{\Omega}; L^\infty(0, T; L^2(\mathcal{O})) \cap L^2(0, T; H_0^1(\mathcal{O}))), \\ \hat{v} &\in L^q(\hat{\Omega}; L^\infty(0, T; L^2(\mathcal{O})) \cap C^0([0, T]; H^{-1}(\mathcal{O}))), \\ \hat{\xi} &\in L^q(\hat{\Omega}; L^2(0, T; L^2(\mathcal{O}; \mathbb{R}^d))), \\ \hat{v} &\in \alpha(\hat{u}), \quad \hat{\xi} \in \beta(\nabla \hat{u}) \quad \text{a.e. in } \hat{\Omega} \times \mathcal{O} \times (0, T), \\ \hat{v}(t) - \int_0^t \operatorname{div} \hat{\xi}(s) ds &= \hat{v}(0) + \int_0^t f(\hat{u}(s)) ds + \int_0^t G(s, \hat{u}(s)) d\hat{W}(s) \\ &\text{in } H^{-1}(\mathcal{O}) \quad \forall t \in [0, T], \hat{\mathbb{P}}\text{-a.s.} \end{aligned}$$

and $\hat{v}(0)$ has the same law of v_0 on $H^{-1}(\mathcal{O})$.

7.2. Stochastic equations with nonlocal terms. Before closing this discussion let us mention that the abstract existence result of Theorem 2.8 applies to other classes of SPDEs as well. As regards A , one could consider some linear operators of positive type, even nonlocal in space. An example in this direction is $A = \partial\varphi$ for

$$\varphi(u) = \frac{cA}{2} \int_{\mathcal{O}} |u|^2 dx + \frac{1}{2} \iint_{\mathcal{O} \times \mathcal{O}} k(x, y) (u(x) - u(y))^2 dx dy$$

where $c > 0$ and $k \in L^2(\mathcal{O} \times \mathcal{O})$ nonnegative and symmetric. In particular, φ is convex and lower semicontinuous and $Au(x) = \partial\varphi(u)(x) = cu(x) + \int_{\mathcal{O}} k(x, y) u(y) dy$ for a.e. $x \in \mathcal{O}$. Assumptions (H1)–(H3) are hence easy to check and assumption (H4) would actually be not needed here, for A is already Lipschitz-continuous.

Nonlocal operators B could also be considered. A relevant example in this direction is the fractional laplacian $(-\Delta)^r$ with $r \in (0, 1)$. When with homogeneous Dirichlet boundary conditions in $\mathbb{R}^d \setminus \mathcal{O}$, this can be variationally formulated by letting $H = \{v \in L^2(\mathbb{R}^d) \mid v = 0 \text{ a.e. in } \mathbb{R}^d \setminus \mathcal{O}\}$ and defining the fractional Sobolev space [19]

$$V := H_0^r(\mathcal{O}) = \{v \in H \mid [v]_{H^r(\mathcal{O})} < \infty\}$$

where the *Gagliardo seminorm* $[v]_r$ reads

$$[v]_{H^r(\mathcal{O})} := \left(\iint_{\mathcal{O} \times \mathcal{O}} \frac{|v(x) - v(y)|^2}{|x - y|^{d+2r}} dx dy \right)^{1/2}.$$

When endowed with the norm $\|v\|_V^2 := \|v\|_{L^2(\mathbb{R}^d)}^2 + [v]_{H^r(\mathcal{O})}^2$, the Hilbert space $V \subset H$ densely and continuously. The subdifferential $B = \partial[\cdot]_{H^r(\mathcal{O})}^2 : V \rightarrow V^*$ is hence linear, positive, and continuous and delivers a weak formulation of the fractional laplacian $(-\Delta)^r$, up to a multiplicative dimensional constant [3].

7.3. Systems of doubly-nonlinear SPDEs. Eventually, as B is not required to be cyclic monotone, one can tackle some classes of SPDE systems as well. An example in this direction is

$$d(\vec{\alpha}(\vec{u})) - \text{Div } b_1(\nabla \vec{u}) dt + \vec{b}_0(\vec{u}) dt \ni \vec{F}(\vec{u}) dt + \vec{G}(\vec{u}) d\vec{W} \quad \text{in } (H^{-1}(\mathcal{O}))^n, \text{ a.e. in } \Omega \times (0, T) \quad (8)$$

where $\vec{u} = (u_1, \dots, u_n) : \Omega \times [0, T] \times \mathcal{O} \rightarrow \mathbb{R}^n$, now with $\vec{u}(t) \in V = (H_0^1(\mathcal{O}))^n$ a.e. in \mathcal{O} . The maximal monotone graph $\vec{\alpha} : \mathbb{R}^n \rightarrow 2^{\mathbb{R}^n}$ is assumed to be diagonal, namely $\vec{\alpha}(\vec{u}) = \text{diag}(\alpha_1(u_1), \dots, \alpha_n(u_n))$, with all $\alpha_i : \mathbb{R} \rightarrow 2^{\mathbb{R}}$ being of the type discussed above. The graphs $b_1 : \mathbb{R}^{n \times d} \rightarrow 2^{\mathbb{R}^{n \times d}}$ and $\vec{b}_0 : \mathbb{R}^n \rightarrow 2^{\mathbb{R}^n}$ are maximal monotone, possibly noncyclic, linearly bounded and b_1 is coercive. Note that in the vectorial case Div is the standard tensorial divergence, namely $(\text{Div } b)_i = \sum_{j=1}^d \partial_j b_{ij}$. We define the operator $B : V \rightarrow 2^{V^*}$ as $w \in B(\vec{u})$ iff $\xi \in (L^2(\mathcal{O}))^{n \times d}$ and $\vec{b} \in (L^2(\mathcal{O}))^n$ exist such that $\xi \in b_1(\nabla \vec{u})$, $\vec{b} \in \vec{b}_0(\vec{u})$ a.e. in \mathcal{O} , and

$$\langle w, \vec{z} \rangle = \sum_{i=1}^n \int_{\mathcal{O}} \xi_i \cdot \nabla z_i dx + \int_{\mathcal{O}} \vec{b} \cdot \vec{z} dx \quad \forall z \in (H_0^1(\mathcal{O}))^n$$

and remark that it fulfills (H5) by not being cyclic monotone. Asking $\vec{u} \mapsto \vec{F}(\vec{u})$ to be Lipschitz continuous, and \vec{G} and \vec{W} to be corresponding vectorial versions of operator-valued coefficients and cylindrical Wiener processes, the initial-value problem for the SPDE system (8) can be variationally reformulated as relation (2) and the abstract Theorem 2.8 provides the existence of martingale solutions.

ACKNOWLEDGEMENT

This work has been funded by the Vienna Science and Technology Fund (WWTF) through Project MA14-009. The support by the Austrian Science Fund (FWF) projects F 65 and I 2375 is also gratefully acknowledged.

REFERENCES

- [1] G. Akagi. Doubly nonlinear evolution equations governed by time-dependent subdifferentials in reflexive Banach space. *J. Differential Equations*, 231:32–56, 2006.
- [2] G. Akagi. Local solvability of a fully nonlinear parabolic equation. *Kodai Math. J.*, 37:702–727, 2014.
- [3] G. Akagi, G. Schimperna, and A. Segatti. Convergence of solutions for the fractional Cahn-Hilliard system. *J. Funct. Anal.*, 276(9):2663–2715, 2019.
- [4] S. Aizicovici and V.-M. Hokkanen. Doubly nonlinear equations with unbounded operators. *Nonlinear Anal.*, 58:591–607, 2004.
- [5] S. Aizicovici and V.-M. Hokkanen. Doubly nonlinear periodic problems with unbounded operators. *J. Math. Anal. Appl.*, 292:540–557, 2004.

- [6] G. Akagi and U. Stefanelli. Doubly nonlinear evolution equations as convex minimization. *SIAM J. Math. Anal.*, 46:1922–1945, 2014.
- [7] H. W. Alt and S. Luckhaus. Quasilinear elliptic-parabolic differential equations. *Math. Z.*, 183:311–341, 1983.
- [8] V. Barbu. Existence theorems for a class of two point boundary problems. *J. Differential Equations*, 17:236–257, 1975.
- [9] V. Barbu, S. Bonaccorsi, L. Tubaro. Existence and asymptotic behavior for hereditary stochastic evolution equations. *Appl. Math. Optim.*, 69(2):273–314, 2014.
- [10] V. Barbu and G. Da Prato. The two phase stochastic Stefan problem. *Probab. Theory Related Fields*, 124(4):544–560, 2002.
- [11] 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.
- [12] V. Barbu and M. Röckner. An operatorial approach to stochastic partial differential equations driven by linear multiplicative noise. *J. Eur. Math. Soc. (JEMS)*, 17(7):1789–1815, 2015.
- [13] C. Bardos and H. Brézis. Sur une classe de problèmes d’évolution non linéaires. *C. R. Acad. Sci. Paris. Sér. I Math.*, 266:56–59, 1968.
- [14] A. Bensoussan and A. Rascanu. Stochastic variational inequalities in infinite-dimensional spaces. *Numer. Funct. Anal. Optim.*, 18(1-2):19–54, 1997.
- [15] F. Bernis. Existence results for doubly nonlinear higher order parabolic equations on unbounded domains. *Math. Ann.*, 279:373–394, 1988.
- [16] 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).
- [17] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 152 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 2014.
- [18] E. DiBenedetto and R. E. Showalter. Implicit degenerate evolution equations and applications. *SIAM J. Math. Anal.*, 12(5):731–751, 1981.
- [19] E. Di Nezza, G. Palatucci, and E. Valdinoci. Hitchhiker’s guide to the fractional Sobolev spaces. *Bull. Sci. Math.*, 136(5):521–573, 2012.
- [20] F. Flandoli and D. Gatarek. Martingale and stationary solutions for stochastic Navier-Stokes equations. *Probab. Theory Related Fields*, 102(3):367–391, 1995.
- [21] H. Gajewski and I. V. Skrypnik. To the uniqueness problem for nonlinear parabolic equations. *Discrete Contin. Dyn. Syst.*, 10:315–336, 2004.
- [22] B. Gess. Strong solutions for stochastic partial differential equations of gradient type. *J. Funct. Anal.*, 263(8):2355–2383, 2012.
- [23] B. Gess and M. Röckner. Singular-degenerate multivalued stochastic fast diffusion equations. *SIAM J. Math. Anal.*, 47(5):4058–4090, 2015.
- [24] B. Gess and M. Röckner. Stochastic variational inequalities and regularity for degenerate stochastic partial differential equations. *Trans. Amer. Math. Soc.*, 369(5):3017–3045, 2017.
- [25] G. Gilardi and U. Stefanelli. Time-discretization and global solution for a doubly nonlinear Volterra equation. *J. Differential Equations*, 228:707–736, 2006.
- [26] O. Grange and F. Mignot. Sur la résolution d’une équation et d’une inéquation paraboliques non linéaires. *J. Funct. Anal.*, 11:77–92, 1972.
- [27] I. Gyöngy and N. Krylov. Existence of strong solutions for Itô’s stochastic equations via approximations. *Probab. Theory Related Fields*, 105:143–158, 1996.
- [28] V.-M. Hokkanen. An implicit nonlinear time dependent equation has a solution. *J. Math. Anal. Appl.*, 161:117–141, 1991.
- [29] V.-M. Hokkanen. Existence for nonlinear time dependent Volterra equations in Hilbert spaces. *An. Ştiinţ. Univ. Al. I. Cuza Iaşi Secţ. I a Mat.*, 38:29–49, 1992.
- [30] V.-M. Hokkanen. On nonlinear Volterra equations in Hilbert spaces. *Differential Integral Equations*, 5:647–669, 1992.
- [31] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, second edition, 1989.
- [32] M. Keller-Ressel and M. S. Müller. A Stefan-type stochastic moving boundary problem. *Stoch. Partial Differ. Equ. Anal. Comput.*, 4(4):746–790, 2016.
- [33] N. V. Krylov and B. L. Rozovskiĭ. Stochastic evolution equations. In: *Current Problems in Mathematics*, vol. 14, Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Informatsii, Moscow, 1979, pp. 71–147, 256 (in Russian).

- [34] E. Maitre and P. Witomski. A pseudo-monotonicity adapted to doubly nonlinear elliptic-parabolic equations. *Nonlinear Anal.*, 50:223–250, 2002.
- [35] C. Marinelli and L. Scarpa. Refined existence and regularity results for a class of semilinear dissipative SPDEs. *ArXiv e-prints*, Nov. 2017.
- [36] C. Marinelli and L. Scarpa. Strong solutions to SPDEs with monotone drift in divergence form. *Stoch. Partial Differ. Equ. Anal. Comput.*, 6(3):364–396, 2018.
- [37] C. Marinelli and L. Scarpa. A variational approach to dissipative SPDEs with singular drift. *Ann. Probab.*, 46(3):1455–1497, 2018.
- [38] C. Marinelli and L. Scarpa. Ergodicity and Kolmogorov equations for dissipative spdes with singular drift: a variational approach. *Potential Anal.*, published online Sept. 2018, doi:10.1007/s11118-018-9731-5.
- [39] C. Marinelli and L. Scarpa. A note on doubly nonlinear SPDEs with singular drift in divergence form. *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.*, 29(4):619–633, 2018.
- [40] C. Orrieri and L. Scarpa. Singular stochastic Allen-Cahn equations with dynamic boundary conditions. *J. Differential Equations*, 266(8):4624–4667, 2019.
- [41] E. Pardoux. *Equations aux dérivées partielles stochastiques nonlinéaires monotones*. PhD thesis, Université Paris XI, 1975.
- [42] É. Pardoux. Sur des équations aux dérivés partielles stochastiques monotones. *C. R. Acad. Sci. Paris Sér. A-B*, 275:A101–A103, 1972.
- [43] C. Prévôt and M. Röckner. *A concise course on stochastic partial differential equations*. Lecture Notes in Mathematics, 1905. Springer, Berlin, 2007.
- [44] P. A. Raviart. Sur la résolution de certaines équations paraboliques non linéaires. *J. Funct. Anal.*, 5:299–328, 1970.
- [45] N. Sapountzoglou, P. Wittbold, and A. Zimmermann. On a doubly nonlinear PDE with stochastic perturbation. *Stoch. Partial Differ. Equ. Anal. Comput.*, 7(2):297–330, 2019.
- [46] L. Scarpa. Well-posedness for a class of doubly nonlinear stochastic PDEs of divergence type. *J. Differential Equations*, 263(4):2113–2156, 2017.
- [47] L. Scarpa. On the stochastic Cahn-Hilliard equation with a singular double-well potential. *Nonlinear Anal.*, 171:102–133, 2018.
- [48] R. E. Showalter. *Monotone operators in Banach space and nonlinear differential equations*, volume 49 of *Mathematical Surveys and Monographs*. American Mathematical Society, 1997.
- [49] J. Simon. Compact sets in the space $L^p(0, T; B)$. *Ann. Mat. Pura Appl. (4)*, 146:65–96, 1987.
- [50] U. Stefanelli. Well-posedness and time discretization of a nonlinear Volterra integrodifferential equation. *J. Integral Equations Appl.*, 13:273–304, 2001.
- [51] U. Stefanelli. On a class of doubly nonlinear nonlocal evolution equations. *Differential Integral Equations*, 15:897–922, 2002.
- [52] U. Stefanelli. On some nonlocal evolution equations in Banach spaces. *J. Evol. Equ.*, 4:1–26, 2004.
- [53] G. Vallet and A. Zimmermann. Well-posedness for a pseudomonotone evolution problem with multiplicative noise. *J. Evol. Equ.*, 19:153–202, 2019.
- [54] A. Visintin. *Models of phase transitions*. Progress in Nonlinear Differential Equations and Their Applications, vol. 28, Birkhäuser, 1996.

(Luca Scarpa) FACULTY OF MATHEMATICS, UNIVERSITY OF VIENNA, OSKAR-MORGENSTERN-PLATZ 1, 1090 WIEN, AUSTRIA.

E-mail address: luca.scarpa@univie.ac.at

URL: <http://www.mat.univie.ac.at/~scarpa>

(Ulisse Stefanelli) FACULTY OF MATHEMATICS, UNIVERSITY OF VIENNA, OSKAR-MORGENSTERN-PLATZ 1, 1090 WIEN, AUSTRIA AND ISTITUTO DI MATEMATICA APPLICATA E TECNOLOGIE INFORMATICHE *E. Magenes*, V. FERRATA 1, 27100 PAVIA, ITALY.

E-mail address: ulisse.stefanelli@univie.ac.at

URL: <http://www.mat.univie.ac.at/~stefanelli>