Stochastic PDEs via convex minimization

Luca Scarpa & Ulisse Stefanelli

Published online: 14 Oct 2020.

Submit your article to this journal ☑

Article views: 10

View related articles ☑

View Crossmark data ☑

Taylor & Francis
Taylor & Francis Group

🔓 OPEN ACCESS | Check for updates

Stochastic PDEs via convex minimization

Luca Scarpa[a] 🆔 and Ulisse Stefanelli[a,b,c] 🆔

[a]Faculty of Mathematics, University of Vienna, Wien, Austria; [b]Vienna Research Platform on Accelerating Photoreaction Discovery, University of Vienna, Wien, Austria; [c]Istituto di Matematica Applicata e Tecnologie Informatiche E. Magenes, Pavia, Italy

ABSTRACT
We prove the applicability of the Weighted Energy-Dissipation (WED) variational principle to nonlinear parabolic stochastic partial differential equations in abstract form. The WED principle consists in the minimization of a parameter-dependent convex functional on entire trajectories. Its unique minimizers correspond to elliptic-in-time regularizations of the stochastic differential problem. As the regularization parameter tends to zero, solutions of the limiting problem are recovered. This in particular provides a direct approach via convex optimization to the approximation of nonlinear stochastic partial differential equations.

1. Introduction

This paper is concerned with stochastic quasilinear partial differential equations of the form

$$\mathrm{d}u - \mathrm{div}\ (D\phi(t, \nabla u))\ \mathrm{d}t + D\psi(t, u)\ \mathrm{d}t \ni f(t)\ \mathrm{d}t + B(t)\ \mathrm{d}W, \qquad (1)$$

complemented with suitable boundary and initial conditions. Here, the real-valued function u is defined on $\Omega \times [0, T] \times \mathcal{O}$, where $(\Omega, \mathscr{F}, \mathbb{P})$ is a probability space, $\mathcal{O} \subset \mathbb{R}^d$ is a smooth bounded domain, and $T > 0$ is a reference time. The functions $\phi(t, \cdot) : \mathbb{R}^d \to \mathbb{R}$ and $\psi(t, \cdot) : \mathbb{R} \to \mathbb{R}$ are asked to be convex, the gradients and $D\psi$ are taken with respect to the second variable only, and the time-dependent sources f and B are given. In particular, $B(\cdot) \in \mathscr{L}^2(U; L^2(\mathcal{O}))$ (Hilbert-Schmidt operators) is stochastically integrable with respect to W, a cylindrical Wiener process on a separable Hilbert space U.

Under different choices for the nonlinearities ϕ and ψ, equation (1) may arise in connection with various classical models, including the Allen-Cahn and the p-Laplace equation. Assume equation (1) to be complemented with homogeneous Dirichlet boundary conditions, for notational simplicity, and with the initial condition $u(0) = u_0$, where u_0 is some suitable initial datum. Letting $\phi(t, \cdot)$ and $\psi(t, \cdot)$ be of p-growth, equation (1)

can be weakly formulated in the dual of the space $W_0^{1,p}(\mathcal{O})$, according to the classical theory by Pardoux [1] and Krylov–Rozovskiĭ [2]. It is well-known that the solution u is an Itô process, in the sense that it can be represented in the general form

$$u = u^d + \int_0^{\cdot} u^s(r) \ \mathrm{d}W(r), \tag{2}$$

where the process u^d is differentiable in time and u^s is $\mathscr{L}^2(U; L^2(\mathcal{O}))$-valued and stochastically integrable with respect to W. This decomposition into the deterministic part u^d and the stochastic part u^s is unique. With this notation, u is a solution to the original problem (1) if and only if u satisfies the constraint (2) and the equations

$$\partial_t u^d - \mathrm{div} \ (D\phi(\cdot, \nabla u)) + D\psi(\cdot, u) \ni f, \qquad u^s = B, \qquad u^d(0) = u_0.$$

The aim of this paper is to tackle the weak solvability of equation (1) via the Weighted Energy-Dissipation (WED) variational approach. This hinges upon the minimization of the parameter-dependent functional I_ε on entire trajectories, the so-called WED functional, given by

$$I_\varepsilon(u) = \mathbb{E} \int_0^T \int_{\mathcal{O}} e^{-r/\varepsilon} \left(\frac{\varepsilon}{2} |\partial_t u^d(r)|^2 + \phi(r, \nabla u(r)) + \psi(r, u(r)) - f(r) \ u(r) \right) \ \mathrm{d}x \, \mathrm{d}r$$
$$+ \mathbb{E} \int_0^T e^{-r/\varepsilon} \frac{1}{2} \|u^s(r) - B(r)\|_{\mathscr{L}^2(U, L^2(\mathcal{O}))}^2 \ \mathrm{d}r.$$

The convex WED functional I_ε has to be minimized under two linear constraints, namely the decomposition (2) and the initial condition $u(0) = u_0$. This results in a *convex minimization* problem. Our main result, Theorem 2.1, states that, under suitable assumptions on data,

for all $\varepsilon > 0$ the minimizer u_ε of I_ε uniquely exists. As $\varepsilon \to 0$ we have that $u_\varepsilon \to u$ where u is the unique solution of the stochastic differential problem (1).

This provides a new variational approximation to the stochastic differential problem (1), making it accessible to a direct optimization approach, and paving the way to the application of the far-reaching tools of the calculus of variations [3–5].

The role of the exponential weight in I_ε is revealed by computing the corresponding Euler-Lagrange equation. In the current setting these formally read

$$-\varepsilon \partial_t (\partial_t u_\varepsilon^d)^d + \partial_t u_\varepsilon^d - \mathrm{div} \ (D\phi(\cdot, \nabla u_\varepsilon)) + D\psi(\cdot, u_\varepsilon) = f, \qquad u_\varepsilon^s = B + \varepsilon(\partial_t u_\varepsilon^d)^s,$$
$$\varepsilon \partial_t u_\varepsilon^d(T) = 0, \qquad u_\varepsilon^d(0) = u_0,$$

where we have also included the initial condition, for completeness. In particular, the minimizers u_ε solve an *elliptic-in-time regularization* of the stochastic differential problem (1), complemented by an extra Neumann boundary condition at T. Note that for all $\varepsilon > 0$ the problem is *not causal* and that causality is restored in the limit $\varepsilon \to 0$.

Elliptic-regularization techniques for nonlinear PDEs are quite classical. Introduced by Lions [6], they have been used by Kohn and Nirenberg [7], Oleinĭk [8], and again Lions [9, 10] in order to investigate regularity. An account on linear results can be found the book by Lions & Magenes [11], whereas an early result on solvability in a nonlinear setting is due to Barbu [12].

The variational formulation of elliptic-regularization via WED functionals can be traced back to Ilmanen [13], who used it in the context of Brakke mean-curvature flow of varifolds, and to Hirano [14] in connection with periodic solutions of gradient flows. A reference to WED functionals is already pointed out in the classical textbook by Evans [15, Problem 3, p. 487].

The WED variational approach has been applied to a variety of different parabolic problems, including gradient flows [16–19], rate-independent flows [20, 21], crack propagation [22], doubly-nonlinear flows [23–27], nonpotential perturbations [28, 29] and variational approximations [30], curves of maximal slope in metric spaces [31–33], mean curvature flow [13, 34], dynamic plasticity [35], and the incompressible Navier–Stokes system [36].

Motivated by a conjecture by De Giorgi [37], the WED variational approach has been extended to semilinear wave equations [38,39]. Extensions to other classes of hyperbolic problems including mixed hyperbolic-parabolic equations [40–43] and nonhomogeneous equations [44,45] have also been addressed.

In the context of stochastic PDEs, the application of tools from calculus of variations in order to characterize variational solutions is much less developed, and has been employed so far mainly in connection with the Brezis-Ekeland principle. In this direction, we mention the pioneering works by Barbu and Röckner [46, 47] dealing with SPDEs with additive and linear multiplicative noise, and by Krylov [48]. More recently, Boroushaki and Ghoussoub [49] generalized these results also to the case of multiplicative noise, by characterizing solutions as minima of self-dual functionals.

This paper contributes to the first application of the WED principle in the stochastic setting. Compared with the deterministic situation, the theory is here much more involved.

The first main difficulty arises in proving existence of minimizers for I_ε. This requires the characterization of the subdifferential of I_ε in terms of the Euler-Lagrange problem. In the stochastic setting, this ε-regularized problem consist of a forward-backward system of SPDEs. The identification of the Euler-Lagrange equation is more delicate compared to the deterministic framework. In the deterministic case, it is well known that the space of compactly-supported C^k test-functions $C_c^k(0, T)$ is dense in $L^2(0, T)$ for all $k \in \mathbb{N}$: this allows to identify the Euler-Lagrange equation pretty straightforwardly at least in a weak sense. By contrast, due to the presence of nonzero martingales in $L^2(\Omega; L^2(0, T))$, the usual deterministic techniques do not apply here, and the Euler-Lagrange equation has to be characterized using different tools, both on the analytical side and the probabilistic side. As a matter of fact, on the one hand we need to introduce suitable functional spaces of processes in Banach spaces (Itô processes), and on the other hand we rely on the abstract variational theory for backward SPDEs and martingale representation theorems in infinite dimensional spaces.

The second main difficulty concerns proving the well-posedness of the Euler–Lagrange problem. As we have pointed out above, the second-order Euler-Lagrange equation is non-causal and corresponds to a system of a forward and a backward first-order stochastic equation. The discussion of this forward-backward system calls for a further approximation on the nonlinearity. Identifications of nonlinear limits are performed via lower semicontinuity arguments, which in turn rely on specific Itô's formulas, both at the approximate and at the limit level.

In the paper, we actually consider a general class of abstract equations, including (1). Indeed, we frame the problem in the abstract variational setting of a Gelfand triple (V, H, V^*) and focus on

$$du + A(t, u) \ dt \ni B \ dW, \qquad u(0) = u_0,$$

where A is a time-dependent subdifferential-type operator from V to V^*, V being a separable reflexive Banach space and H a separable Hilbert space. We collect all relevant notation, list assumptions, and state Theorem 2.1, our main result, in Section 2. The proof of Theorem 2.1 is then split into Section 3 (Euler-Lagrange problem), Section 4 (convergence as $\varepsilon \to 0$), and Section 5 (existence of minimizers).

2. Main result

In the following, we directly focus on the abstract Cauchy problem

$$du + \partial\Phi(t, u) \ dt \ni B \ dW, \qquad u(0) = u_0. \tag{3}$$

The latter arises as variational formulation of an initial and boundary value problem for equation (1) by choosing the convex map $\Phi(t, \cdot)$ as

$$\Phi_p(t, \cdot) : W_0^{1,p}(\mathcal{O}) \to (-\infty, \infty], \quad \Phi_p(t, u) := \int_{\mathcal{O}} (\phi(t, \nabla u) + \psi(t, u)) \ dx.$$

Note that we have neglected the deterministic forcing f in (1) for the sake of notational simplicity. Indeed, this could be included in the analysis with no specific difficulty.

In this section we introduce the necessary notation and assumptions to make the meaning of problem (3) precise and we state our main result, Theorem 2.1. This is then proved in Sections 3-5.

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space endowed with a complete and right-continuous filtration $(\mathscr{F}_t)_{t \in [0, T]}$, where $T > 0$ is a fixed final time. Let also W be a cylindrical Wiener process on a separable Hilbert space U. We will assume that $(\mathscr{F}_t)_{t \in [0, T]}$ is the natural augmented filtration associated to W. The progressive σ-algebra on $\Omega \times [0, T]$ will be denoted by $\mathscr{P}$. For any Banach space E, the norm in E will be denoted by $\| \cdot \|_E$. For any $r, s \in [1, +\infty)$ we denote by $L_{\mathscr{P}}^r(\Omega; L^s(0, T; E))$ the usual space of Bochner-integrable functions which are strongly $\mathscr{P}$-measurable from $\Omega \times [0, T]$ to E. When $r > 1$ and $s = +\infty$, we explicitly define

$$L_{\mathscr{P}}^r(\Omega; L^\infty(0, T; E^*)) := \left\{ v : \Omega \to L^\infty(0, T; E^*) \text{ weakly} * \text{ meas.} \ : \ \mathbb{E}\|v\|_{L^\infty(0, T; E^*)}^r < \infty \right\}$$

where for any $f \in L^1(\Omega)$ we use the standard notation $\mathbb{E}f := \int_\Omega f \ d\mathbb{P}$ for the expected value. Recall that by [50, Thm. 8.20.3] we have the identification

$$L_{\mathscr{P}}^r(\Omega; L^\infty(0, T; E^*)) = \left(L_{\mathscr{P}}^{r/(r-1)}(\Omega; L^1(0, T; E)) \right)^*.$$

Moreover, for any $r \geq 1$, the symbol $L^r(\Omega; C^0([0, T]; E))$ denotes the space of r-integrable continuous adapted process (hence also progressively measurable) with values in E. For any pair of separable Hilbert spaces E_1 and E_2, we will use the symbols $\mathscr{L}(E_1, E_2)$ and $\mathscr{L}^2(E_1, E_2)$ for the spaces of linear continuous and Hilbert-Schmidt operators from E_1 and E_2, respectively.

Let us fix now a useful notation in order to denote suitable spaces of Itô processes. For every separable reflexive Banach space E_1 and any Hilbert spaces E, E_2, with $E_1, E_2 \hookrightarrow E$ continuously, and for any $r_1, r_2 \in [1, +\infty)$, we use the notation

$$\mathcal{I}^{r_1, r_2}(E_1, E_2) := L^{r_1}_{\mathscr{P}}(\Omega; W^{1, r_1}(0, T; E_1)) \oplus \left[L^{r_2}_{\mathscr{P}}(\Omega; L^2(0, T; \mathscr{L}^2(U, E_2))) \cdot W \right],$$

where we have used the classical symbol $\cdot W$ to denote stochastic integration with respect to W. Equivalently, we have the representation

$$\mathcal{I}^{r_1, r_2}(E_1, E_2) = \Big\{ z = z^d + z^s \cdot W :$$
$$z^d \in L^{r_1}_{\mathscr{P}}(\Omega; W^{1, r_1}(0, T; E_1)), \quad z^s \in L^{r_2}_{\mathscr{P}}(\Omega; L^2(0, T; \mathscr{L}^2(U, E_2))) \Big\}.$$

The latter specifies that the two components z^d and z^s are uniquely determined from the process z, see (2), so that the sum appearing above is actually a direct sum, and the projections

$$\Pi^d : \mathcal{I}^{r_1, r_2}(E_1, E_2) \to L^{r_1}_{\mathscr{P}}(\Omega; W^{1, r_1}(0, T; E_1)), \quad z \mapsto z^d,$$
$$\Pi^s : \mathcal{I}^{r_1, r_2}(E_1, E_2) \to L^{r_2}_{\mathscr{P}}(\Omega; L^2(0, T; \mathscr{L}^2(U, E_2))), \quad z \mapsto z^s,$$

are well-defined, linear, and continuous. Let us also point out that the space $\mathcal{I}^{r_1, r_2}(E_1, E_2)$ is a Banach space, and even a Hilbert space if $r_1 = r_2 = 2$ and E_1 is a Hilbert space. A natural norm on $\mathcal{I}^{r_1, r_2}(E_1, E_2)$ is given by

$$\|z\|_{\mathcal{I}^{r_1, r_2}(E_1, E_2)} := \|z^d\|_{L^{r_1}_{\mathscr{P}}(\Omega; W^{1, r_1}(0, T; E_1))} + \|z^s\|_{L^{r_2}_{\mathscr{P}}(\Omega; L^2(0, T; \mathscr{L}^2(U, E_2)))}, \qquad z \in \mathcal{I}^{r_1, r_2}(E_1, E_2).$$

Throughout the paper, we assume the following setting.

H0: H and V_0 are separable Hilbert spaces and V is a separable reflexive Banach space, with $V_0 \hookrightarrow V \hookrightarrow H$ continuously and densely.

In the sequel, we will identify H with its dual H^* in the canonical way, so that we have the continuous and dense inclusions

$$V_0 \hookrightarrow V \hookrightarrow H \hookrightarrow V^* \hookrightarrow V_0^*.$$

The scalar product in H and the duality pairing between V^* and V (and between V_0^* and V_0) will be denoted by the symbols $(\cdot, \cdot)$ and $\langle \cdot, \cdot \rangle$, respectively.

We assume the following hypotheses.

H1: $\Phi : \Omega \times [0, T] \times V \to [0, +\infty)$ is $\mathscr{P} \otimes \mathscr{B}(V)$-measurable, and $\Phi(\omega, t, \cdot) : V \to [0, +\infty)$ is convex and lower semicontinuous. We let $A(\omega, t, \cdot) := \partial \Phi(\omega, t, \cdot) : V \to 2^V$ for almost every $(\omega, t) \in \Omega \times [0, T]$. Moreover, we ask for constants $c_A, C_A > 0$ and $p \in [2, +\infty)$, and a $\mathscr{P}$-measurable process $f_A \in L^1(\Omega \times (0, T))$ such that, setting $q := p/(p-1)$,

$$\langle v, z \rangle \geq c_a \|z\|_V^p, \qquad \|v\|_{V^*}^q \leq f_A(\omega, t) + C_A \|z\|_V^p,$$

for almost every $(\omega, t) \in \Omega \times [0, T]$, for every $z \in V$, and for every $v \in A(\omega, t, z)$.

H2: $u_0 \in L^2(\Omega, \mathscr{F}_0; H)$ and $B \in L^2_{\mathscr{P}}(\Omega; L^2(0, T; \mathscr{L}^2(U, H)))$.

Let us point out that the progressive measurability of Φ required in **H1** implies that the operator A is $\mathscr{P} \otimes \mathscr{B}(V)/\mathscr{B}(V^*)$-Effros-measurable, in the sense of [51, 52].

Before moving on, let us comment on the choice of the space V_0. The introduction of V_0 will be needed in the paper since at some point we would have to rely on Itô's formula for the square of the V^*-norm. However, this cannot be done in general if V is a Banach space: indeed, in such case the duality mapping of V^* is nonlinear and possibly not twice Fréchet-differentiable, hence the required Itô formula is not trivial and not known in general, even in the extended framework of stochastic integration in UMD Banach spaces (see [53–55]). The introduction of the space V_0 is then employed to bypass this problem by exploiting its structure as Hilbert space, and allows to write an Itô formula in V_0^*. Clearly, if V is a Hilbert space itself, the optimal choice of V_0 is given by $V_0 = V$. In general, if V is only a Banach space, roughly speaking one should ideally choose the space V_0 *as large as possible*. For example, if $V = W^{s,\ell}(\mathcal{O})$ for a certain domain $\mathcal{O} \subset \mathbb{R}^d$ with Lipschitz boundary, with $\ell \in (2, +\infty)$ and $s > 0$, one could choose

$$V_0 = H^{s'}(\mathcal{O}), \qquad \forall \, s' \geq s + \frac{d}{2} - \frac{d}{\ell},$$

with the choice $s' = s + d/2 - d/\ell$ being optimal in this sense.

The classical variational theory on SPDEs (see [1, 2]) ensures that under the assumptions **H0**– **H2** the Cauchy problem (3) admits a unique solution (u, ξ), with

$$u \in L^2(\Omega; C^0([0, T]; H)) \cap L^p_{\mathscr{P}}(\Omega; L^p(0, T; V)) \qquad \xi \in L^q_{\mathscr{P}}(\Omega; L^q(0, T; V^*)), \qquad (4)$$

such that

$$\xi(\omega, t) \in A(\omega, t, u(\omega, t)) \quad \text{for a.e. } (\omega, t) \in \Omega \times (0, T), \qquad (5)$$

and

$$u(t) + \int_0^t \xi(s) \, ds = u_0 + \int_0^t B(s) \, dW(s) \quad \text{in } V^*, \quad \forall \, t \in [0, T], \quad \mathbb{P}\text{-a.s.} \qquad (6)$$

Let us reformulate this solution concept in a different fashion. We introduce the space

$$\mathcal{U} := \Big\{ z \in L^2(\Omega; C^0([0, T]; H)) \cap L^p_{\mathscr{P}}(\Omega; L^p(0, T; V)) : z = z^d + z^s \cdot W,$$

$$z^d \in L^q_{\mathscr{P}}(\Omega; W^{1,q}(0, T; V^*)), \quad z^s \in L^2_{\mathscr{P}}(\Omega; L^2(0, T; \mathscr{L}^2(U, H))) \Big\}.$$

Note that $\mathcal{U}$ can be written in compact form as

$$\mathcal{U} = L^2(\Omega; C^0([0, T]; H)) \cap L^p_{\mathscr{P}}(\Omega; L^p(0, T; V)) \cap \mathcal{I}^{q,2}(V^*, H),$$

so that in particular $\mathcal{U}$ is a Banach space.

With this notation, the process u solves the problem (4)–(6) if and only if

$$u \in \mathcal{U}, \qquad \begin{cases} \partial_t u^d + A(u) \ni 0, \\ u^d(0) = u_0, \\ u^s = B. \end{cases} \qquad (7)$$

In such a case, (4)–(6) are satisfied with the choice $\xi := -\partial_t u^d$.

As mentioned, the WED approach consists in minimizing an ε-dependent functional over entire trajectories and passing to the limit in the parameter ε. This procedure results in an elliptic regularization in time, hence delivering regular approximations. In

particular, the differential problem (3) is reformulated as a linearly constrained convex minimization. In the abstract setting of (4)-(6), letting $\varepsilon > 0$ we introduce the WED functional

$$I_\varepsilon : L_{\mathscr{P}}^p(\Omega; L^p(0, T; V)) \cap \mathcal{I}^{2,2}(H, H) =: \mathcal{V} \to [0, +\infty],$$

as

$$I_\varepsilon(u) := \begin{cases} \mathbb{E} \int_0^T e^{-t/\varepsilon} \left[\dfrac{\varepsilon}{2} \|\partial_t u^d(t)\|_H^2 + \Phi(t, u(t)) + \dfrac{1}{2} \|(u^s - B)(t)\|_{\mathscr{L}^2(U,H)}^2 \right] \, \mathrm{d}t \\ \qquad \text{if } u^d(0) = u_{0,\varepsilon}, \\ +\infty \\ \qquad \text{if } u^d(0) \neq u_{0,\varepsilon}. \end{cases}$$

We qualify the ε-dependent initial data $(u_{0,\varepsilon})_\varepsilon$ above by requiring that the sequence

$$(u_{0,\varepsilon})_{\varepsilon>0} \subset L^p(\Omega, \mathscr{F}_0; V_0) \tag{8}$$

is well-prepared, in the sense that, as $\varepsilon \searrow 0$,

$$u_{0,\varepsilon} \to u_0 \quad \text{in } L^2(\Omega, \mathscr{F}_0; H), \qquad \varepsilon \|u_{0,\varepsilon}\|_{L^p(\Omega, \mathscr{F}_0; V_0)}^p \to 0. \tag{9}$$

The existence of sequences fulfilling (8)–(9) follows directly from **H2** and the density of $V_0 \hookrightarrow H$, by standard regularization techniques.

Minimizers of I_ε will be proved to belong to the space

$$\begin{aligned} \mathcal{U}_{reg} := \{ z \in L^2(\Omega; C^0([0, T]; H)) &\cap L_{\mathscr{P}}^p(\Omega; L^p(0, T; V)) : z = z^d + z^s \cdot W, \\ &z^d \in L_{\mathscr{P}}^q(\Omega; C^1([0, T]; V_0^*)) \cap L_{\mathscr{P}}^2(\Omega; H^1(0, T; H)), \\ &z^s \in L_{\mathscr{P}}^2(\Omega; L^2(0, T; \mathscr{L}^2(U, H))), \\ &\partial_t z^d = (\partial_t z^d)^d + (\partial_t z^d)^s \cdot W, \\ &(\partial_t z^d)^d \in L_{\mathscr{P}}^q(\Omega; W^{1,q}(0, T; V^*)), \quad (\partial_t z^d)^s \in L_{\mathscr{P}}^2(\Omega; L^2(0, T; \mathscr{L}^2(U, H))) \}. \end{aligned}$$

Again, note that a more compact notation for $\mathcal{U}_{reg}$ reads

$$\mathcal{U}_{reg} = \{ z \in L_{\mathscr{P}}^p(\Omega; L^p(0, T; V)) \cap \mathcal{I}^{2,2}(H, H) : \partial_t z^d \in \mathcal{I}^{q,2}(V^*, H) \}.$$

Let us point out in particular that $\mathcal{U}_{reg} \hookrightarrow \mathcal{V} \hookrightarrow \mathcal{U}$ with continuous inclusions.

The Euler-Lagrange equation for functional I_ε corresponds to the ε-regularized problem

$$u_\varepsilon \in \mathcal{U}_{reg}, \qquad \begin{cases} -\varepsilon \partial_t (\partial_t u_\varepsilon^d)^d + \partial_t u_\varepsilon^d + A(u_\varepsilon) \ni 0, \\ \varepsilon \partial_t u_\varepsilon^d(T) = 0, \\ u_\varepsilon^d(0) = u_{0,\varepsilon}, \\ u_\varepsilon^s = B + \varepsilon(\partial_t u_\varepsilon^d)^s. \end{cases} \tag{10}$$

Note that the second-order problem (10) can be seen as a system of two equations of first order in time, one forward and one backward, by using the classical substitution $v_\varepsilon := \partial_t u_\varepsilon^d$. Indeed, with this notation (10) is equivalent to

$$\begin{cases} \mathrm{d}u_\varepsilon = v_\varepsilon \, \mathrm{d}t + (B + \varepsilon G_\varepsilon) \, \mathrm{d}W, \\ u_\varepsilon(0) = u_{0,\varepsilon}, \end{cases} \qquad \begin{cases} -\varepsilon \mathrm{d}v_\varepsilon + v_\varepsilon \, \mathrm{d}t + A(u_\varepsilon) \, \mathrm{d}t \ni -\varepsilon G_\varepsilon \, \mathrm{d}W, \\ \varepsilon v_\varepsilon(T) = 0. \end{cases} \tag{11}$$

Note that the variables of the forward-backward system (11) are three, namely $u_\varepsilon, v_\varepsilon$,

and G_ε. Indeed, while the forward equation has a unique variable (u_ε), the concept of solution for the backward stochastic equation requires the two variables v_ε and G_ε due to the need of representation theorems for martingales. In particular, we have that $G_\varepsilon = v_\varepsilon^s$ is uniquely determined by the backward stochastic equation.

The main result of the paper reads as follows.

Theorem 2.1 (Weighted Energy-Dissipation approach). *Assume* **H0–H2**. *Then:*

i) (Minimization) *For all $\varepsilon > 0$ the functional I_ε admits a unique global minimizer $u_\varepsilon \in \mathcal{V}$.*

ii) (Euler-Lagrange equation) *The minimizer also satisfies $u_\varepsilon \in \mathcal{U}_{reg}$ and it is the unique solution to the problem* (10). *Namely, there exists a unique triplet $(\xi_\varepsilon, v_\varepsilon, G_\varepsilon)$ with*

$$\xi_\varepsilon \in L^q_{\mathscr{P}}(\Omega; L^q(0, T; V^*)),$$
$$v_\varepsilon \in L^q(\Omega; C^0([0, T]; V_0^*)) \cap L^2_{\mathscr{P}}(\Omega; L^2(0, T; H)),$$
$$G_\varepsilon \in L^q_{\mathscr{P}}(\Omega; L^2(0, T; \mathscr{L}^2(U, V_0^*))),$$

such that

$$\xi_\varepsilon(\omega, t) \in A(\omega, t, u_\varepsilon(\omega, t)) \quad \text{for a.e. } (\omega, t) \in \Omega \times (0, T),$$

and

$$u_\varepsilon(t) = u_{0,\varepsilon} + \int_0^t v_\varepsilon(s) \, ds + \int_0^t (B + \varepsilon G_\varepsilon)(s) \, dW(s),$$
$$\varepsilon v_\varepsilon(t) + \int_t^T v_\varepsilon(s) \, ds + \int_t^T \xi_\varepsilon(s) \, ds = -\varepsilon \int_t^T G_\varepsilon(s) \, dW(s),$$

for every $t \in [0, T]$, $\mathbb{P}$-almost surely. In particular, it holds that $\partial_t u_\varepsilon^d = v_\varepsilon, u_\varepsilon^s = B + \varepsilon G_\varepsilon$, and $v_\varepsilon^s = G_\varepsilon$.

iii) (Convergence) *As $\varepsilon \to 0$ it holds that*

$$u_\varepsilon \rightharpoonup u \quad \text{in } L^p_{\mathscr{P}}(\Omega; L^p(0, T; V)) \cap L^q_{\mathscr{P}}(\Omega; W^{s,q}(0, T; V_0^*)) \quad \forall s \in (0, 1/2),$$
$$v_\varepsilon \rightharpoonup -\xi \quad \text{in } L^q_{\mathscr{P}}(\Omega; L^q(0, T; V_0^*)),$$
$$\xi_\varepsilon \rightharpoonup \xi \quad \text{in } L^q_{\mathscr{P}}(\Omega; L^q(0, T; V^*)),$$
$$\varepsilon v_\varepsilon \to 0 \quad \text{in } L^q(\Omega; C^0([0, T]; V_0^*)) \cap L^2_{\mathscr{P}}(\Omega; L^2(0, T; H)),$$
$$\varepsilon G_\varepsilon \to 0 \quad \text{in } L^q_{\mathscr{P}}(\Omega; L^2(0, T; \mathscr{L}^2(U, V_0^*))),$$

where (u, ξ) is the unique solution to the problem (7) *in the sense of* (4)-(6). *Furthermore, if $V \hookrightarrow H$ compactly and $p < 4$, it also holds that*

$$u_\varepsilon \to u \quad \text{in } L^r_{\mathscr{P}}(\Omega; L^p(0, T; H)) \qquad \forall r \in [1, p).$$

The proof of Theorem 2.1 is recorded in the coming Sections 3-5. In particular, Part ii of the theorem is proved in Section 3, where we focus on the well-posedness of the forward-backward regularized problem (10). Then, the convergence Part iii of Theorem 2.1 is proved in Section 4. Eventually, the existence of minimizers is checked in Section 5.

This counterintuitive structuring of the proof of Theorem 2.1 is motivated by the fact that the existence of minimizers of I_ε follows from proving that the corresponding Euler-Lagrange problem has a unique solution. One hence has to check the well-posedness of problem (10) first.

3. The forward-backward regularized problem

This section is devoted to proof of the well-posedness of the ε-regularized problem (10) in the sense of Theorem 2.1.ii. Throughout the section, $\varepsilon > 0$ is fixed.

First of all, let A_H be the random and time-dependent unbounded operator on H defined as

$$A_H : \Omega \times [0, T] \times H \to 2^H, \quad A_H(\omega, t, z) := A(\omega, t, z) \cap H, \quad (\omega, t) \in \Omega \times [0, T], \quad z \in H.$$

It is not difficult to show that, for every $(\omega, t) \in \Omega \times [0, T]$, the unbounded operator $A_H(\omega, t, \cdot)$ is maximal monotone on H. Indeed, the monotonicity is an immediate consequence of the monotonicity of A. As for the maximality, note that the operator $I + A(\omega, t, \cdot) : V \to V^*$, where I is the identity in H (namely, $\langle Iv, w \rangle = (v, w) \forall v, w \in H$), is maximal monotone and coercive by assumption on A, hence it is surjective, which yields the maximality of $A_H(\omega, t, \cdot)$. Furthermore, since A is $\mathscr{P} \otimes \mathscr{B}(V)/\mathscr{B}(V^*)$-Effros-measurable, it follows that A_H is $\mathscr{P} \otimes \mathscr{B}(H)/\mathscr{B}(H)$-Effros-measurable as well.

3.1. The approximation

Since A_H is maximal monotone on H in its last component, for any $\lambda > 0$ its resolvent and its Yosida approximation are well defined, respectively, as

$$J_\lambda : \Omega \times [0, T] \times H \to H, \quad J_\lambda(\omega, t, z) := (I + \lambda A_H(\omega, t, \cdot))^{-1}(z), \quad (\omega, t, z) \in \Omega \times [0, T] \times H,$$

and

$$A_\lambda : \Omega \times [0, T] \times H \to H, \quad A_\lambda(\omega, t, z) := \frac{z - J_\lambda(\omega, t, z)}{\lambda}, \quad (\omega, t, z) \in \Omega \times [0, T] \times H.$$

It is well-known that J_λ and A_λ are 1- and $1/\lambda$-Lipschitz-continuous in their third component, respectively, uniformly in $\Omega \times [0, T]$. Moreover, the Effros-measurability of A_H implies that J_λ and A_λ are $\mathscr{P} \otimes \mathscr{B}(H)/\mathscr{B}(H)$-measurable (see for example [56, Prop. 3.12]).

For any $\lambda > 0$, we consider the approximated problem

$$\begin{cases} du_{\varepsilon\lambda} = v_{\varepsilon\lambda} \ dt + (B + \varepsilon G_{\varepsilon\lambda}) \ dW, \\ u_{\varepsilon\lambda}(0) = u_{0,\varepsilon}, \end{cases} \qquad \begin{cases} -\varepsilon dv_{\varepsilon\lambda} + v_{\varepsilon\lambda} \ dt + A_\lambda(u_{\varepsilon\lambda}) \ dt = -\varepsilon G_{\varepsilon\lambda} \ dW, \\ \varepsilon v_{\varepsilon\lambda}(T) = 0. \end{cases}$$

$$(12)$$

We say that a triplet $(u_{\varepsilon\lambda}, v_{\varepsilon\lambda}, G_{\varepsilon\lambda})$ is a solution to the approximated problem (12) if

$$(u_{\varepsilon\lambda}, v_{\varepsilon\lambda}, G_{\varepsilon\lambda}) \in L^2(\Omega; C^0([0, T]; H)) \times L^2(\Omega; C^0([0, T]; H)) \times L^2_{\mathscr{P}}(\Omega; L^2(0, T; \mathscr{L}^2(U, H)))$$

and it holds that

$$u_{\varepsilon\lambda}(t) = u_{0,\varepsilon} + \int_0^t v_{\varepsilon\lambda}(s) \ ds + \int_0^t (B + \varepsilon G_{\varepsilon\lambda})(s) \ dW(s),$$

$$\varepsilon v_{\varepsilon\lambda}(t) + \int_t^T v_{\varepsilon\lambda}(s) \ ds + \int_t^T A_\lambda(s, u_{\varepsilon\lambda}(s)) \ ds = -\varepsilon \int_t^T G_{\varepsilon\lambda}(s) \ dW(s),$$

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

3.2. Existence of solutions to the approximated problem

We prove here that the approximated problem (12) admits a solution $(u_{\varepsilon\lambda}, v_{\varepsilon\lambda}, G_{\varepsilon\lambda})$. To this end, we characterize the unique solution $(u_{\varepsilon\lambda}, v_{\varepsilon\lambda}, G_{\varepsilon\lambda})$ as the unique minimizer of a suitable approximated WED functional.

Let us first introduce some preliminary notation. Note that we have the representation

$$I_\varepsilon = I_\varepsilon^1 + S_\varepsilon + I_\varepsilon^2,$$

where

$$I_\varepsilon^1 : \mathcal{I}^{2,2}(H,H) \to [0,+\infty), \qquad I_\varepsilon^2 : L_{\mathscr{P}}^p(\Omega; L^p(0,T;V)) \to [0,+\infty),$$

$$I_\varepsilon^1(z) := \mathbb{E} \int_0^T e^{-t/\varepsilon} \left[\frac{\varepsilon}{2} \|\partial_t z^d(t)\|_H^2 + \frac{1}{2} \|(z^s - B)(t)\|_{\mathscr{L}^2(U,H)}^2 \right] \mathrm{d}t, \quad z \in \mathcal{I}^{2,2}(H,H),$$

$$I_\varepsilon^2(z) := \mathbb{E} \int_0^T e^{-t/\varepsilon} \Phi(t,z(t)) \, \mathrm{d}t, \quad z \in L_{\mathscr{P}}^p(\Omega; L^p(0,T;V)),$$

and

$$S_\varepsilon : \mathcal{I}^{2,2}(H,H) \to [0,+\infty], \qquad S_\varepsilon(z) := \begin{cases} 0 & \text{if } z^d(0) = u_{0,\varepsilon}, \\ +\infty & \text{otherwise.} \end{cases}$$

Moreover, it will be useful to introduce the notation

$$\mathcal{I}_0^{2,2}(H,H) := \{ h \in \mathcal{I}^{2,2}(H,H) : h^d(0) = 0 \}.$$

The natural candidate as WED functional related to the approximated problem (12) is clearly

$$I_{\varepsilon\lambda} : \mathcal{I}^{2,2}(H,H) \to [0,+\infty), \qquad I_{\varepsilon\lambda} := I_\varepsilon^1 + S_\varepsilon + I_{\varepsilon\lambda}^2,$$

with

$$I_{\varepsilon\lambda}^2 : \mathcal{I}^{2,2}(H,H) \to [0,+\infty),$$

$$I_{\varepsilon\lambda}^2(z) := \mathbb{E} \int_0^T e^{-t/\varepsilon} \Phi_\lambda(t,z(t)) \mathrm{d}t, \quad z \in \mathcal{I}^{2,2}(H,H),$$

where Φ_λ is the Moreau-Yosida regularization of Φ, i.e.

$$\Phi_\lambda : \Omega \times [0,T] \times H \to [0,+\infty),$$

$$\Phi_\lambda(\omega,t,z) := \Phi(\omega,t,J_\lambda(z)) + \frac{1}{2\lambda} \|z - J_\lambda(z)\|_H^2, \qquad (\omega,t,z) \in \Omega \times [0,T] \times H.$$

It is well known that $\Phi_\lambda(\omega,t,\cdot)$ is Gâteaux-differentiable on H with derivative $A_\lambda(\omega,t,\cdot)$, for every $(\omega,t) \in \Omega \times [0,T]$.

We now show that the approximated problem (12) is equivalent to the minimization of $I_{\varepsilon\lambda}$. In this direction, we aim now at characterizing the subdifferential of $I_{\varepsilon\lambda}$. This will follow after some intermediate steps.

First of all, we characterize the subdifferential of the sum $I_\varepsilon^1 + S_\varepsilon$.

Lemma 3.1. *The subdifferential of $I_\varepsilon^1 + S_\varepsilon : \mathcal{I}^{2,2}(H,H) \to [0,+\infty]$ is the operator*

$$\partial\big(I_\varepsilon^1 + S_\varepsilon\big) : \mathcal{I}^{2,2}(H,H) \to 2^{\mathcal{I}^{2,2}(H,H)^*}$$

defined in the following way:

$$D\big(\partial\big(I_\varepsilon^1 + S_\varepsilon\big)\big) := \{z \in \mathcal{I}^{2,2}(H,H) : z^d(0) = u_{0,\varepsilon}\},$$

and, for every $z \in D\big(\partial\big(I_\varepsilon^1 + S_\varepsilon\big)\big)$ and $w \in \mathcal{I}^{2,2}(H,H)^*$,

$$w \in \partial\big(I_\varepsilon^1 + S_\varepsilon\big)(z)$$

if and only if there exists $\tilde{w} \in \mathcal{I}_0^{2,2}(H,H)^\perp$ such that

$$\langle w, h\rangle_{\mathcal{I}^{2,2}(H,H)}$$
$$= \mathbb{E}\int_0^T e^{-t/\varepsilon}\Big[\varepsilon\big(\partial_t z^d(t), \partial_t h^d(t)\big) + \big((z^s - B)(t), h^s(t)\big)_{\mathscr{L}^2(U,H)}\Big]\,\mathrm{d}t + \langle\tilde{w}, h\rangle_{\mathcal{I}^{2,2}(H,H)}$$

for every $h \in \mathcal{I}^{2,2}(H,H)$.

Proof. First of all, it is clear that I_ε^1 is proper, convex, and lower semicontinuous on $\mathcal{I}^{2,2}(H,H)$. Moreover, we have that I_ε^1 is actually Gâteaux-differentiable. Indeed, for every $z, h \in \mathcal{I}^{2,2}(H,H)$ and $\delta \neq 0$ we have

$$\frac{I_\varepsilon^1(z + \delta h) - I_\varepsilon^1(z)}{\delta} = \mathbb{E}\int_0^T e^{-t/\varepsilon}\Big[\varepsilon\big(\partial_t z^d(t), \partial_t h^d(t)\big) + \big((z^s - B)(t), h^s(t)\big)_{\mathscr{L}^2(U,H)}\Big]\,\mathrm{d}t$$
$$+ \delta\mathbb{E}\int_0^T e^{-t/\varepsilon}\Big[\frac{\varepsilon}{2}\|\partial_t h^d(t)\|_H^2 + \frac{1}{2}\|h^s(t)\|_{\mathscr{L}^2(U,H)}^2\Big]\,\mathrm{d}t,$$

where the second term on the right-hand side converges to 0 as $\delta \to 0$ since $h \in \mathcal{I}^{2,2}(H,H)$. Hence, I_ε^1 is Gâteaux-differentiable, its Gâteaux-differential coincides with its subdifferential and it is given by

$$\partial I_\varepsilon^1 : \mathcal{I}^{2,2}(H,H) \to \mathcal{I}^{2,2}(H,H)^*,$$
$$\langle\partial I_\varepsilon^1(z), h\rangle_{\mathcal{I}^{2,2}(H,H)} = \mathbb{E}\int_0^T e^{-t/\varepsilon}\Big[\varepsilon\big(\partial_t z^d(t), \partial_t h^d(t)\big) + \big((z^s - B)(t), h^s(t)\big)_{\mathscr{L}^2(U,H)}\Big]\,\mathrm{d}t,$$
$$z, h \in \mathcal{I}^{2,2}(H,H).$$

Secondly, S_ε is proper, convex, and lower semicontinuous on $\mathcal{I}^{2,2}(H,H)$. Moreover, its subdifferential is given by

$$\partial S_\varepsilon : \mathcal{I}^{2,2}(H;H) \to 2^{\mathcal{I}^{2,2}(H,H)^*},$$
$$\partial S_\varepsilon(z) := \Big\{w \in \mathcal{I}^{2,2}(H;H) : \langle w, h\rangle_{\mathcal{I}^{2,2}(H,H)} = 0 \quad \forall h \in \mathcal{I}^{2,2}(H,H), h^d(0) = 0\Big\}$$
$$z \in D(\partial S_\varepsilon) := \Big\{z \in I^{2,2}(H,H) : z^d(0) = u_{0,\varepsilon}\Big\}.$$

In other words, we have that

$$\partial S_\varepsilon(z) = \mathcal{I}_0^{2,2}(H,H)^\perp, \quad z \in D(\partial S_\varepsilon).$$

Consequently, since ∂I_ε^1 and ∂S_ε are maximal monotone on $\mathcal{I}^{2,2}(H,H)$, and

$$\mathrm{Int}\big(D\big(I_\varepsilon^1\big)\big) \cap D(S_\varepsilon) = \mathcal{I}^{2,2}(H,H) \cap \big\{z \in \mathcal{I}^{2,2}(H,H) : z^d(0) = u_{0,\varepsilon}\big\} \neq \emptyset,$$

a classical result on convex analysis (see [57, Thm. 2.10]) ensures that

$$\partial\big(I_\varepsilon^1 + S_\varepsilon\big) = \partial I_\varepsilon^1 + \partial S_\varepsilon,$$

with

$$D\big(\partial\big(I_\varepsilon^1 + S_\varepsilon\big)\big) = D\big(\partial I_\varepsilon^1\big) \cap D(\partial S_\varepsilon) = D(\partial S_\varepsilon).$$

This implies that, for every $z \in D\big(\partial\big(I_\varepsilon^1 + S_\varepsilon\big)\big)$ and $w \in \partial\big(I_\varepsilon^1 + S_\varepsilon\big)$, we have

$$w = \partial I_\varepsilon^1(z) + \tilde{w}$$

for a certain $\tilde{w} \in \mathcal{I}_0^{2,2}(H,H)^\perp$, as required.   □

Now, we characterize the subdifferential of $I_{\varepsilon\lambda}^2$.

Lemma 3.2. *The subdifferential of $I_{\varepsilon\lambda}^2 : \mathcal{I}^{2,2}(H,H) \to [0, +\infty)$ is the single-valued operator*

$$\partial I_{\varepsilon\lambda} : \mathcal{I}^{2,2}(H,H) \to \mathcal{I}^{2,2}(H,H)^*,$$

$$\langle \partial I_{\varepsilon\lambda}^2(z), h \rangle_{\mathcal{I}^{2,2}(H,H)} = \mathbb{E} \int_0^T e^{-t/\varepsilon}(A_\lambda(t, z(t)), h(t)) \; \mathrm{d}t, \qquad z, h \in \mathcal{I}^{2,2}(H,H).$$

In particular, it holds that $D\big(\partial I_{\varepsilon\lambda}^2\big) = \mathcal{I}^{2,2}(H,H)$.

Proof. The proof is consequence of a classical computation: see for example [58, Prop. 1.1].   □

We are now able to characterize the subdifferential of the functional $I_{\varepsilon\lambda}$.

Lemma 3.3. *The subdifferential of $I_{\varepsilon\lambda} : \mathcal{I}^{2,2}(H,H) \to [0, +\infty]$ is the operator*

$$\partial I_{\varepsilon\lambda} : \mathcal{I}^{2,2}(H,H) \to 2^{\mathcal{I}^{2,2}(H,H)^*}$$

defined in the following way:

$$D(\partial I_{\varepsilon\lambda}) := \{z \in \mathcal{I}^{2,2}(H,H) : z^d(0) = u_{0,\varepsilon}\},$$

and, for every $z \in D(\partial I_{\varepsilon\lambda})$ and $w \in \mathcal{I}^{2,2}(H,H)^*$,

$$w \in \partial I_{\varepsilon\lambda}(z)$$

if and only if there exists $\tilde{w} \in \mathcal{I}_0^{2,2}(H,H)^\perp$ such that, for every $h \in \mathcal{I}^{2,2}(H,H)$,

$$
\begin{aligned}
\langle w, h \rangle_{\mathcal{I}^{2,2}(H,H)} = \mathbb{E} \int_0^T e^{-t/\varepsilon}\Big[&\varepsilon\big(\partial_t z^d(t), \partial_t h^d(t)\big) + (A_\lambda(t, z(t)), h(t)) \\
&+ \big((z^s - B)(t), h^s(t)\big)_{\mathscr{L}^2(U,H)} \Big] \; \mathrm{d}t + \langle \tilde{w}, h \rangle_{\mathcal{I}^{2,2}(H,H)}.
\end{aligned}
$$

In particular, for every $z \in \mathcal{I}^{2,2}(H,H)$ with $z^d(0) = u_{0,\varepsilon}$ and $w \in \partial I_{\varepsilon\lambda}(z)$ it holds that

$$
\begin{aligned}
&\langle w, h \rangle_{\mathcal{I}^{2,2}(H,H)} \\
&= \mathbb{E} \int_0^T e^{-t/\varepsilon}\Big[\varepsilon\big(\partial_t z^d(t), \partial_t h^d(t)\big) + (A_\lambda(t, z(t)), h(t)) + \big((z^s - B)(t), h^s(t)\big)_{\mathscr{L}^2(U,H)}\Big] \; \mathrm{d}t
\end{aligned}
$$

for every $h \in \mathcal{I}_0^{2,2}(H,H)$.

Proof. Since $D\big(I_\varepsilon^1 + S_\varepsilon\big) = \{z \in \mathcal{I}^{2,2}(H,H) : z^d(0) = u_{0,\varepsilon}\}$ and $D\big(I_{\varepsilon\lambda}^2\big) = \mathcal{I}^{2,2}(H,H)$, we have

$$D\big(I_\varepsilon^1 + S_\varepsilon\big) \cap \mathrm{Int}\big(D\big(I_{\varepsilon\lambda}^2\big)\big) = \{z \in \mathcal{I}^{2,2}(H,H) : z^d(0) = u_{0,\varepsilon}\} \neq \emptyset.$$

Hence, by the classical result [57, Thm. 2.10], we infer that

$$\partial I_{\varepsilon\lambda} = \partial\big(I_\varepsilon^1 + S_\varepsilon\big) + \partial I_{\varepsilon\lambda}^2, \qquad D(\partial I_{\varepsilon\lambda}) = D\big(\partial\big(I_\varepsilon^1 + S_\varepsilon\big)\big) \cap D\big(\partial I_{\varepsilon\lambda}^2\big).$$

The thesis follows then directly from Lemma 3.1 and Lemma 3.2. □

We have now all the tools in order to show existence of solutions to the approximated problem (12) via minimization of the regularized functional $I_{\varepsilon\lambda}$. Namely, we have the following result.

Proposition 3.4 (Well-posedness of the approximated problem). *For every $\lambda > 0$, the functional $I_{\varepsilon\lambda}$ admits a unique global minimizer*

$$z_{\varepsilon\lambda} \in \mathcal{I}^{2,2}(H,H).$$

Moreover, the triplet $\Big(z_{\varepsilon\lambda}, \partial_t z_{\varepsilon\lambda}^d, \big(\partial_t z_{\varepsilon\lambda}^d\big)^s\Big)$ is a solution of the approximated problem (12).

Proof. We note first that the functional $I_\varepsilon^1 + S_\varepsilon$ is strictly convex and coercive on $\mathcal{I}^{2,2}(H,H)$, hence so is the functional $I_{\varepsilon\lambda}$ since Φ_λ is convex and bounded from below. Since $\mathcal{I}^{2,2}(H,H)$ is reflexive, this ensures the existence and uniqueness of a global minimizer $z_{\varepsilon\lambda} \in \mathcal{I}^{2,2}(H,H)$ for $I_{\varepsilon\lambda}$. Clearly, we have that $z_{\varepsilon\lambda} \in D(I_{\varepsilon\lambda})$, so that $z_{\varepsilon\lambda}^d(0) = u_{0,\varepsilon}$. Moreover, by definition of minimizer we have that

$$0 \in \partial I_{\varepsilon\lambda}(z_{\varepsilon\lambda}).$$

By virtue of Lemma 3.3, we deduce that

$$\mathbb{E}\int_0^T e^{-t/\varepsilon}\Big[\varepsilon\Big(\partial_t z_{\varepsilon\lambda}^d(t), \partial_t h^d(t)\Big) + (A_\lambda(t, z_{\varepsilon\lambda}(t)), h(t))\Big]\ \mathrm{d}t \tag{13}$$
$$+ \mathbb{E}\int_0^T e^{-t/\varepsilon}\big(\big(z_{\varepsilon\lambda}^s - B\big)(t), h^s(t)\big)_{\mathscr{L}^2(U,H)}\ \mathrm{d}t = 0$$

for every $h \in \mathcal{I}^{2,2}(H,H)$ with $h(0) = 0$. Now, since $z_{\varepsilon\lambda} \in L^2\big(\Omega; C^0([0,T];H)\big)$ and A_λ is uniformly Lipschitz-continuous in its third variable, it is clear that

$$\int_0^{\cdot} e^{-s/\varepsilon}A_\lambda(s, z_{\varepsilon\lambda}(s))\ \mathrm{d}s \in L^2\big(\Omega; C^1([0,T];H)\big).$$

Hence, by Itô's formula we have, in differential (formal) form, that

$$\mathrm{d}\left(\int_0^t e^{-s/\varepsilon}A_\lambda(s, z_{\varepsilon\lambda}(s))\ \mathrm{d}s, h(t)\right)$$
$$= \left(e^{-t/\varepsilon}A(t, z_{\varepsilon\lambda}), h(t)\right)\ \mathrm{d}t + \left(\int_0^t e^{-s/\varepsilon}A_\lambda(s, z_{\varepsilon\lambda}(s))\ \mathrm{d}s, \partial_t h^d(t)\right)\ \mathrm{d}t$$
$$+ \left(\int_0^t e^{-r/\varepsilon}A_\lambda(r, z_{\varepsilon\lambda}(r))\ \mathrm{d}r, h^s(t)\ \mathrm{d}W(t)\right).$$

Integrating on $[0,T]$ and taking expectations we infer that

$$\mathbb{E}\left(\int_0^T e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds, h(T)\right) = \mathbb{E}\int_0^T e^{-t/\varepsilon}(A(t, z_{\varepsilon\lambda}(t)), h(t)) \ dt$$
$$+ \mathbb{E}\int_0^T \left(\int_0^t e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds, \partial_t h^d(t)\right) \ dt.$$

Noting that the first term on the right-hand side appears in (13) as well, by substitution we infer then that

$$\mathbb{E}\int_0^T \left(\varepsilon e^{-t/\varepsilon}\partial_t z_{\varepsilon\lambda}^d(t) - \int_0^t e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds, \partial_t h^d(t)\right) \ dt$$
$$+ \mathbb{E}\int_0^T e^{-t/\varepsilon}\left((z_{\varepsilon\lambda}^s - B)(t), h^s(t)\right)_{\mathscr{L}^2(U,H)} \ dt \tag{14}$$
$$+ \mathbb{E}\left(\int_0^T e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds, h(T)\right) = 0$$

for every $h \in \mathcal{I}^{2,2}(H, H)$ such that $h(0) = 0$. Now, note that for any such h, we have that

$$h(T) = h^d(T) + \int_0^T h^s(r) \ dW(r) = \int_0^T \partial_t h^d(s) \ ds + \int_0^T h^s(r) \ dW(r),$$

which yields in turn that

$$\mathbb{E}\left(\int_0^T e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds, h(T)\right) = \mathbb{E}\int_0^T \left(\int_0^T e^{-r/\varepsilon} A_\lambda(r, z_{\varepsilon\lambda}(r)) \ dr, \partial_t h^d(s)\right) \ ds$$
$$+ \mathbb{E}\left(\int_0^T e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds, \int_0^T h^s(r) \ dW(r)\right).$$

Using this equality for the last term of (14) we obtain that

$$\mathbb{E}\int_0^T \left(\varepsilon e^{-t/\varepsilon}\partial_t z_{\varepsilon\lambda}^d(t) - \int_0^t e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ dr + \int_0^T e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds, \partial_t h^d(t)\right) \ dt$$
$$+ \mathbb{E}\left(\int_0^T e^{-t/\varepsilon}(z_{\varepsilon\lambda}^s - B)(t) \ dW(t) + \int_0^T e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds, \int_0^T h^s(t) \ dW(t)\right) = 0$$
$$\tag{15}$$

for every $h \in \mathcal{I}^{2,2}(H, H)$ with $h^d(0) = 0$. Now, for any arbitrary $K \in L^2_{\mathscr{P}}(\Omega; L^2(0, T; H))$, note that the process

$$h_K := \int_0^{\cdot} K(s) \ ds \in \mathcal{I}^{2,2}(H, H)$$

satisfies $h_K(0) = 0$, hence it is a possible test in equation (15). Since $h_K^s = 0$, we deduce that

$$\mathbb{E}\int_0^T \left(\varepsilon e^{-t/\varepsilon}\partial_t z_{\varepsilon\lambda}^d(t) - \int_0^t e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds + \int_0^T e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds, K(t)\right) \ dt = 0$$

for every $K \in L^2_{\mathscr{P}}(\Omega; L^2(0, T; H))$. Let us stress that the first component of the scalar product appearing in this equality is *not* progressively measurable, hence one cannot simply deduce that it vanishes by arbitrariness of K. Nonetheless, note that by definition of conditional expectation and by the adaptedness of K, we have

$$\mathbb{E} \int_0^T \left(\int_0^T e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds, K(t) \right) \ dt$$

$$= \mathbb{E} \int_0^T \left(\mathbb{E} \left[\int_0^T e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds \ \Big| \ \mathscr{F}_t \right], K(t) \right) \ dt.$$

Since $(\mathscr{F}_t)_{t \in [0, T]}$ is the filtration generated by W and

$$\int_0^T e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds \in L^2(\Omega, \mathscr{F}_T; H),$$

the process

$$t \mapsto \mathbb{E} \left[\int_0^T e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds \ \Big| \ \mathscr{F}_t \right]$$

is an H-valued continuous square-integrable martingale, and in particular is progressively measurable. We deduce then that the variational equality reads equivalently

$$\mathbb{E} \int_0^T \left(\varepsilon e^{-t/\varepsilon} \partial_t z^d_{\varepsilon\lambda}(t) - \int_0^t e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds + \mathbb{E} \left[\int_0^T e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds \Big| \mathscr{F}_t \right], K(t) \right) \ dt$$

$$= 0 \qquad \forall \ K \in L^2_{\mathscr{P}}(\Omega; L^2(0, T; H)).$$

At this point, since the process appearing on the left term of the scalar product belongs to the space $L^2_{\mathscr{P}}(\Omega; L^2(0, T; H))$, by arbitrariness of K we have that

$$\varepsilon e^{-t/\varepsilon} \partial_t z^d_{\varepsilon\lambda}(t) - \int_0^t e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds + \mathbb{E} \left[\int_0^T e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds \ \Big| \ \mathscr{F}_t \right] = 0$$

almost everywhere in $\Omega \times [0, T]$. We deduce that there is a $d\mathbb{P} \otimes dt$-version of $\partial_t z^d_{\varepsilon\lambda}$ (which will be denoted with the same symbol for brevity of notation) such that

$$\partial_t z^d_{\varepsilon\lambda} \in L^2(\Omega; C^0([0, T]; H)). \tag{16}$$

Furthermore, by the classical martingale representation theorem in Hilbert spaces (see e.g. [59, Prop. 4.1] and [60]), there exists a process $G_{\varepsilon\lambda} \in L^2_{\mathscr{P}}(\Omega; L^2(0, T; \mathscr{L}^2(U, H)))$ such that

$$\mathbb{E} \left[\int_0^T e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds \ \Big| \ \mathscr{F}_t \right] = \int_0^T e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds + \varepsilon \int_t^T e^{-s/\varepsilon} G_{\varepsilon\lambda}(s) \ dW(s)$$

for every $t \in [0, T]$, from which it follows that

$$\varepsilon e^{-t/\varepsilon} \partial_t z^d_{\varepsilon\lambda}(t) + \int_t^T e^{-s/\varepsilon} A_\lambda(s, z_{\varepsilon\lambda}(s)) \ ds = -\varepsilon \int_t^T e^{-s/\varepsilon} G_{\varepsilon\lambda}(s) \ dW(s) \qquad \forall t \in [0, T].$$

It follows in particular that

$$\partial_t z^d_{\varepsilon\lambda} \in \mathcal{I}^{2,2}(H,H), \qquad \partial_t z^d_{\varepsilon\lambda}(T) = 0, \tag{17}$$

and

$$\varepsilon \mathrm{d}\left(\partial_t z^d_{\varepsilon\lambda}\right) = \varepsilon \mathrm{d}\left(e^{t/\varepsilon}e^{-t/\varepsilon}\partial_t z^d_{\varepsilon\lambda}\right) = e^{t/\varepsilon}\left(e^{-t/\varepsilon}\partial_t z^d_{\varepsilon\lambda}\right)\,\mathrm{d}t + \varepsilon e^{t/\varepsilon}\mathrm{d}\left(e^{-t/\varepsilon}\partial_t z^d_{\varepsilon\lambda}\right)$$

$$= \left(\partial_t z^d_{\varepsilon\lambda}\right)\,\mathrm{d}t + A_\lambda(\cdot, z_{\varepsilon\lambda})\,\mathrm{d}t + \varepsilon G_{\varepsilon\lambda}\,\mathrm{d}W,$$

which reads, equivalently,

$$-\varepsilon \mathrm{d}\left(\partial_t z^d_{\varepsilon\lambda}\right) + \left(\partial_t z^d_{\varepsilon\lambda}\right)\,\mathrm{d}t + A_\lambda(\cdot, z_{\varepsilon\lambda})\,\mathrm{d}t = -\varepsilon G_{\varepsilon\lambda}\,\mathrm{d}W. \tag{18}$$

Now, we go back to the variational formulation (15) and take $h_L := L \cdot W \in \mathcal{I}^{2,2}(H,H)$ as test process for any arbitrary $L \in L^2_{\mathscr{P}}\big(\Omega; L^2\big(0,T; \mathscr{L}^2(U,H)\big)\big)$. Clearly, the process h_L satisfies $h_L(0) = 0$ and is hence a possible test in (15). Since $\partial_t h^d_L = 0$, we infer that

$$\mathbb{E}\left(\int_0^T e^{-t/\varepsilon}\big(z^s_{\varepsilon\lambda} - B\big)(t)\,\mathrm{d}W(t) + \int_0^T e^{-s/\varepsilon}A_\lambda(s, z_{\varepsilon\lambda}(s))\,\mathrm{d}s, (L \cdot W)(T)\right) = 0$$

for every $L \in L^2_{\mathscr{P}}\big(\Omega; L^2\big(0,T; \mathscr{L}^2(U,H)\big)\big)$. Now, note that by arbitrariness of L, by the usual martingale representation theorems, the random variable $L \cdot W(T)$ is arbitrary in the subspace $L^2_0(\Omega, \mathscr{F}_T; H)$ of elements with null mean. It follows in particular that the process on the left-term is constant and equal to its mean, i.e.

$$\int_0^T e^{-t/\varepsilon}\big(z^s_{\varepsilon\lambda} - B\big)(t)\,\mathrm{d}W(t) + \int_0^T e^{-s/\varepsilon}A_\lambda(s, z_{\varepsilon\lambda}(s))\,\mathrm{d}s$$

$$= \mathbb{E}\left[\int_0^T e^{-t/\varepsilon}\big(z^s_{\varepsilon\lambda} - B\big)(t)\,\mathrm{d}W(t) + \int_0^T e^{-s/\varepsilon}A_\lambda(s, z_{\varepsilon\lambda}(s))\,\mathrm{d}s\right]$$

$$= \mathbb{E}\left[\int_0^T e^{-s/\varepsilon}A_\lambda(s, z_{\varepsilon\lambda}(s))\,\mathrm{d}s\right].$$

Recalling the definition of $G_{\varepsilon\lambda}$, we have that

$$\mathbb{E}\left[\int_0^T e^{-s/\varepsilon}A_\lambda(s, z_{\varepsilon\lambda}(s))\,\mathrm{d}s\right] = \int_0^T e^{-s/\varepsilon}A_\lambda(s, z_{\varepsilon\lambda}(s))\,\mathrm{d}s + \varepsilon\int_0^T e^{-s/\varepsilon}G_{\varepsilon\lambda}(s)\,\mathrm{d}W(s),$$

so that by comparison we obtain that

$$\int_0^T e^{-t/\varepsilon}\big(z^s_{\varepsilon\lambda} - B\big)(t)\,\mathrm{d}W(t) = \varepsilon\int_0^T e^{-s/\varepsilon}G_{\varepsilon\lambda}(s)\,\mathrm{d}W(s) \qquad \mathbb{P}\text{-a.s.},$$

yielding

$$z^s_{\varepsilon\lambda} = B + \varepsilon G_{\varepsilon\lambda}. \tag{19}$$

It is then clear now from (16), (17), (18), and (19), that $\big(z_{\varepsilon\lambda}, \partial_t z^d_{\varepsilon\lambda}, G_{\varepsilon\lambda}\big)$ is a solution to the approximated problem (12). □

3.3. Uniform estimates

We want to pass now to the limit as $\lambda \searrow 0$ in (12). To this end, let us show some uniform estimates in λ, still with $\varepsilon > 0$ fixed.

Itô's formula for the square of the H-norm yields

$$\frac{1}{2}\mathbb{E}\|u_{\varepsilon\lambda}(T)\|_H^2 = \frac{1}{2}\mathbb{E}\|u_{0,\varepsilon}\|_H^2 + \mathbb{E}\int_0^T (v_{\varepsilon\lambda}(s), u_{\varepsilon\lambda}(s)) \ ds + \frac{1}{2}\mathbb{E}\int_0^T \|B(s)\|_{\mathscr{L}^2(U,H)}^2 \ ds$$
$$+ \frac{\varepsilon^2}{2}\mathbb{E}\int_0^T \|G_{\varepsilon\lambda}(s)\|_{\mathscr{L}^2(U,H)}^2 \ ds + \varepsilon\mathbb{E}\int_0^T (B(s), G_{\varepsilon\lambda}(s))_{\mathscr{L}^2(U;H)} \ ds. \tag{20}$$

Note now that

$$d(\varepsilon v_{\varepsilon\lambda}, u_{\varepsilon\lambda}) = \varepsilon v_{\varepsilon\lambda} \ du_{\varepsilon\lambda} + u_{\varepsilon\lambda}\varepsilon dv_{\varepsilon\lambda} + \varepsilon d[G_{\varepsilon\lambda}, B + \varepsilon G_{\varepsilon\lambda}],$$

which yields, taking (12) into account,

$$d(\varepsilon v_{\varepsilon\lambda}, u_{\varepsilon\lambda}) = \varepsilon\|v_{\varepsilon\lambda}\|_H^2 \ dt + (\varepsilon v_{\varepsilon\lambda}, (B + \varepsilon G_{\varepsilon\lambda}) \ dW) + (v_{\varepsilon\lambda}, u_{\varepsilon\lambda}) \ dt + (A_\lambda(u_{\varepsilon\lambda}), u_{\varepsilon\lambda}) \ dt$$
$$+ \varepsilon(u_{\varepsilon\lambda}, G_{\varepsilon\lambda} \ dW) + \varepsilon(B, G_{\varepsilon\lambda})_{\mathscr{L}^2(U,H)} \ dt + \varepsilon^2\|G_{\varepsilon\lambda}\|_{\mathscr{L}^2(U,H)}^2 \ dt.$$

Recalling that $\varepsilon v_\varepsilon(T) = 0$, we deduce then that

$$\varepsilon\mathbb{E}\int_0^T \|v_{\varepsilon\lambda}(s)\|_H^2 \ ds + \mathbb{E}\int_0^T (v_{\varepsilon\lambda}(s), u_{\varepsilon\lambda}(s)) \ ds + \mathbb{E}\int_0^T (A_\lambda(u_{\varepsilon\lambda}(s)), u_{\varepsilon\lambda}(s)) \ ds$$
$$+ \varepsilon^2\mathbb{E}\int_0^T \|G_{\varepsilon\lambda}(s)\|_{\mathscr{L}^2(U,H)}^2 \ ds + \varepsilon\mathbb{E}\int_0^T (B(s), G_{\varepsilon\lambda}(s))_{\mathscr{L}^2(U;H)} ds = -\varepsilon\mathbb{E}(v_{\varepsilon\lambda}(0), u_{0,\varepsilon}). \tag{21}$$

Now, noting that $I - J_\lambda = \lambda A_\lambda$, recalling that $A_\lambda(\cdot) \in A(J_\lambda(\cdot))$, the coercivity of A entails

$$(A_\lambda(u_{\varepsilon\lambda}), u_{\varepsilon\lambda}) = (A_\lambda(u_\lambda), J_\lambda(u_{\varepsilon\lambda})) + (A_\lambda(u_{\varepsilon\lambda}), u_{\varepsilon\lambda} - J_\lambda(u_{\varepsilon\lambda}))$$
$$= (A_\lambda(u_\lambda), J_\lambda(u_{\varepsilon\lambda})) + \lambda\|A_\lambda(u_{\varepsilon\lambda})\|_H^2$$
$$\geq c_A\|J_\lambda(u_{\varepsilon\lambda})\|_V^p + \lambda\|A_\lambda(u_{\varepsilon\lambda})\|_H^2.$$

Hence, by comparing (20) and (21) we obtain

$$\frac{1}{2}\mathbb{E}\|u_{\varepsilon\lambda}(T)\|_H^2 + \varepsilon\mathbb{E}\int_0^T \|v_{\varepsilon\lambda}(s)\|_H^2 \ ds + c_A\mathbb{E}\int_0^T \|J_\lambda(u_{\varepsilon\lambda}(s))\|_V^p \ ds$$
$$+ \lambda\mathbb{E}\int_0^T \|A_\lambda(u_{\varepsilon\lambda}(s))\|_H^2 \ ds + \frac{\varepsilon^2}{2}\mathbb{E}\int_0^T \|G_{\varepsilon\lambda}(s)\|_{\mathscr{L}^2(U,H)}^2 \ ds \tag{22}$$
$$\leq \frac{1}{2}\mathbb{E}\|u_{0,\varepsilon}\|_H^2 - \varepsilon\mathbb{E}(v_{\varepsilon\lambda}(0), u_{0,\varepsilon}) + \frac{1}{2}\mathbb{E}\int_0^T \|B(s)\|_{\mathscr{L}^2(U;H)}^2 \ ds.$$

Next, denoting by $R_0 : V_0 \to V_0^*$ the duality mapping, Itô's formula for the square of the V_0^*-norm of $v_{\varepsilon\lambda}$ yields, by (12),

$$\frac{\varepsilon}{2}\|v_{\varepsilon\lambda}(t)\|_{V_0^*}^2 + \int_t^T \|v_{\varepsilon\lambda}(s)\|_{V_0^*}^2 \ ds + \frac{\varepsilon}{2}\int_t^T \|G_{\varepsilon\lambda}(s)\|_{\mathscr{L}^2(U,V_0^*)}^2 \ ds$$
$$= -\int_t^T (A_\lambda(u_{\varepsilon\lambda}(s)), R_0^{-1}(v_{\varepsilon\lambda}(s))) \ ds - \varepsilon\int_t^T (R_0^{-1}(v_{\varepsilon\lambda}(s), G_{\varepsilon\lambda}(s) \ dW(s))$$

for every $t \in [0, T]$, $\mathbb{P}$-almost surely. We would like to write Itô's formula for the q-power of the V_0^*-norm of $v_{\varepsilon\lambda}$. Clearly, if $p = 2$ then also $q = 2$ and nothing has to be

done. If $p > 2$ then we have $q \in (1, 2)$ and this can be achieved by writing Itô's formula for the real function $|\cdot|^{q/2}$. However, since $q \in (1, 2)$ the function $|\cdot|^{q/2}$ is not of class C^2, and this cannot be done straightaway. We need then to rely on a suitable approximation of the function $|\cdot|^{q/2}$. Let us introduce to this end the approximations

$$\gamma_\delta : [0, +\infty) \to \mathbb{R}, \qquad \gamma_\delta(r) := (r^2 + \delta^2)^{q/4}, \quad r \geq 0, \quad \delta > 0.$$

Clearly, we have that $\gamma_\delta \in C^\infty([0, +\infty))$, with

$$\gamma_\delta'(r) = \begin{cases} \frac{q}{2}(r^2 + \delta^2)^{\frac{q-4}{4}} r & \text{if } r > 0, \\ 0 & \text{if } r = 0, \end{cases}$$

and

$$\gamma_\delta''(r) = \begin{cases} \frac{q}{2}\left(r^2 + \delta^2\right)^{\frac{q-4}{4}} + \frac{q(q-4)}{4}\left(r^2 + \delta^2\right)^{\frac{q-8}{4}} r^2 & \text{if } r > 0, \\ 0 & \text{if } r = 0. \end{cases}$$

Consequently, for every $r \geq 0$ it holds that

$$\lim_{\delta \searrow 0} \gamma_\delta(r) = r^{q/2}, \qquad \lim_{\delta \searrow 0} \gamma_\delta'(r) = \frac{q}{2} r^{q/2-1} 1_{\{r>0\}}, \qquad \lim_{\delta \searrow 0} \gamma_\delta''(r) = \frac{q}{2}\frac{q-2}{2} r^{q/2-2} 1_{\{r>0\}}.$$

Since γ_δ is of class C^2, we can use the classical finite dimensional Itô's formula (see, e.g. [61]) and infer that

$$\gamma_\delta\left(\frac{\varepsilon}{2}\|v_{\varepsilon\lambda}(t)\|_{V_0^*}^2\right) + \int_t^T \gamma_\delta'\left(\frac{\varepsilon}{2}\|v_{\varepsilon\lambda}(s)\|_{V_0^*}^2\right)\left(\|v_{\varepsilon\lambda}(s)\|_{V_0}^2 + \frac{\varepsilon}{2}\|G_{\varepsilon\lambda}(s)\|_{\mathscr{L}^2(U, V_0^*)}^2\right) \, ds$$

$$+ \frac{\varepsilon^2}{2}\int_t^T \gamma_\delta''\left(\frac{\varepsilon}{2}\|v_{\varepsilon\lambda}(s)\|_{V_0^*}^2\right)\|\left(R_0^{-1}v_{\varepsilon\lambda}(s), G_{\varepsilon\lambda}(s)\right)\|_{\mathscr{L}^2(U, \mathbb{R})}^2 \, ds$$

$$= -\int_t^T \gamma_\delta'\left(\frac{\varepsilon}{2}\|v_{\varepsilon\lambda}(s)\|_{V_0^*}^2\right)\left(A_\lambda(u_{\varepsilon\lambda}(s)), R_0^{-1}(v_{\varepsilon\lambda}(s))\right) \, ds$$

$$- \varepsilon\int_t^T \gamma_\delta'\left(\frac{\varepsilon}{2}\|v_{\varepsilon\lambda}(s)\|_{V_0^*}^2\right)\left(R_0^{-1}(v_{\varepsilon\lambda}(s)), G_{\varepsilon\lambda}(s) \, dW(s)\right).$$

Now, letting $\delta \searrow 0$ it follows by the Dominated Convergence Theorem that

$$\left(\frac{\varepsilon}{2}\right)^{q/2}\|v_{\varepsilon\lambda}(t)\|_{V_0^*}^q + \frac{q}{2}\left(\frac{\varepsilon}{2}\right)^{q/2-1}\int_t^T \|v_{\varepsilon\lambda}(s)\|_{V_0^*}^q \, ds$$

$$+ \frac{q}{2}\left(\frac{\varepsilon}{2}\right)^{q/2-1}\int_t^T \frac{\varepsilon}{2} 1_{\{\|v_{\varepsilon\lambda}(s)\|_{V_0^*}>0\}}\|v_{\varepsilon\lambda}(s)\|_{V_0^*}^{q-2}\|G_{\varepsilon\lambda}(s)\|_{\mathscr{L}^2(U, V_0^*)}^2 \, ds$$

$$+ \frac{q}{2}\frac{q-2}{2}\frac{\varepsilon^2}{2}\left(\frac{\varepsilon}{2}\right)^{q/2-2}\int_t^T 1_{\{\|v_{\varepsilon\lambda}(s)\|_{V_0^*}>0\}}\|v_{\varepsilon\lambda}(s)\|_{V^*}^{q-4}\|\left(R_0^{-1}v_{\varepsilon\lambda}(s), G_{\varepsilon\lambda}(s)\right)\|_{\mathscr{L}^2(U, \mathbb{R})}^2 \, ds$$

$$= -\frac{q}{2}\left(\frac{\varepsilon}{2}\right)^{q/2-1}\int_t^T 1_{\{\|v_{\varepsilon\lambda}(s)\|_{V_0^*}>0\}}\|v_{\varepsilon\lambda}(s)\|_{V_0^*}^{q-2}\left(A_\lambda(u_{\varepsilon\lambda}(s)), R_0^{-1}(v_{\varepsilon\lambda}(s))\right) \, ds$$

$$- \varepsilon\frac{q}{2}\left(\frac{\varepsilon}{2}\right)^{q/2-1}\int_t^T 1_{\{\|v_{\varepsilon\lambda}(s)\|_{V_0^*}>0\}}\|v_{\varepsilon\lambda}(s)\|_{V_0^*}^{q-2}\left(R_0^{-1}(v_{\varepsilon\lambda}(s)), G_{\varepsilon\lambda}(s) \, dW(s)\right).$$

Multiplying by $\left(\frac{\varepsilon}{2}\right)^{1-\frac{q}{2}}$, taking expectations, and using the Young inequality yields

$$\frac{\varepsilon}{2}\mathbb{E}\|v_{\varepsilon\lambda}(t)\|_{V_0^*}^q + \frac{q}{2}\mathbb{E}\int_t^T \left(\|v_{\varepsilon\lambda}(s)\|_{V_0^*}^q + \frac{\varepsilon}{2}\mathbf{1}_{\{\|v_{\varepsilon\lambda}(s)\|_{V^*}>0\}}\|v_{\varepsilon\lambda}(s)\|_{V_0^*}^{q-2}\|G_{\varepsilon\lambda}(s)\|_{\mathscr{L}^2(U,V_0^*)}^2\right)\, ds$$

$$+ \frac{q(q-2)}{4}\varepsilon\mathbb{E}\int_t^T \mathbf{1}_{\{\|v_{\varepsilon\lambda}(s)\|_{V_0^*}>0\}}\|v_{\varepsilon\lambda}(s)\|_{V_0^*}^{q-4}\|(R_0^{-1}v_{\varepsilon\lambda}(s),G_{\varepsilon\lambda}(s))\|_{\mathscr{L}^2(U,\mathbb{R})}^2\, ds$$

$$\leq \frac{q}{2}\mathbb{E}\int_t^T \|R_0^{-1}v_{\varepsilon\lambda}(s)\|_V^{q-1}\|A_\lambda(u_{\varepsilon\lambda}(s))\|_{V^*}\, ds$$

$$\leq \frac{c_0 q}{2}\mathbb{E}\int_t^T \|R_0^{-1}v_{\varepsilon\lambda}(s)\|_{V_0}^{q-1}\|A_\lambda(u_{\varepsilon\lambda}(s))\|_{V^*}\, ds$$

$$\leq \frac{q}{2p}\mathbb{E}\int_t^T \|v_{\varepsilon\lambda}(s)\|_{V_0^*}^q\, ds + \frac{c_0^q}{2}\mathbb{E}\int_t^T \|A_\lambda(u_{\varepsilon\lambda}(s))\|_{V^*}^q\, ds,$$

where c_0 denotes the norm of the continuous inclusion $V_0 \hookrightarrow V$. Since

$$\frac{q}{2} - \frac{q}{2p} = \frac{q}{2}\left(1 - \frac{1}{p}\right) = \frac{1}{2},$$

by rearranging the terms and using the boundedness of A we deduce that

$$\frac{\varepsilon}{2}\mathbb{E}\|v_{\varepsilon\lambda}(t)\|_{V_0^*}^q + \frac{1}{2}\mathbb{E}\int_t^T \|v_{\varepsilon\lambda}(s)\|_{V_0^*}^q\, ds$$

$$+ \varepsilon\frac{q}{4}\mathbb{E}\int_t^T \mathbf{1}_{\{\|v_{\varepsilon\lambda}(s)\|_{V^*}>0\}}\|v_{\varepsilon\lambda}(s)\|_{V_0^*}^{q-2}\|G_{\varepsilon\lambda}(s)\|_{\mathscr{L}^2(U,V_0^*)}^2\, ds \qquad (23)$$

$$\leq \frac{c_0^q}{2}\|f\|_{L^1(\Omega\times(0,T))} + \frac{C_A c_0^q}{2}\mathbb{E}\int_t^T \|J_\lambda(u_{\varepsilon\lambda}(s))\|_V^p\, ds$$

for every $t \in [0, T]$, $\mathbb{P}$-almost surely. Now, since $0 < q/2 < 1$, its conjugate exponent $-q/(2-q)$ is negative: the reverse Young's inequality implies then that

$$\mathbb{E}\int_t^T \mathbf{1}_{\{\|v_{\varepsilon\lambda}(s)\|_{V^*}>0\}}\|v_{\varepsilon\lambda}(s)\|_{V_0^*}^{q-2}\|G_{\varepsilon\lambda}(s)\|_{\mathscr{L}^2(U,V_0^*)}^2\, ds$$

$$\geq \frac{2}{q}\mathbb{E}\int_t^T \|G_{\varepsilon\lambda}(s)\|_{\mathscr{L}^2(U,V_0^*)}^q\, ds - \frac{q}{2-q}\mathbb{E}\int_t^T \|v_{\varepsilon\lambda}(s)\|_{V_0^*}^q\, ds$$

Taking this information into account we deduce from (23) that

$$\frac{\varepsilon}{2}\mathbb{E}\|v_{\varepsilon\lambda}(t)\|_{V_0^*}^q + \frac{1}{2}\mathbb{E}\int_t^T \|v_{\varepsilon\lambda}(s)\|_{V_0^*}^q\, ds + \frac{\varepsilon}{2}\mathbb{E}\int_t^T \|G_{\varepsilon\lambda}(s)\|_{\mathscr{L}^2(U,V_0^*)}^q\, ds$$

$$\leq \frac{c_0^q}{2}\|f\|_{L^1(\Omega\times(0,T))} + \frac{C_A c_0^q}{2}\mathbb{E}\int_0^T \|J_\lambda(u_{\varepsilon\lambda}(s))\|_V^p\, ds + \frac{q^2}{4(2-q)}\varepsilon\mathbb{E}\int_t^T \|v_{\varepsilon\lambda}(s)\|_{V_0^*}^q\, ds,$$

yielding, by the Gronwall lemma,

$$\frac{\varepsilon}{2}\sup_{t\in[0,T]}\mathbb{E}\|v_{\varepsilon\lambda}(t)\|_{V_0^*}^q + \frac{1}{2}\mathbb{E}\int_0^T \|v_{\varepsilon\lambda}(s)\|_{V_0^*}^q\, ds + \frac{\varepsilon}{2}\mathbb{E}\int_0^T \|G_{\varepsilon\lambda}(s)\|_{\mathscr{L}^2(U,V_0^*)}^q\, ds$$

$$\leq e^{\frac{Tq^2}{2(2-q)}}\left(\frac{c_0^q}{2}\|f\|_{L^1(\Omega\times(0,T))} + \frac{C_A c_0^q}{2}\mathbb{E}\int_0^T \|J_\lambda(u_{\varepsilon\lambda}(s))\|_V^p\, ds\right). \qquad (24)$$

Now, by multiplying the inequality (24) by $e^{-\frac{Tq^2}{2(2-q)}}\frac{c_A}{C_A c_0^q}$ and summing it with inequality (22), the last term on the right-hand side of (24) can be incorporated into the corresponding term on the left-hand side of (22): rearranging the terms, we obtain

$$
\frac{1}{2}\mathbb{E}\|u_{\varepsilon\lambda}(T)\|_H^2
$$
$$
+ \mathbb{E}\int_0^T \left(\varepsilon\|v_{\varepsilon\lambda}(s)\|_H^2 \ ds + \frac{c_A}{2}\|J_\lambda(u_{\varepsilon\lambda}(s))\|_V^p + \lambda\|A_\lambda(u_{\varepsilon\lambda}(s))\|_H^2 + \frac{\varepsilon^2}{2}\|G_{\varepsilon\lambda}(s)\|_{\mathscr{L}^2(U,H)}^2 \right) \ ds
$$
$$
+ \frac{e^{-\frac{Tq^2}{2(2-q)}}c_A}{2C_A c_0^q}\left(\varepsilon \sup_{t\in[0,T]} \mathbb{E}\|v_{\varepsilon\lambda}(t)\|_{V_0^*}^q + \mathbb{E}\int_0^T \|v_{\varepsilon\lambda}(s)\|_{V_0^*}^q \ ds + \varepsilon\mathbb{E}\int_0^T \|G_{\varepsilon\lambda}(s)\|_{\mathscr{L}^q(U,V_0^*)}^q \ ds \right)
$$
$$
\leq \frac{1}{2}\mathbb{E}\|u_{0,\varepsilon}\|_H^2 - \varepsilon\mathbb{E}\big(v_{\varepsilon\lambda}(0),u_{0,\varepsilon}\big) + \frac{1}{2}\mathbb{E}\int_0^T \|B(s)\|_{\mathscr{L}^2(U;H)}^2 \ ds + \frac{c_A}{2C_A}\|f\|_{L^1(\Omega\times(0,T))}.
$$

At this point, note the second term on the right-hand side above can be handled using the averaged Young inequality: indeed, we infer that, for every $\sigma > 0$,

$$
-\varepsilon\mathbb{E}\big(v_{\varepsilon\lambda}(0),u_{0,\varepsilon}\big) \leq \varepsilon\mathbb{E}\Big[\|v_{\varepsilon\lambda}(0)\|_{V_0^*}\|u_{0,\varepsilon}\|_{V_0}\Big]
$$
$$
\leq \frac{\sigma^q}{q}\varepsilon \sup_{t\in[0,T]} \|v_{\varepsilon\lambda}(t)\|_{L^q(\Omega,\mathscr{F}_0;V_0^*)}^q + \frac{\varepsilon}{\sigma^p p}\|u_{0,\varepsilon}\|_{L^p(\Omega,\mathscr{F}_0;V_0)}^p.
$$

Choosing and fixing σ sufficiently small, independent of λ and ε, for example

$$
\sigma := \left(e^{-\frac{Tq^2}{2(2-q)}}\frac{qc_A}{4C_A c_0^q} \right)^{1/q},
$$

rearranging the terms we deduce that there exists a positive constant $M = M(c_A, C_A, c_0, q, T)$, independent of both λ and ε, such that

$$
\mathbb{E}\|u_{\varepsilon\lambda}(T)\|_H^2 + \varepsilon\mathbb{E}\int_0^T \|v_{\varepsilon\lambda}(s)\|_H^2 \ ds + \mathbb{E}\int_0^T \|J_\lambda(u_{\varepsilon\lambda}(s))\|_V^p \ ds
$$
$$
+ \lambda\mathbb{E}\int_0^T \|A_\lambda(u_{\varepsilon\lambda}(s))\|_H^2 \ ds + \varepsilon^2\mathbb{E}\int_0^T \|G_{\varepsilon\lambda}(s)\|_{\mathscr{L}^2(U,H)}^2 \ ds
$$
$$
+ \varepsilon \sup_{t\in[0,T]} \mathbb{E}\|v_{\varepsilon\lambda}(t)\|_{V_0^*}^q + \mathbb{E}\int_0^T \|v_{\varepsilon\lambda}(s)\|_{V_0^*}^q \ ds + \varepsilon\mathbb{E}\int_0^T \|G_{\varepsilon\lambda}(s)\|_{\mathscr{L}^q(U,V_0^*)}^q \ ds
$$
$$
\leq M\mathbb{E}\Big(\|u_{0,\varepsilon}\|_H^2 + \varepsilon\|u_{0,\varepsilon}\|_{V_0}^p + \|B\|_{L^2(0,T;\mathscr{L}^2(U;H))}^2\Big).
$$

At this point, note that by the assumption (9) on $(u_{0,\varepsilon})_\varepsilon$, we have that the right-hand side is uniformly bounded in ε and λ.

Then, we deduce that, by updating the value of the constant M (here below and the following possibly changing from line to line),

$$
\varepsilon\|v_{\varepsilon\lambda}\|_{L^2(\Omega;L^2(0,T;H))}^2 + \|J_\lambda(u_{\varepsilon\lambda})\|_{L^p(\Omega;L^p(0,T;V))}^p + \lambda\|A_\lambda(u_{\varepsilon\lambda})\|_{L^2(\Omega;L^2(0,T;H))}^2 \leq M, \tag{25}
$$

$$
\varepsilon\|v_{\varepsilon\lambda}\|_{C^0([0,T];L^q(\Omega;V_0^*))}^q + \|v_{\varepsilon\lambda}\|_{L^q(\Omega;L^q(0,T;V_0^*))}^q \leq M, \tag{26}
$$

$$
\varepsilon^2\|G_{\varepsilon\lambda}\|_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))}^2 + \varepsilon\|G_{\varepsilon\lambda}\|_{L^q(\Omega;L^q(0,T;\mathscr{L}^q(U,V_0^*)))}^q \leq M. \tag{27}
$$

In particular, since $(v_{\varepsilon\lambda})$ is uniformly bounded in $L^q_{\mathscr{P}}(\Omega; L^q(0, T; V_0^*))$ by (26) and $(B + \varepsilon G_{\varepsilon\lambda})$ is uniformly bounded in $L^2_{\mathscr{P}}(\Omega; L^2(0, T; \mathscr{L}^2(U, H)))$ by (27), it follows from the definition of $u_{\varepsilon\lambda}$ itself in (12) that

$$\|u_{\varepsilon\lambda}\|^q_{L^q(\Omega; C^0([0, T]; V_0^*))} \leq M. \tag{28}$$

The boundedness of the operator A yields also

$$\|A_\lambda(u_{\varepsilon\lambda})\|^q_{L^q(\Omega; L^q(0, T; V^*))} \leq M. \tag{29}$$

Furthermore, following a classical argument employed in backward SPDEs, we can refine the estimate on $(v_{\varepsilon\lambda})$. Indeed, let us recall the already obtained Itô's formula for $v_{\varepsilon\lambda}$ in V_0^*, which reads

$$
\begin{aligned}
&\frac{\varepsilon}{2}\|v_{\varepsilon\lambda}(t)\|^q_{V_0^*} + \frac{q}{2}\int_t^T \left(\|v_{\varepsilon\lambda}(s)\|^q_{V_0^*} + \frac{\varepsilon}{2} 1_{\{\|v_{\varepsilon\lambda}(s)\|_{V_0^*}>0\}}\|v_{\varepsilon\lambda}(s)\|^{q-2}_{V_0^*}\|G_{\varepsilon\lambda}(s)\|^2_{\mathscr{L}^2(U, V_0^*)} \right) \, ds \\
&\quad + \frac{q}{2}\frac{q-2}{2}\varepsilon \int_t^T 1_{\{\|v_{\varepsilon\lambda}(s)\|_{V_0^*}>0\}}\|v_{\varepsilon\lambda}(s)\|^{q-4}_{V_0^*}\|(R_0^{-1}v_{\varepsilon\lambda}(s), G_{\varepsilon\lambda}(s))\|^2_{\mathscr{L}^2(U, \mathbb{R})} \, ds \\
&= -\frac{q}{2}\int_t^T 1_{\{\|v_{\varepsilon\lambda}(s)\|_{V_0^*}>0\}}\|v_{\varepsilon\lambda}(s)\|^{q-2}_{V_0^*}\left(A_\lambda(u_{\varepsilon\lambda}(s)), R_0^{-1}(v_{\varepsilon\lambda}(s))\right) \, ds \\
&\quad - \frac{q}{2}\varepsilon \int_t^T 1_{\{\|v_{\varepsilon\lambda}(s)\|_{V_0^*}>0\}}\|v_{\varepsilon\lambda}(s)\|^{q-2}_{V_0^*}\left(R_0^{-1}(v_{\varepsilon\lambda}(s)), G_{\varepsilon\lambda}(s) \, dW(s)\right).
\end{aligned}
$$

Instead of taking expectations at t fixed, we can now take supremum in time and then expectations. The first term on the right-hand side can be easily bounded using the Hölder inequality and the estimates (26) and (29) as

$$
\begin{aligned}
&\mathbb{E}\int_0^T 1_{\{\|v_{\varepsilon\lambda}(s)\|_{V_0^*}>0\}}\|v_{\varepsilon\lambda}(s)\|^{q-2}_{V_0^*}\left(A_\lambda(u_{\varepsilon\lambda}(s)), R_0^{-1}(v_{\varepsilon\lambda}(s))\right) \, ds \\
&\leq c_0 \mathbb{E}\int_0^T \|v_{\varepsilon\lambda}(s)\|^{q-1}_{V_0^*}\|A_\lambda(u_{\varepsilon\lambda}(s))\|_{V^*} \, ds \\
&\leq c_0 \|v_{\varepsilon\lambda}\|^{q-1}_{L^q(\Omega; L^q(0, T; V_0^*))}\|A_\lambda(u_{\varepsilon\lambda})\|_{L^q(\Omega; L^q(0, T; V^*))} \leq M.
\end{aligned}
$$

The second term on the right-hand side can be bounded, thanks to Burkholder-Davis-Gundy and Young inequalities, as

$$
\begin{aligned}
&\mathbb{E}\sup_{t\in[0, T]}\left|\int_t^T 1_{\{\|v_{\varepsilon\lambda}(s)\|_{V_0^*}>0\}}\|v_{\varepsilon\lambda}(s)\|^{q-2}_{V_0^*}(R_0^{-1}(v_{\varepsilon\lambda}(s)), G_{\varepsilon\lambda}(s) \, dW(s))\right| \\
&\leq M\mathbb{E}\left(\int_0^T 1_{\{\|v_{\varepsilon\lambda}(s)\|_{V_0^*}>0\}}\|v_{\varepsilon\lambda}(s)\|^{2(q-2)}_{V_0^*}\|(R_0^{-1}(v_{\varepsilon\lambda}(s)), G_{\varepsilon\lambda}(s))\|^2_{\mathscr{L}^2(U, \mathbb{R})} \, ds\right)^{1/2} \\
&\leq M\mathbb{E}\left[\|v_{\varepsilon\lambda}\|^{q/2}_{C^0([0, T]; V_0^*)}\left(\int_0^T 1_{\{\|v_{\varepsilon\lambda}(s)\|_{V_0^*}>0\}}\|v_{\varepsilon\lambda}(s)\|^{q-4}_{V_0^*}\|(R_0^{-1}(v_{\varepsilon\lambda}(s)), G_{\varepsilon\lambda}(s))\|^2_{\mathscr{L}^2(U, \mathbb{R})} \, ds\right)^{1/2}\right] \\
&\leq \sigma\mathbb{E}\|v_{\varepsilon\lambda}\|^q_{C^0([0, T]; V_0^*)} \\
&\quad + \frac{M^2}{4\sigma}\mathbb{E}\int_0^T 1_{\{\|v_{\varepsilon\lambda}(s)\|_{V_0^*}>0\}}\|v_{\varepsilon\lambda}(s)\|^{q-4}_{V_0^*}\|(R_0^{-1}(v_{\varepsilon\lambda}(s)), G_{\varepsilon\lambda}(s))\|^2_{\mathscr{L}^2(U, \mathbb{R})} \, ds
\end{aligned}
$$

for every $\sigma > 0$ (independent of λ and ε). Hence, choosing σ sufficiently small (for example $\sigma := q/2$), rearranging the terms, and using the Hölder inequality yields

$$\varepsilon \mathbb{E}\|v_{\varepsilon\lambda}\|^q_{C^0([0,T];V_0^*)} \leq M\left(1 + \varepsilon\mathbb{E}\int_0^T 1_{\{\|v_{\varepsilon\lambda}(s)\|_{V_0^*}>0\}}\|v_{\varepsilon\lambda}(s)\|^{q-2}_{V_0^*}\|G_{\varepsilon\lambda}(s)\|^2_{\mathscr{L}^2(U,V_0^*)} \ ds\right).$$

Now, note that the right-hand side is uniformly bounded in λ and ε thanks to the inequality (23) and the already proved estimate (25). Consequently, we deduce that

$$\varepsilon\|v_{\varepsilon\lambda}\|^q_{L^q(\Omega;C^0([0,T];V_0^*))} \leq M. \tag{30}$$

Moreover, from inequality (23), since the function $r \mapsto |r|^{q-2}$, $r > 0$, is decreasing, using again the reverse Young inequality and the estimate (25) we deduce that

$$\begin{aligned}
M &\geq \varepsilon\mathbb{E}\int_0^T 1_{\{\|v_{\varepsilon\lambda}(s)\|_{V_0^*}>0\}}\|v_{\varepsilon\lambda}(s)\|^{q-2}_{V_0^*}\|G_{\varepsilon\lambda}(s)\|^2_{\mathscr{L}^2(U,V_0^*)} \ ds \\
&\geq \varepsilon\mathbb{E}\left[\left\|1_{\{\|v_{\varepsilon\lambda}\|_{V_0^*}>0\}}v_{\varepsilon\lambda}\right\|^{q-2}_{C^0([0,T];V_0^*)}\|G_{\varepsilon\lambda}\|^2_{L^2(0,T;\mathscr{L}^2(U,V_0^*))}\right] \\
&\geq \frac{2}{q}\varepsilon\mathbb{E}\|G_{\varepsilon\lambda}\|^q_{L^2(0,T;\mathscr{L}^2(U,V_0^*))} - \frac{q}{2-q}\varepsilon\mathbb{E}\|v_{\varepsilon\lambda}\|^q_{C^0([0,T];V_0^*)}.
\end{aligned}$$

Hence, estimate (30) readily implies also

$$\varepsilon\|G_{\varepsilon\lambda}\|^q_{L^q(\Omega;L^2(0,T;\mathscr{L}^2(U,V_0^*)))} \leq M. \tag{31}$$

3.4. Passage to the limit as $\lambda \searrow 0$

We pass now to the limit as $\lambda \searrow 0$, keeping $\varepsilon > 0$ fixed, and deduce existence of solutions for the regularized problem (11).

The estimates (25)–(31) imply that there exist $(u_\varepsilon, \hat{u}_\varepsilon, v_\varepsilon, \xi_\varepsilon, G_\varepsilon)$ such that, as $\lambda \searrow 0$,

$$\begin{aligned}
u_{\varepsilon\lambda} &\overset{*}{\rightharpoonup} u_\varepsilon &&\text{in } L^q(\Omega;L^\infty(0,T;V_0^*)), \\
J_\lambda(u_{\varepsilon\lambda}) &\rightharpoonup \hat{u}_\varepsilon &&\text{in } L^p(\Omega;L^p(0,T;V)), \\
v_{\varepsilon\lambda} &\overset{*}{\rightharpoonup} v_\varepsilon &&\text{in } L^q(\Omega;L^\infty(0,T;V_0^*)) \cap L^2(\Omega;L^2(0,T;H)), \\
A_\lambda(u_{\varepsilon\lambda}) &\rightharpoonup \xi_\varepsilon &&\text{in } L^q(\Omega;L^q(0,T;V^*)), \\
G_{\varepsilon\lambda} &\rightharpoonup G_\varepsilon &&\text{in } L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H))).
\end{aligned}$$

Note that by the definition of Yosida approximation and estimate (25) we have

$$\|u_{\varepsilon\lambda} - J_\lambda(u_{\varepsilon\lambda})\|_{L^2(\Omega;L^2(0,T;H))} = \lambda\|A_\lambda(u_{\varepsilon\lambda})\|_{L^2(\Omega;L^2(0,T;H))} \leq M\lambda^{1/2} \to 0,$$

which implies that $\hat{u}_\varepsilon = u_\varepsilon$. Moreover, by letting $\lambda \searrow 0$ in the forward equation in (12), we get

$$u_\varepsilon = u_{0,\varepsilon} + \int_0^\cdot v_\varepsilon(s) \ ds + \int_0^\cdot (B + \varepsilon G_\varepsilon)(s) \ dW(s),$$

yielding, a posteriori, also that $u_\varepsilon \in L^2(\Omega;C^0([0,T];H))$. Similarly, letting $\lambda \searrow 0$ in the backward equation in (12) we obtain, by the weak convergences above,

$$\varepsilon v_\varepsilon + \int_.^T v_\varepsilon(s) \; ds + \int_.^T \xi_\varepsilon(s) \; ds = \varepsilon \int_.^T G_\varepsilon(s) \; dW(s),$$

which yields a posteriori that $v_\varepsilon \in L^q\big(\Omega; C^0\big([0,T]; V_0^*\big)\big)$. Furthermore, by comparison in the equation (12) it follows in particular that

$$u_{\varepsilon\lambda}(T) \rightharpoonup u_\varepsilon(T) \quad \text{in } L^2(\Omega, \mathscr{F}_T; H), \qquad v_{\varepsilon\lambda}(0) \rightharpoonup v_\varepsilon(0) \quad \text{in } L^q\big(\Omega, \mathscr{F}_0; V_0^*\big).$$

It only remains to show that $\xi_\varepsilon \in A(\cdot, u_\varepsilon)$ almost everywhere. To this end, we recall that by comparison of (20) and (21) we have

$$\frac{1}{2}\mathbb{E}\|u_{\varepsilon\lambda}(T)\|_H^2 + \varepsilon\mathbb{E}\int_0^T \|v_{\varepsilon\lambda}(s)\|_H^2 \; ds$$

$$+ \mathbb{E}\int_0^T (A_\lambda(u_{\varepsilon\lambda}(s)), u_{\varepsilon\lambda}(s)) \; ds + \frac{\varepsilon^2}{2}\mathbb{E}\int_0^T \|G_{\varepsilon\lambda}(s)\|_{\mathscr{L}^2(U,H)}^2 \; ds$$

$$\leq \frac{1}{2}\mathbb{E}\|u_{0,\varepsilon}\|_H^2 - \varepsilon\mathbb{E}\big(v_{\varepsilon\lambda}(0), u_{0,\varepsilon}\big) + \frac{1}{2}\mathbb{E}\int_0^T \|B(s)\|_{\mathscr{L}^2(U;H)}^2 \; ds.$$

By the weak lower semicontinuity of the norms and the regularities of the data B_ε and $u_{0,\varepsilon}$ in condition (8) we infer then that

$$\limsup_{\lambda\searrow 0}\mathbb{E}\int_0^T (A_\lambda(u_{\varepsilon\lambda}(s)), u_{\varepsilon\lambda}(s)) ds$$

$$= \frac{1}{2}\mathbb{E}\|u_{0,\varepsilon}\|_H^2 - \varepsilon\mathbb{E}\big(v_\varepsilon(0), u_{0,\varepsilon}\big) + \frac{1}{2}\mathbb{E}\int_0^T \|B(s)\|_{\mathscr{L}^2(U;H)}^2 \; ds$$

$$- \frac{1}{2}\liminf_{\lambda\searrow 0}\mathbb{E}\|u_{\varepsilon\lambda}(T)\|_H^2 - \varepsilon\liminf_{\lambda\searrow 0}\mathbb{E}\int_0^T \|v_{\varepsilon\lambda}(s)\|_H^2 \; ds$$

$$- \frac{\varepsilon^2}{2}\liminf_{\lambda\searrow 0}\mathbb{E}\int_0^T \|G_{\varepsilon\lambda}(s)\|_{\mathscr{L}^2(U,H)}^2 \; ds \tag{32}$$

$$\leq \frac{1}{2}\mathbb{E}\|u_{0,\varepsilon}\|_H^2 - \varepsilon\mathbb{E}\big(v_\varepsilon(0), u_{0,\varepsilon}\big) + \frac{1}{2}\mathbb{E}\int_0^T \|B(s)\|_{\mathscr{L}^2(U;H)}^2 \; ds$$

$$- \frac{1}{2}\mathbb{E}\|u_\varepsilon(T)\|_H^2 - \varepsilon\mathbb{E}\int_0^T \|v_\varepsilon(s)\|_H^2 \; ds - \frac{\varepsilon^2}{2}\mathbb{E}\int_0^T \|G_\varepsilon(s)\|_{\mathscr{L}^2(U,H)}^2 \; ds.$$

We claim now that the right-hand side of inequality (32) coincides with

$$\mathbb{E}\int_0^T \langle \xi_\varepsilon(s), u_\varepsilon(s)\rangle \; ds.$$

In order to show this, we replicate in the limit $\lambda = 0$ the Itô's formulas obtained for $\lambda > 0$ in (20) and (21).

Indeed, the Itô formula for the square of the H-norm of u_ε yields

$$\frac{1}{2}\mathbb{E}\|u_\varepsilon(T)\|_H^2 = \frac{1}{2}\mathbb{E}\|u_{0,\varepsilon}\|_H^2 + \mathbb{E}\int_0^T (v_\varepsilon(s), u_\varepsilon(s)) \; ds + \frac{1}{2}\mathbb{E}\int_0^T \|(B + \varepsilon G_\varepsilon)(s)\|_{\mathscr{L}^2(U,H)}^2 \; ds,$$

while Itô's formula for $(u_\varepsilon, \varepsilon v_\varepsilon)$ yields

$$\varepsilon \mathbb{E} \int_0^T \|v_\varepsilon(s)\|_H^2 \, ds + \mathbb{E} \int_0^T (v_\varepsilon(s), u_\varepsilon(s)) \, ds + \mathbb{E} \int_0^T \langle \xi_\varepsilon(s), u_\varepsilon(s) \rangle \, ds$$

$$= \varepsilon^2 \mathbb{E} \int_0^T \|G_\varepsilon(s)\|_{\mathscr{L}^2(U,H)}^2 \, ds + \varepsilon \mathbb{E} \int_0^T (B(s), G_\varepsilon(s))_{\mathscr{L}^2(U;H)} ds - \varepsilon \mathbb{E} \langle v_\varepsilon(0), u_{0,\varepsilon} \rangle.$$

By comparison we infer exactly that

$$\frac{1}{2} \mathbb{E} \|u_\varepsilon(T)\|_H^2 + \varepsilon \mathbb{E} \int_0^T \|v_\varepsilon(s)\|_H^2 \, ds + \mathbb{E} \int_0^T \langle \xi_\varepsilon(s), u_\varepsilon(s) \rangle \, ds + \frac{\varepsilon^2}{2} \mathbb{E} \int_0^T \|G_\varepsilon(s)\|_{\mathscr{L}^2(U,H)}^2 \, ds$$

$$= \frac{1}{2} \mathbb{E} \|u_{0,\varepsilon}\|_H^2 - \varepsilon \mathbb{E} \langle v_\varepsilon(0), u_{0,\varepsilon} \rangle + \frac{1}{2} \mathbb{E} \int_0^T \|B(s)\|_{\mathscr{L}^2(U;H)}^2 \, ds,$$

(33)

as required. Substituting now this expression in the inequality (32), we get

$$\limsup_{\lambda \searrow 0} \mathbb{E} \int_0^T \left(A_\lambda(u_{\varepsilon\lambda}(s)), u_{\varepsilon\lambda(s)} \right) \, ds \leq \mathbb{E} \int_0^T \langle \xi_\varepsilon(s), u_\varepsilon(s) \rangle \, ds.$$

The maximal monotonicity of A implies then that $\xi_\varepsilon \in A(\cdot, u_\varepsilon)$ almost everywhere, see [62, Prop. 2.5, p. 27]. Hence, $(u_\varepsilon, \xi_\varepsilon, v_\varepsilon, G_\varepsilon)$ is a solution to (11) in the sense of Theorem 2.1.ii.

3.5. Uniqueness

Let us check that the quadruplet $(u_\varepsilon, \xi_\varepsilon, v_\varepsilon, G_\varepsilon)$ is unique. Assume that $(u_\varepsilon^i, \xi_\varepsilon^i, v_\varepsilon^i, G_\varepsilon^i)$, for $i = 1, 2$, solve (11) in the sense of Theorem 2.1.ii. Then, we have

$$\begin{cases} d(u_\varepsilon^1 - u_\varepsilon^2) = (v_\varepsilon^1 - v_\varepsilon^2) \, dt + \varepsilon(G_\varepsilon^1 - G_\varepsilon^2) \, dW \\ (u_\varepsilon^1 - u_\varepsilon^2)(0) = 0, \end{cases}$$

and

$$\begin{cases} -\varepsilon d(v_\varepsilon^1 - v_\varepsilon^2) + (v_\varepsilon^1 - v_\varepsilon^2) \, dt + (\xi_\varepsilon^1 - \xi_\varepsilon^2) \, dt = -\varepsilon(G_\varepsilon^1 - G_\varepsilon^2) \, dW \\ (v_\varepsilon^1 - v_\varepsilon^2)(T) = 0. \end{cases}$$

Using the same argument employed to deduce (33), we infer that

$$\frac{1}{2} \mathbb{E} \|(u_\varepsilon^1 - u_\varepsilon^2)(T)\|_H^2 + \varepsilon \mathbb{E} \int_0^T \|(v_\varepsilon^1 - v_\varepsilon^2)(s)\|_H^2 \, ds$$

$$+ \mathbb{E} \int_0^T \langle (\xi_\varepsilon^1 - \xi_\varepsilon^2)(s), u_\varepsilon(s) \rangle \, ds + \frac{\varepsilon^2}{2} \mathbb{E} \int_0^T \|(G_\varepsilon^1 - G_\varepsilon^2)(s)\|_{\mathscr{L}^2(U,H)}^2 \, ds = 0,$$

which implies that $v_\varepsilon^1 - v_\varepsilon^2 = 0$ and $G_\varepsilon^1 - G_\varepsilon^2 = 0$ by the monotonicity of A. From the forward equation we deduce that $u_\varepsilon^1 - u_\varepsilon^2 = 0$. By comparison in the backward equation we find that $\xi_\varepsilon^1 - \xi_\varepsilon^2 = 0$, as required. This completes the proof of well-posedness in Theorem 2.1.ii.

4. The asymptotics as $\varepsilon \searrow 0$ of the forward-backward problem

The aim of this section is to show that the solution of the ε-regularized forward-backward problem (11) converges to the solution of the nonregularized problem (3).

First of all, note that the estimates (25)–(31) are independent of ε. Hence, by weak lower semicontinuity we deduce that

$$\|u_\varepsilon\|^q_{L^q\left(\Omega;C^0\left([0,T];V_0^*\right)\right)} + \|u_\varepsilon\|^p_{L^p(\Omega;L^p(0,T;V))} \leq M,$$

$$\varepsilon\|v_\varepsilon\|^q_{L^q\left(\Omega;C^0\left([0,T];V_0^*\right)\right)} + \varepsilon\|v_\varepsilon\|^2_{L^2(\Omega;L^2(0,T;H))} + \|v_\varepsilon\|^q_{L^q\left(\Omega;L^q\left(0,T;V_0^*\right)\right)} \leq M,$$

$$\|\xi_\varepsilon\|^q_{L^q(\Omega;L^q(0,T;V^*))} \leq M,$$

$$\varepsilon^2\|G_\varepsilon\|^2_{L^2\left(\Omega;L^2\left(0,T;\mathscr{L}^2(U,H)\right)\right)} + \varepsilon\|G_\varepsilon\|^q_{L^q\left(\Omega;L^2\left(0,T;\mathscr{L}^2\left(U,V_0^*\right)\right)\right)} \leq M.$$

Moreover, thanks also to assumption (9) and [63, Lem. 2.1], we have that

$$\|u_\varepsilon\|_{L^q_{\mathscr{P}}\left(\Omega;W^{s,q}\left(0,T;V_0^*\right)\right)} \leq M_s \qquad \forall\, s \in (0,1/2).$$

We deduce that there exist

$$u \in L^q\left(\Omega;W^{s,q}\left(0,T;V_0^*\right)\right) \cap L^q_{\mathscr{P}}\left(\Omega;L^\infty\left(0,T;V_0^*\right)\right) \cap L^p_{\mathscr{P}}\left(\Omega;L^p(0,T;V)\right),$$

$$v \in L^q_{\mathscr{P}}\left(\Omega;L^q\left(0,T;V_0^*\right)\right), \qquad \xi \in L^q_{\mathscr{P}}(\Omega;L^q(0,T;V^*))$$

such that, as $\varepsilon \searrow 0$,

$$u_\varepsilon \overset{*}{\rightharpoonup} u \quad \text{in } L^q\left(\Omega;W^{s,q}\left(0,T;V_0^*\right)\right) \cap L^q(\Omega;L^\infty(0,T;H)) \cap L^p\left(\Omega;L^p(0,T;V)\right),$$

$$v_\varepsilon \rightharpoonup v \quad \text{in } L^q\left(\Omega;L^q\left(0,T;V_0^*\right)\right),$$

$$\xi_\varepsilon \rightharpoonup \xi \quad \text{in } L^q(\Omega;L^q(0,T;V^*)).$$

Moreover, note that

$$\|\varepsilon v_\varepsilon\|_{L^q\left(\Omega;C^0\left([0,T];V_0^*\right)\right)} = \varepsilon^{1/p}\varepsilon^{1/q}\|v_\varepsilon\|_{L^q\left(\Omega;C^0\left([0,T];V_0^*\right)\right)} \leq \varepsilon^{1/p}M^{1/q} \to 0,$$

$$\|\varepsilon G_\varepsilon\|_{L^q\left(\Omega;L^2\left(0,T;\mathscr{L}^2\left(U,V_0^*\right)\right)\right)} = \varepsilon^{1/p}\varepsilon^{1/q}\|G_\varepsilon\|_{L^q\left(\Omega;L^2\left(0,T;\mathscr{L}^2\left(U,V_0^*\right)\right)\right)} \leq \varepsilon^{1/p}M^{1/q} \to 0,$$

so that, by the Burkholder-Davis-Gaundy inequality,

$$\varepsilon v_\varepsilon \to 0 \quad \text{in } L^q\left(\Omega;C^0\left([0,T];V_0^*\right)\right),$$

$$\varepsilon G_\varepsilon \to 0 \quad \text{in } L^q\left(\Omega;L^2\left(0,T;\mathscr{L}^2\left(U,V_0^*\right)\right)\right),$$

$$\varepsilon G_\varepsilon \cdot W \to 0 \quad \text{in } L^q\left(\Omega;C^0\left([0,T];V_0^*\right)\right).$$

Then, by passing to the weak limit as $\varepsilon \searrow 0$ in the backward equation in (11) yields

$$\int_t^T v(s)\ \mathrm{d}s + \int_t^T \xi(s)\ \mathrm{d}s = 0 \qquad \forall t \in [0,T], \quad \mathbb{P}\text{-a.s.},$$

from which $v + \xi = 0$ almost everywhere in $\Omega \times (0,T)$. In particular, we have that

$$v = -\xi \in L^q_{\mathscr{P}}(\Omega;L^q(0,T;V^*)).$$

Furthermore, recalling the convergences (9) on the data and passing to the weak limit

in the forward equation in (11) we have that

$$u = u_0 + \int_0^{\cdot} v(s) \ \mathrm{d}s + \int_0^{\cdot} B(s) \ \mathrm{d}W(s).$$

In particular, since $v \in L^q_{\mathscr{P}}(\Omega; L^q(0, T; V^*))$ and $u \in L^p_{\mathscr{P}}(\Omega; L^p(0, T; V))$, by the classical Itô's formula (see [64]) we deduce by comparison that $u \in L^2\big(\Omega; C^0([0, T]; H)\big)$, while from (33) we have

$$u_\varepsilon(T) \rightharpoonup u(T) \quad \text{in } L^2(\Omega; H).$$

Eventually, let us show that $\xi \in A(\cdot, u)$ almost everywhere. This follows again by lower-semicontinuity arguments. In particular, from (33), the weak lower semicontinuity of the norms, and the convergences (9) we have that

$$\limsup_{\varepsilon \searrow 0} \mathbb{E} \int_0^T \langle \xi_\varepsilon(s), u_\varepsilon(s) \rangle \ \mathrm{d}s$$

$$\leq -\frac{1}{2} \liminf_{\varepsilon \searrow 0} \mathbb{E} \|u_\varepsilon(T)\|_H^2 + \frac{1}{2} \mathbb{E} \|u_0\|_H^2 + \frac{1}{2} \mathbb{E} \int_0^T \|B(s)\|_{\mathscr{L}^2(U;H)}^2 \ \mathrm{d}s - \limsup_{\varepsilon \to 0} \varepsilon \mathbb{E}\big(v_\varepsilon(0), u_{0,\varepsilon}\big)$$

$$\leq -\frac{1}{2} \mathbb{E} \|u(T)\|_H^2 + \frac{1}{2} \mathbb{E} \|u_0\|_H^2 + \frac{1}{2} \mathbb{E} \int_0^T \|B(s)\|_{\mathscr{L}^2(U;H)}^2 \ \mathrm{d}s$$

$$+ \limsup_{\varepsilon \searrow 0} \varepsilon \|v_\varepsilon(0)\|_{L^q(\Omega; V_0^*)} \|u_{0,\varepsilon}\|_{L^p(\Omega; V_0)}.$$

The last term on the right-hand side can be handled using the estimates above and the condition (9) as

$$\varepsilon \|v_\varepsilon(0)\|_{L^q(\Omega; V_0^*)} \|u_{0,\varepsilon}\|_{L^p(\Omega; V_0)} \leq M^{1/q} \varepsilon^{1/p} \|u_{0,\varepsilon}\|_{L^p(\Omega; V_0)} \to 0.$$

Hence, we infer that

$$\limsup_{\varepsilon \searrow 0} \mathbb{E} \int_0^T \langle \xi_\varepsilon(s), u_\varepsilon(s) \rangle \ \mathrm{d}s \leq -\frac{1}{2} \mathbb{E} \|u(T)\|_H^2 + \frac{1}{2} \mathbb{E} \|u_0\|_H^2 + \frac{1}{2} \mathbb{E} \int_0^T \|B(s)\|_{\mathscr{L}^2(U;H)}^2 \ \mathrm{d}s.$$

Now, since by the Itô formula for u and the fact that $\xi = -v$ we know that

$$\frac{1}{2} \mathbb{E} \|u(T)\|_H^2 + \mathbb{E} \int_0^T \langle \xi(s), u(s) \rangle \ \mathrm{d}s = \frac{1}{2} \mathbb{E} \|u_0\|_H^2 + \frac{1}{2} \mathbb{E} \int_0^T \|B(s)\|_{\mathscr{L}^2(U;H)}^2 \ \mathrm{d}s,$$

we obtain

$$\limsup_{\varepsilon \searrow 0} \mathbb{E} \int_0^T \langle \xi_\varepsilon(s), u_\varepsilon(s) \rangle \ \mathrm{d}s \leq \mathbb{E} \int_0^T \langle \xi(s), u(s) \rangle \ \mathrm{d}s.$$

This yields $\xi \in A(\cdot, u)$ almost everywhere in $\Omega \times (0, T)$, and the first part of Theorem 2.1.iii is proved.

We only need to show the strong convergences in the last assertion of Theorem 2.1.iii, under the extra assumption that $V \hookrightarrow H$ is compact and $p < 4$. The idea is to use the following classical result by Gyöngy and Krylov [65, Lem. 1.1].

Lemma 4.1. *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 $(Z_{n_{k_i}}, Z_{n_{j_i}})_i$ converging in law to a probability measure ν on $\mathcal{X} \times \mathcal{X}$ such that $\nu(\{(z_1, z_2) \in \mathcal{X} \times \mathcal{X} : z_1 = z_2\}) = 1$.*

Let then $(u_{\varepsilon_k})_k$ and $(u_{\varepsilon_j})_j$ be arbitrary subsequences of (u_ε). By the compactness result [66, Cor. 5, p. 86] we have the compact inclusion

$$L^p(0, T; V) \cap W^{s,q}(0, T; V_0^*) \overset{c}{\hookrightarrow} L^p(0, T; H)$$

provided that $s > 1/q - 1/p = 1 - 2/p$. Since $p < 4$ by assumption, an easy computation shows that $1 - 2/p < 1/2$: hence there exists $\bar{s} \in (0, 1/2)$ such that the compact inclusion holds. Now, from the estimates we know that

$$\|u_\varepsilon\|_{L^p\left(\Omega; L^p(0,T;V) \cap W^{\bar{s},q}(0,T;V_0^*)\right)} \leq M,$$

which implies, using a standard argument based on the Markov inequality, that the family of laws of $(u_\varepsilon)_\varepsilon$ on $L^p(0, T; H)$ is tight. By the Skorokhod theorem [67, Thm. 2.7], there exists a probability space $(\Omega', \mathscr{F}', \mathbb{P}')$ and measurable functions $\phi_\varepsilon : (\Omega', \mathscr{F}') \to (\Omega, \mathscr{F})$ such that $\mathbb{P}' \circ \phi_\varepsilon^{-1} = \mathbb{P}$ for all $\varepsilon > 0$ and

$$u'_{\varepsilon_{k_i}} := u_{\varepsilon_{k_i}} \circ \phi_{\varepsilon_{k_i}} \to u'_1 \quad \text{in } L^p(0, T; H), \quad \mathbb{P}'\text{-a.s.},$$

$$u'_{\varepsilon_{j_i}} := u_{\varepsilon_{j_i}} \circ \phi_{\varepsilon_{j_i}} \to u'_2 \quad \text{in } L^p(0, T; H), \quad \mathbb{P}'\text{-a.s..}$$

Relying on the uniform estimates proved above and on the uniqueness of the limit problem, it is not difficult to show that

$$\mathbb{P}'\{u'_1(t) = u'_2(t) \forall\ t \in [0, T]\} = 1,$$

which is exactly the condition of Lemma 4.1. For further details we refer for example to [68, § 5] and [69]. Hence, the Lemma implies that, on the original probability space, we have

$$u_\varepsilon \to u \quad \text{in } L^p(0, T; H), \quad \mathbb{P}\text{-a.s.}$$

As $(u_\varepsilon)_\varepsilon$ is bounded in $L^p(\Omega; L^p(0, T; H))$, this yields

$$u_\varepsilon \to u \quad \text{in } L^r(\Omega; L^p(0, T; H)) \qquad \forall\ r \in [1, p).$$

This completes the proof of Theorem 2.1.iii.

5. Equivalence between regularized equation and minimization problem

This section is devoted to check that I_ε admits a unique minimizer in $\mathcal{V}$, and that this coincides with the unique solution to the ε-regularized problem. This proves Theorem 2.1.i. In all of this section $\varepsilon > 0$ is kept fixed.

A natural idea would be to identify the subdifferential of I_ε in terms of $\partial(I_\varepsilon^1 + S_\varepsilon)$ and ∂I_ε^2. However, let us point out that the domain of I_ε^2, i.e. the space $L_{\mathscr{P}}^p(\Omega; L^p(0, T; V))$, may have empty interior in the topology of $\mathcal{I}^{2,2}(H, H)$. For this reason, the analogous of [57, Thm. 2.10] is not applicable in this case, and we need to rely

again on a further approximation, obtained by replacing Φ with its Moreau-Yosida approximation Φ_λ, for $\lambda > 0$.

We follow the following strategy instead. First of all, we show that the unique solution u_ε to problem (10) is a minimizer for I_ε. This ensures in particular that I_ε admits at least a minimizer. Secondly, we note that actually I_ε admits at most one minimizer. This eventually entails that minimizing I_ε is equivalent to solving (10).

Proposition 5.1. The *unique solution u_ε to (10) is a minimizer for I_ε.*

Proof. From Section 3 we know that u_ε can be constructed as limit in suitable topologies of a sequence $(u_{\varepsilon\lambda})_{\lambda>0}$, where $u_{\varepsilon\lambda}$ is the unique first solution component of (12). By Proposition 3.4 we also know that such $u_{\varepsilon\lambda}$ is the unique global minimizer of $I_{\varepsilon\lambda}$ for all $\lambda > 0$, so that

$$I_{\varepsilon\lambda}(u_{\varepsilon\lambda}) \le I_{\varepsilon\lambda}(z) \qquad \forall\, z \in \mathcal{I}^{2,2}(H, H). \tag{34}$$

Let us now consider $z \in D(I_\varepsilon) = \mathcal{V}$: since $\Phi_\lambda \le \Phi$, we immediately have

$$I_{\varepsilon\lambda}(z) \le I_\varepsilon(z) \qquad \forall\, z \in \mathcal{V}.$$

Furthermore, by Section 3 we know that

$$J_\lambda(u_{\varepsilon\lambda}) \rightharpoonup u_\varepsilon \quad \text{in } L^p\big(\Omega; L^p(0, T; V)\big),$$
$$v_{\varepsilon\lambda} \rightharpoonup v_\varepsilon \quad \text{in } L^2\big(\Omega; L^2(0, T; H)\big).$$

Hence, by the definition of Φ_λ, the weak lower semicontinuity of Φ, and by the Fatou lemma, we have

$$\mathbb{E} \int_0^T e^{-t/\varepsilon} \Phi(t, u_\varepsilon(t)) \ dt \le \liminf_{\lambda \searrow 0} \mathbb{E} \int_0^T e^{-t/\varepsilon} \Phi(t, J_\lambda(u_{\varepsilon\lambda}(t))) \ dt$$
$$\le \liminf_{\lambda \searrow 0} \mathbb{E} \int_0^T e^{-t/\varepsilon} \Phi_\lambda(t, u_{\varepsilon\lambda}(t)) \ dt$$

and

$$\mathbb{E} \int_0^T e^{-t/\varepsilon} \frac{\varepsilon}{2} \|v_\varepsilon(t)\|^2 \ dt \le \liminf_{\lambda \searrow 0} \mathbb{E} \int_0^T e^{-t/\varepsilon} \frac{\varepsilon}{2} \|v_{\varepsilon\lambda}(t)\|^2 \ dt.$$

Taking these remarks into account, and recalling that $v_{\varepsilon\lambda} = \partial_t u_{\varepsilon\lambda}^d$, $v_\varepsilon = \partial_t u_\varepsilon^d$, and $u_{\varepsilon\lambda}^s = u_\varepsilon^s = B_\varepsilon$, we have

$$I_\varepsilon(u_\varepsilon) = \mathbb{E} \int_0^T e^{-t/\varepsilon} \left[\frac{\varepsilon}{2} \|v_\varepsilon(t)\|^2 + \Phi(t, u_\varepsilon(t)) \right] \ dt$$
$$\le \liminf_{\lambda \searrow 0} \mathbb{E} \int_0^T e^{-t/\varepsilon} \left[\frac{\varepsilon}{2} \|v_{\varepsilon\lambda}(t)\|^2 + \Phi_\lambda(t, u_{\varepsilon\lambda}(t)) \right] \ dt = \liminf_{\lambda \searrow 0} I_{\varepsilon\lambda}(u_{\varepsilon\lambda}).$$

Passing then to the lim inf in (34) yields then

$$I_\varepsilon(u_\varepsilon) \le I_\varepsilon(z) \qquad \forall\, z \in \mathcal{V},$$

hence u_ε is a global mininimizer of I_ε, as required. □

In order to conclude the proof of Theorem 2.1.i, note that the functional $I_\varepsilon^1 + S_\varepsilon$ is strictly convex and coercive on $\mathcal{I}^{2,2}(H,H)$, hence so is I_ε on $\mathcal{V}$ since Φ is convex and bounded from below. Since $\mathcal{V}$ is reflexive, we deduce that I_ε admits a unique global minimizer $z_\varepsilon \in \mathcal{V}$. Moreover, by virtue of Proposition 5.1, we know that the unique solution $u_\varepsilon \in \mathcal{U}_{reg}$ to (10) is a global minimizer of I_ε. By uniqueness of z_ε, we infer that $z_\varepsilon = u_\varepsilon \in \mathcal{U}_{reg}$. This concludes the proof of Theorem 2.1.i.

ORCID

Luca Scarpa (iD) http://orcid.org/0000-0001-6928-8944
Ulisse Stefanelli (iD) http://orcid.org/0000-0002-3969-2716

References

[1] Pardoux, E. *Equations aux derivées partielles stochastiques nonlinéaires monotones.* PhD thesis, Université Paris XI, (1975).

[2] Krylov, N. V., Rozovskiĭ, B. L. (1979). Stochastic evolution equations. In *Current Problems in Mathematics*, Vol. 14, pp. 71–147, 256. Moscow: Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Informatsii.

[3] Buttazzo, G. (1989). *Semicontinuity, Relaxation and Integral Representation in the Calculus of Variations, Volume 207 of Pitman Research Notes in Mathematics Series.* New York: Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc.

[4] Dacorogna, B. (2008). *Direct Methods in the Calculus of Variations, Volume 78 of Applied Mathematical Sciences*, 2nd ed. New York: Springer.

[5] Dal Maso, G. (1993). *An Introduction to Γ-Convergence, Volume 8 of Progress in Nonlinear Differential Equations and Their Applications.* Boston, MA: Birkhäuser Boston, Inc.

[6] Lions, J. L. (1963). Singular perturbations and some non linear boundary value problems. MRC Technical Summary Report 421. University of Wisconsin, Madison, Wisconsin.

[7] Kohn, J. J., Nirenberg, L. (1965). Non-coercive boundary value problems. *Comm. Pure Appl. Math.* 18(3):443–492. DOI: 10.1002/cpa.3160180305.

[8] Oleĭnik, O. A. (1964). On a problem of G. Fichera (Russian). *Dokl. Akad. Nauk SSSR.* 157:1297–1300.

[9] Lions, J. L. (1963). Équations différentielles opérationelles dans les espaces de Hilbert. *Centro Internazionale Matematico Estivo: Equazioni differenziali astratte*, n. 2. Rome: Cremonese,

[10] Lions, J. L. (1965). Sur certaines équations paraboliques non linéares (French). *Bul. Soc. Math. France* . 79:155–175. DOI: 10.24033/bsmf.1620.

[11] Lions, J. L., Magenes, E. (1972). *Non-Homogeneus Boundary Value Problems and Applications*, vol. 1. New York: Springer-Verlag.

[12] Barbu, V. (1975). Existence theorems for a class of two point boundary problems. *J. Differential Equations.* 17(1):236–257. DOI: 10.1016/0022-0396(75)90043-1.

[13] Ilmanen, T. (1994). Elliptic regularization and partial regularity for motion by mean curvature. *Memoirs of the AMS.* 108(520). DOI: 10.1090/memo/0520.

[14] Hirano, N. (1994). Existence of periodic solutions for nonlinear evolution equations in Hilbert spaces. *Proc. Amer. Math. Soc.* 120(1):185–192. DOI: 10.1090/S0002-9939-1994-1174494-8.

[15] Evans, L. C. (1998). *Partial Differential Equations, Volume 19 of Graduate Studies in Mathematics.* Providence, RI: American Mathematical Society.

[16] Akagi, G., Stefanelli, U. (2016). A variational principle for gradient flows of nonconvex energies. *J. Convex Anal.* 23(1):53–75.

[17] Bögelein, V., Duzaar, F., Marcellini, P. (2014). Existence of evolutionary variational solutions via the calculus of variations. *J. Differ. Equ.* 256(12):3912–3942. DOI: 10.1016/j.jde.2014.03.005.

[18] Conti, S., Ortiz, M. (2008). Minimum principles for the trajectories of systems governed by rate problems. *J. Mech. Phys. Solids.* 56(5):1885–1904. DOI: 10.1016/j.jmps.2007.11.006.

[19] Mielke, A., Stefanelli, U. (2011). Weighted energy-dissipation functionals for gradient flows. *Esaim: Cocv.* 17(1):52–85. DOI: 10.1051/cocv/2009043.

[20] Mielke, A., Ortiz, M. (2008). A class of minimum principles for characterizing the trajectories of dissipative systems. *Esaim: Cocv.* 14(3):494–516. DOI: 10.1051/cocv:2007064.

[21] Mielke, A., Stefanelli, U. (2008). A discrete variational principle for rate-independent evolution. *Adv. Calc. Var.* 1:399–431.

[22] Larsen, C. J., Ortiz, M., Richardson, C. L. (2009). Fracture paths from front kinetics: relaxation and rate independence. *Arch. Rational Mech. Anal.* . 193(3):539–583. DOI: 10.1007/s00205-009-0216-y.

[23] Akagi, G., Melchionna, S., Stefanelli, U. (2018). Weighted energy-dissipation approach to doubly nonlinear problems on the half line. *J. Evol. Equ.* 18(1):49–74. DOI: 10.1007/s00028-017-0390-6.

[24] Akagi, G., Stefanelli, U. (2010). A variational principle for doubly nonlinear evolution. *Appl. Math. Lett.* 23(9):1120–1124. DOI: 10.1016/j.aml.2010.04.047.

[25] Akagi, G., Stefanelli, U. (2011). Weighted energy-dissipation functionals for doubly nonlinear evolution. *J. Funct. Anal.* 260(9):2541–2578. DOI: 10.1016/j.jfa.2010.12.027.

[26] Akagi, G., Stefanelli, U. (2014). Doubly nonlinear evolution equations as convex minimization. *SIAM J. Math. Anal.* 46(3):1922–1945. DOI: 10.1137/13091909X.

[27] Melchionna, S. (2017). A variational approach to symmetry, monotonicity, and comparison for doubly-nonlinear equations. *J. Math. Anal. Appl.* 456(2):1303–1328. DOI: 10.1016/j.jmaa.2017.07.052.

[28] Akagi, G., Melchionna, S. (2018). Elliptic-regularization of nonpotential perturbations of doubly-nonlinear flows of nonconvex energies: A variational approach. *J. Convex Anal.* 25(3):861–898.

[29] Melchionna, S. (2017). A variational principle for nonpotential perturbations of gradient flows of nonconvex energies. *J. Differ. Equ.* 262(6):3737–3758. DOI: 10.1016/j.jde.2016.12.002.

[30] Liero, M., Melchionna, S. (2019). The weighted energy-dissipation principle and evolutionary Γ-convergence for doubly nonlinear problems. *Esaim: Cocv.* 25:36–38. DOI: 10.1051/cocv/2018023.

[31] Rossi, R., Savaré, G., Segatti, A., Stefanelli, U. (2019). Weighted energy-dissipation functionals for gradient flows in metric spaces. *J. Math. Pures Appl.* 127(9):1–6. DOI: 10.1016/j.matpur.2018.06.022.

[32] Rossi, R., Savaré, G., Segatti, A., Stefanelli, U. (2011). A variational principle for gradient flows in metric spaces. *C. R. Math. Acad. Sci. Paris.* 349:1224–1228.

[33] Segatti, A. (2013). A variational approach to gradient flows in metric spaces. *Boll. Unione Mat. Ital.* 6(3):765–780.

[34] Spadaro, E. N., Stefanelli, U. (2011). A variational view at mean curvature evolution for linear growth functionals. *J. Evol. Equ.* 11(4):793–809. DOI: 10.1007/s00028-011-0111-5.

[35] Davoli, E., Stefanelli, U. (2019). Dynamic perfect plasticity as convex minimization. *SIAM J. Math. Anal.* 51(2):672–730. DOI: 10.1137/17M1148864.

[36] Ortiz Schmidt, M. B., Stefanelli, U. (2018). A variational approach to Navier–Stokes. *Nonlinearity*. 31(12):5664–5682. DOI: 10.1088/1361-6544/aae722.

[37] De Giorgi, E. (1996). Conjectures concerning some evolution problems. *Duke Math. J.* 81(2):255–268. DOI: 10.1215/S0012-7094-96-08114-4.

[38] Serra, E., Tilli, P. (2012). Nonlinear wave equations as limits of convex minimization problems: Proof of a conjecture by De Giorgi. *Ann. Math.* 175(3):1551–1574. DOI: 10.4007/annals.2012.175.3.11.

[39] Stefanelli, U. (2011). The De Giorgi conjecture on elliptic regularization. *Math. Models Methods Appl. Sci.* 21(06):1377–1394. DOI: 10.1142/S0218202511005350.

[40] Liero, M., Stefanelli, U. (2013). A new minimum principle for Lagrangian mechanics. *J. Nonlinear Sci.* 23(2):179–204. DOI: 10.1007/s00332-012-9148-z.

[41] Liero, M., Stefanelli, U. (2013). Weighted inertia-dissipation-energy functionals for semi-linear equations. *Boll. Unione Mat. Ital.* 6(9):1–27.

[42] Serra, E., Tilli, P. (2016). A minimization approach to hyperbolic Cauchy problems. *J. Eur. Math. Soc.* 18(9):2019–2044. DOI: 10.4171/JEMS/637.

[43] Tentarelli, L. (2016). On the extensions of the De Giorgi approach to nonlinear hyperbolic equations. *Rend. Semin. Mat. Univ. Politec. Torino.* 74(2):151–160.

[44] Tentarelli, L., Tilli, P. (2018). De Giorgi's approach to hyperbolic Cauchy problems: The case of nonhomogeneous equations. *Comm. Partial Differ. Equ.* 43(4):677–698. DOI: 10.1080/03605302.2018.1459686.

[45] Tentarelli, L., Tilli, P. (2019). An existence result for dissipative nonhomogeneous hyperbolic equations via a minimization approach. *J. Differ. Equ.* 266(8):5185–5208. DOI: 10.1016/j.jde.2018.10.023.

[46] Barbu, V. (2012). Optimal control approach to nonlinear diffusion equations driven by Wiener noise. *J. Optim. Theory Appl.* . 153(1):1–26. DOI: 10.1007/s10957-011-9946-8.

[47] Barbu, V., Röckner, M. (2015). An operatorial approach to stochastic partial differential equations driven by linear multiplicative noise. *J. Eur. Math. Soc.* 17(7):1789–1815. DOI: 10.4171/JEMS/545.

[48] Krylov, N. V. (1985). Extremal properties of solutions of stochastic equations. *Theory Probab. Appl.* 29(2):205–217. DOI: 10.1137/1129033.

[49] Boroushaki, S., Ghoussoub, N. (2019). A self-dual variational approach to stochastic partial differential equations. *J. Funct. Anal.* 276(4):1201–1243. DOI: 10.1016/j.jfa.2018.05.010.

[50] Edwards, R. E. (1965). *Functional analysis. Theory and Applications*. New York: Holt, Rinehart and Winston.

[51] Hess, C. (1995). On the measurability of the conjugate and the subdifferential of a normal integrand. *J. Convex Anal.* 2(1/2):153–165.

[52] Rockafellar, R. T. (1976). Integral functionals, normal integrands and measurable selections. In: Gossez, J.P., Lami Dozo, E.J., Mawhin, J., Waelbroeck, L. (Eds), *Nonlinear Operators and the Calculus of Variations* (Summer School, Univ. Libre Bruxelles, Brussels, 1975), Lecture Notes in Math., vol. 543. Berlin: Springer; pp. 157–207.

[53] Brzeźniak, Z., van Neerven, J. M. A. M., Veraar, M. C., Weis, L. (2008). Itô's formula in UMD Banach spaces and regularity of solutions of the Zakai equation. *J. Differ. Equ.* 245(1):30–58. DOI: 10.1016/j.jde.2008.03.026.

[54] van Neerven, J. M. A. M., Veraar, M. C., Weis, L. (2007). Stochastic integration in UMD Banach spaces. *Ann. Probab.* 35(4):1438–1478. DOI: 10.1214/009117906000001006.

[55] van Neerven, J. M. A. M., Veraar, M. C., Weis, L. (2008). Stochastic evolution equations in UMD Banach spaces. *J. Funct. Anal.* 255(4):940–993. DOI: 10.1016/j.jfa.2008.03.015.

[56] Liu, W., Stephan, M. (2014). Yosida approximations for multivalued stochastic partial differential equations driven by Lévy noise on a Gelfand triple. *J. Math. Anal. Appl.* 410(1):158–178. DOI: 10.1016/j.jmaa.2013.08.016.

[57] Barbu, V. (2010). Nonlinear differential equations of monotone types in Banach spaces. In *Springer Monographs in Mathematics*. New York: Springer

[58]   Kenmochi, N. (1975). Some nonlinear parabolic variational inequalities. *Israel J. Math.* 22(3/4):304–331. DOI: 10.1007/BF02761596.

[59]   Fuhrman, M., Tessitore, G. (2002). Nonlinear Kolmogorov equations in infinite dimensional spaces: The backward stochastic differential equations approach and applications to optimal control. *Ann. Probab.* 30(3):1397–1465.

[60]   Hu, Y., Peng, S. G. (1991). Adapted solution of a backward semilinear stochastic evolution equation. *Stochastic Anal. Appl.* 9(4):445–459. DOI: 10.1080/07362999108809250.

[61]   Da Prato, G., Zabczyk, J. (2014). *Stochastic Equations in Infinite Dimensions, Volume 152 of Encyclopedia of Mathematics and Its Applications*, 2nd ed. Cambridge: Cambridge University Press,

[62]   Brézis, H. (1973). *Opérateurs Maximaux Monotones et Semi-Groupes de Contractions Dans Les Espaces de Hilbert.* North-Holland Mathematics Studies, No. 5. Notas de Matemática (50). New York: North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc.

[63]   Flandoli, F., Gatarek, D. (1995). Martingale and stationary solutions for stochastic Navier-Stokes equations. *Probab. Th. Rel. Fields* . 102(3):367–391. DOI: 10.1007/BF01192467.

[64]   Liu, W., Röckner, M. (2015). *Stochastic Partial Differential Equations: An Introduction.* Cham: Springer

[65]   GyöNgy, Iön., Krylov, N. (1996). Existence of strong solutions for Itô's stochastic equations via approximations. *Probab. Th. Rel. Fields.* 105(2):143–158. DOI: 10.1007/BF01203833.

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

[67]   Ikeda, N., Watanabe, S. (1989). *Stochastic Differential Equations and Diffusion Processes, Volume 24 of North-Holland Mathematical Library*, 2nd ed. Amsterdam: North-Holland Publishing Co.

[68]   Vallet, G., Zimmermann, A. (2019). Well-posedness for a pseudomonotone evolution problem with multiplicative noise. *J. Evol. Equ.* 19(1):153–202. DOI: 10.1007/s00028-018-0472-0.

[69]   Scarpa, L., Stefanelli, U. (2020). Doubly nonlinear stochastic evolution equations. *Math. Models Methods Appl. Sci.* 30(5):991–1031. DOI: 10.1142/S0218202520500219.