Well-posedness of monotone semilinear SPDEs with semimartingale noise

Carlo Marinelli* Luca Scarpa†

October 28, 2019

Abstract

We prove existence and uniqueness of strong solutions for a class of semilinear stochastic evolution equations driven by general Hilbert space-valued semimartingales, with drift equal to the sum of a linear maximal monotone operator in variational form and of the superposition operator associated to a random time-dependent monotone function defined on the whole real line. Such a function is only assumed to satisfy a very mild symmetry-like condition, but its rate of growth towards infinity can be arbitrary. Moreover, the noise is of multiplicative type and can be path-dependent. The solution is obtained via a priori estimates on solutions to regularized equations, interpreted both as stochastic equations as well as deterministic equations with random coefficients, and ensuing compactness properties. A key role is played by an infinite-dimensional Doob-type inequality due to Métivier and Pellaumail.

AMS Subject Classification: 60H15, 47H06, 46N30.

Key words and phrases: stochastic evolution equations, singular drift, semimartingale noise, variational approach, monotonicity methods

1 Introduction

Let us consider semilinear stochastic evolution equations of the type

$$dX(t) + AX(t)\,dt + \beta(t, X(t))\,dt \ni B(t, X)\,dZ(t), \qquad X(0) = X_0, \tag{1.1}$$

in $L^2(D)$, where D is a smooth bounded domain of $\mathbb{R}^n$. Here A is linear coercive maximal monotone operator on $L^2(D)$, β is a random time-dependent maximal monotone graph everywhere defined on the real line, Z is a Hilbert space-valued semimartingale, and the coefficient B satisfies a suitable Lipschitz continuity assumption (precise hypotheses on the data are given in §2 below). Our main result is the existence and uniqueness of a strong solution to (1.1) (in the sense of Definition 3.1 below), and its continuous dependence on the initial datum in a suitable topology. Stochastic partial differential equations driven by semimartingales arise naturally in several fields, such as physics, biology, and finance, where a noise with possibly discontinuous trajectories can be preferable, for modeling purposes, to the classical Wiener noise (see, e.g., [4, 17]). For further possible applications where equations of the form (1.1) are used we refer to [5, 16] and references therein.

Maximal monotone graphs such as β arise naturally in the study of equations with nonlinearities associated to monotone discontinuous functions. In fact, it is well known that every

*Department of Mathematics, University College London, Gower Street, London WC1E 6BT, United Kingdom. URL: http://goo.gl/4GKJP

†Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria. E-mail: luca.scarpa@univie.ac.at

maximal monotone graph γ in $\mathbb{R} \times \mathbb{R}$ arises (in a unique way) from an increasing function $\gamma_0 : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$, setting $\gamma(r) := [\gamma_0(r-), \gamma_0(r+)]$ for every $r \in \mathbb{R}$, i.e. by the procedure of "filling the jumps". Therefore our treatment provides a notion of (strong) solution to stochastic evolution equations of the type,

$$dX(t) + AX(t)\,dt + \beta_0(t, X(t))\,dt = B(t, X)\,dZ(t), \qquad X(0) = X_0,$$

where $\beta_0(\omega, t, \cdot) \colon \mathbb{R} \to \mathbb{R}$ is an increasing function, with possibly countably many discontinuities, and with essentially no assumption on its rate of growth at infinity. Stochastic evolution equations of this type are particularly interesting as they cannot be handled using existing techniques, as well as for their potential applications (equations with exponentially growing drift appear, for instance, in mathematical models of Euclidean quantum field theory – see, e.g., [1]). In fact, to the best of our knowledge, all results currently available in the literature on stochastic equations with semimartingale noise are obtained under assumptions on the coefficients that are too restrictive to treat equation (1.1). In particular, after the pioneering results by Métivier [23] for equations with bounded A and locally Lipschitz continuous drift and diffusion coefficients, the first contribution to treat "genuine" stochastic evolution equations (i.e., with A unbounded) is probably [8], where the well-posedness result in the variational setting for equations with Wiener noise of [13, 25] is extended to the case where the driving noise is a quasi left-continuous locally square-integrable martingale, although under a rather restrictive growth assumption on the (nonlinear) drift term. In particular, semilinear equations such as (1.1) can be treated with this approach only if β is Lipschitz continuous. More recently, nonlinear equations in the variational setting driven by compensated Poisson random measures have been considered, also under relaxed monotonicity conditions, in [5]. Semilinear equations with drift $A + \beta$, as in (1.1), can be treated within this framework under polynomial growth assumptions on β that depend on the dimension of the domain $D \subset \mathbb{R}^n$: the larger n is, the slower (polynomial) growth is allowed for β (cf. [15] for a discussion of this issue). Our results do not suffer of this drawback, as the growth rate of β is not limited in any way by the dimension n. Multivalued stochastic equations with possibly càdlàg additive noise have been studied also in [6], under a linear growth condition on the drift, so that semilinear equations such as (1.1) can be treated only if β has at most linear growth. Using semigroup methods, well-posedness for (1.1) in the mild sense is proved in [18, 19], under the assumptions that β grows polynomially and the noise is the sum of a Wiener process and a compensated Poisson random measure (one should note, however, that A needs not admit a variational formulation). The well-posedness result for (1.1) obtained here should be interesting also in the finite-dimensional setting, i.e. for stochastic (ordinary) differential equations driven by finite-dimensional semimartingales. In fact, apart of the classical well-posedness results for equations with locally Lipschitz coefficients (see, e.g., [24, 26]), it seems that the only work dealing with equations with monotone coefficients is [12], where, however, linear growth is required.

The strong solution to (1.1) is constructed as limit of solutions to approximating equations. In particular, replacing both A and β with their Yosida approximations, one obtains a family of approximating equations with bounded coefficients that admit classical solutions in $L^2(D)$, thanks to results by Métivier and Pellaumail (see [23, 24]). This double regularization is necessary because, due to the general semimartingale noise, one cannot simply regularize β and rely on the classical variational theory in [8, 13, 25]. Since we allow β to be random, care is needed to make sure that its Yosida approximation is at least a progressively measurable function (see §2 below for detail on this technical issue). Furthermore, we first consider such regularized equations with additive noise, i.e. with B possibly random, but not dependent on the unknown, and with the semimartingale Z satisfying extra integrability conditions that are removed in a second step. Interpreting such approximating equations either as "true" stochastic equations or as deterministic evolution equations with random coefficients (cf. [20, 22]), we obtain a priori estimates for their solutions in various topologies. This idea has already been used in [22] to deal with the well-posedness of semilinear equations with singular drift and Wiener noise, and later in [20] to study regularity properties of their solutions. The much more general assumptions on the noise in the present situation give rise to several difficulties that require new ideas with respect to [20, 22]. A fundamental tool is an infinite-dimensional maximal inequality for stochastic integrals with respect to semimartingales due to Métivier and Pellaumail (see [23, 24]). These a priori estimates imply enough compactness to pass to the limit in the regularized equations, thus solving a version of (1.1) with additive noise. The assumption that β is everywhere defined plays here a crucial role, as it allows to use weak compactness techniques in L^1 spaces. In order to treat the general case with multiplicative noise, we proceed as follows: using localization techniques, we first show the existence of strong solutions on closed stochastic intervals. This technique also allows to remove the extra integrability assumption on Z. Uniqueness of solutions on closed stochastic intervals implies that such local solutions form a directed system, so that it is natural to construct a maximal solution. Finally, the linear growth of B is shown to imply that the maximal solution can be extended to any compact time interval. One can also show that the solution depends continuously on the initial datum in the sense of the topology of uniform (in time) convergence in probability.

Several auxiliary results are needed to carry out the program outlined above, some of which are interesting in their own right. For instance, we prove a general version of Itô's formula for the square of the $L^2(D)$-norm in a variational setting with possibly singular terms. This can be seen as an extension of the classical formulas by Pardoux, Krylov, and Rozovskiĭ [13, 25], as well as by Krylov and Gyöngy [9], at least in the case where the variational triple is Hilbertian. We shall investigate in more detail Itô-type formulas in (generalized) variational settings in a work in preparation. We also give a characterization of weakly càdlàg processes in terms of essential boundedness (in time) and a weak càdlàg property in a larger space, extending the classical result on weak continuity for vector-valued functions by Strauss (see [29]).

The remaining text is organized as follows: in §2 we fix the notation, collect all standing assumptions, and discuss some notable consequences thereof that are going to be used extensively. The definition of strong solution, both in the global and the local sense, and the statement of the main well-posedness result are given in §3. In §4 we recall some elements of the above-mentioned approach by Métivier and Pellaumail to stochastic integration with respect to semimartingales in Hilbert space, centered around a fundamental stopped Doob-type inequality. We also prove an extension to the càdlàg case of a classical criterion for weak continuity of vector-valued function due to Strauss, as well as a slight generalization of a classical criterion for uniform integrability by de la Vallé-Poussin. In §5 we prove an Itô-type formula for the square of the $L^2(D)$ norm of a process that can be decomposed into the sum of a stochastic integral with respect to a (Hilbert-space-valued) semimartingale and of a Lebesgue integral of a singular drift term. This result is an essential tool to obtain, in §6, an auxiliary well-posedness result for a version of (1.1) with additive noise. Finally, the proof of the main result is presented in §7.

Acknowledgment. Large part of the work for this paper was done during several stays of the first-named author at the Interdisziplinäres Zentrum für Komplexe Systeme (IZKS), Universität Bonn, Germany, as guest of Prof. S. Albeverio. His kind hospitality and the excellent working conditions at IZKS are gratefully acknowledged. The second-named author was funded by Vienna Science and Technology Fund (WWTF) through Project MA14-009.

2 Assumptions and first consequences

2.1 Notation

Every Banach space is intended as a real Banach space. For any Banach spaces E and F, we shall denote the Banach space of continuous linear operators from E to F by $\mathscr{L}(E,F)$, if endowed with the operator norm, and by $\mathscr{L}_s(E,F)$, if endowed with the strong operator topology (i.e. with the topology of simple convergence). If $E = F$, we shall just write $\mathscr{L}(E)$ in place of $\mathscr{L}(E,E)$. The usual Lebesgue-Bochner spaces of E-valued functions on a measure space $(Y,\mathscr{A},m)$ will be denoted by $L^p(Y;E)$, $p \in [0,\infty]$, where $L^0(Y;E)$ is endowed with the (metrizable) topology of convergence in measure. The set of continuous functions and of weakly continuous functions on $[0,T]$ with values in E will be denoted by $C([0,T];E)$ and $C_w([0,T];E)$, respectively. Analogously, the symbols $D([0,T];E)$ and $D_w([0,T];E)$ stand for the corresponding spaces of càdlàg functions. A function $f: Y \to \mathscr{L}(E,F)$ will be called strongly measurable if it is the limit in the norm topology of $\mathscr{L}(E,F)$ of a sequence of elementary functions. For every $f \in D([0,T];E)$ we shall use the symbol f^* for $\sup_{t\in[0,T]}\|f(t)\|_E$.

We shall denote by D a smooth bounded domain of $\mathbb{R}^n$, and by H the Hilbert space $L^2(D)$ with its usual scalar product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$.

All random elements will be defined on a fixed probability space $(\Omega,\mathscr{F},\mathbb{P})$ endowed with a filtration $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$ satisfying the "usual assumptions" of right-continuity and completeness. Identities and inequalities between random variables will always be meant to hold $\mathbb{P}$-almost surely, unless otherwise stated. Two (measurable) processes will be declared equal if they are indistinguishable. By Z we shall denote a fixed semimartingale taking values in a (fixed) separable Hilbert space K. The standard notation and terminology of stochastic calculus for semimartingales will be used (see, e.g., [23]).

For any $a,b \in \mathbb{R}$ we shall write $a \lesssim b$ to indicate that there exists a constant $c>0$ such that $a \leqslant cb$.

2.2 Assumptions

The following hypotheses will be in force throughout the paper.

Assumption (A). We assume that $A \in \mathscr{L}(V,V')$, where V is a separable Hilbert space densely, continuously and compactly embedded in H, and that there exists a constant $c>0$ such that

$$\langle Au,u\rangle \geqslant c\|u\|_V^2 \qquad \forall u \in V.$$

We denote by A_2 the part of A in H, i.e. the unbounded linear operator $(A_2,\mathsf{D}(A_2))$ on H defined as $A_2v := Av$ for $v \in \mathsf{D}(A_2) := \{v \in V : Av \in H\}$. Furthermore, we assume that there exists a sequence $(T_n)_{n\in\mathbb{N}}$ of linear injective operators on $L^1(D)$ such that, for every $n \in \mathbb{N}$,

(a) $T_n : L^1(D) \to L^1(D)$ is sub-Markovian, i.e., if $f \in L^1(D)$ with $0 \leqslant f \leqslant 1$ a.e. in D, then $0 \leqslant T_n f \leqslant 1$ a.e. in D;

(b) T_n is ultracontractive, i.e $T_n \in \mathscr{L}(L^1(D),L^\infty(D))$.

Moreover, denoting the restriction of T_n to H by the same symbol, we assume that

(c) $T_n \in \mathscr{L}(H,V)$ for every $n \in \mathbb{N}$ and it can be extended to a continuous linear operator on V', still denoted by the same symbol;

(d) T_n converges to the identity in $\mathscr{L}_s(E)$, with $E \in \{L^1(D),V,H,V'\}$, as $n \to \infty$;

(e) $T_n(H) = T_m(H)$ for every $n,m \in \mathbb{N}$.

Throughout the work, we shall denote by V_0 a Hilbert space continuously embedded in $V \cap L^\infty(D)$ and dense in V. Thanks to the assumptions on (T_n) such a space always exists, for instance setting $V_0 := T_{\bar n}(H)$, with $\bar n$ an arbitrary (but fixed) natural number. Indeed, V_0 is independent of $\bar n$ thanks to (e), so that for every $v \in V$ the sequence $(T_n v)_n \subset V_0$ converges to v in V thanks to (d). An arbitrary but fixed terminal time will be denoted by T.

Assumption (J). Let $j : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}_+$ be a function satisfying the following conditions:

(a) $j(\cdot,\cdot,x)$ is progressively measurable for all $x \in \mathbb{R}$;

(b) $j(\omega,t,\cdot)$ is convex for every $(\omega,t) \in \Omega \times [0,T]$, with $j(\cdot,\cdot,0) = 0$;

(c) one has

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

uniformly with respect to $(\omega, t) \in \Omega \times [0,T]$.

For every $(\omega, t) \in \Omega \times [0,T]$, the maximal monotone graph $\beta(\omega,t,\cdot) \subset \mathbb{R}^2$ is defined as the subdifferential of $j(\omega,t,\cdot)$, i.e. $y \in \beta(\omega,t,x)$ if and only if

$$j(\omega,t,x) + y(z-x) \leqslant j(\omega,t,z) \qquad \forall z \in \mathbb{R}.$$

Seeing the maximal monotone graph $\beta(\omega,t,\cdot)$ as a multivalued map, condition (b) implies that

(d) $\beta(\omega,t,\cdot)$ is everywhere defined for every $(\omega,t) \in \Omega \times [0,T]$.

We further assume that

(e) $\beta(\omega,t,\cdot)$ is bounded on bounded sets uniformly with respect to (ω,t).

The forthcoming assumptions on the coefficient B are formulated in terms of control processes for semimartingales, whose definition is given in §4.1 below.

Assumption (B). Let $B : \Omega \times [0,T] \times D([0,T];H) \to \mathscr{L}(K,H)$ be a map satisfying the following conditions:

(a) the process $B(\cdot,\cdot,u)$ is a strongly predictable $\mathscr{L}(K,H)$-valued process for every adapted càdlàg H-valued process u;

(b) for every stopping time $\tau \leqslant T$, and for every adapted càdlàg H-valued processes u, v,

$$u1_{[\![0,\tau[\![} = v1_{[\![0,\tau[\![} \qquad \text{implies} \qquad B(\cdot,u)1_{[\![0,\tau]\!]} = B(\cdot,v)1_{[\![0,\tau]\!]};$$

(c) for every control process C of Z there exists an increasing, nonnegative, right-continuous, adapted process L such that, for every $t \in \,]0,T]$ and every adapted càdlàg H-valued processes u, v, one has

$$\int_0^t \big\|B(s,u) - B(s,v)\big\|^2_{\mathscr{L}(K,H)}\,dC(s) \leqslant \int_0^t \sup_{r<s}\big\|u(r)-v(r)\big\|^2\,dL(s),$$
$$\int_0^t \big\|B(s,u)\big\|^2_{\mathscr{L}(K,H)}\,dC(s) \leqslant \int_0^t \Big(1 + \sup_{r<s}\big\|u(r)\big\|^2\Big)\,dL(s).$$

Assumptions (a) and (b) are immediately satisfied if B is of the form $B(\omega,t,u) = \widetilde{B}(\omega,t,u(t-))$ for all $u \in D([0,T];H)$ and $(\omega,t) \in \Omega \times [0,T]$, with the convention $u(0-) := u(0)$, where $\widetilde{B} : \Omega \times [0,T] \times H \to \mathscr{L}(K,H)$ is strongly measurable with respect to the product σ-algebra of the predictable σ-algebra and of the Borel σ-algebra of H. A more refined criterion can be found in [24, §§6.2–6.4].

Finally, the initial datum X_0 is an H-valued $\mathscr{F}_0$-measurable random variable.

2.3 On assumptions (A) and (J)

Assumptions (A) and (J) have important consequences that will be extensively used in the sequel. The most important ones are collected in this subsection.

The hypotheses on V and A ensure that (V,H,V') is a Hilbertian variational triple and that the operator A is maximal monotone from V to V'. Moreover, as it follows by coercivity, linearity, and monotonicity, A is bijective form V to V'. However, in applications it is often necessary to consider only the weaker coercivity on A

$$\langle Au, u\rangle \geqslant c\|u\|_V^2 - \delta\|u\|^2 \qquad \forall u \in V,$$

with $\delta > 0$ a constant. This case can be included in our analysis by considering the operator $A + \delta I$ instead of A.

The hypotheses on A are met by large classes of differential operators (second order symmetric and non-symmetric divergence-form operators, as well as the fractional Laplacian, for example) – see, e.g., [22] for a detailed list of concrete examples.

The standard example of a family of operators (T_n) that can be shown to satisfy conditions (a)–(d) above for large classes of operators A is $T_n := (I + (1/n)A)^{-m}$, with $m \in \mathbb{N}$ sufficiently large. We refer again to, e.g., [22] for a discussion of this issue. Moreover, note that for T_n to belong to $\mathscr{L}(V')$ it suffices that the commutator $R_n := T_nA_2 - AT_n : \mathsf{D}(A_2) \to V'$ can be continuously extended to a linear bounded operator from $V \to V'$. In fact, this allows to extend T_n to a linear bounded operator on V' as follows: for any $y \in V'$, by surjectivity of A one has $y = Au$, with $u \in V$. Setting $T_ny := R_nu + AT_nu \in V'$, in order to check that this is well defined it is sufficient to prove that if $u \in V$ is such that $Au = 0$, then $R_nu + AT_nu = 0$. Let $u \in V$ be such that $Au = 0$. Then $Au \in H$, hence $u \in \mathsf{D}(A_2)$, and $0 = Au = A_2u$. Since T_n has already been defined on H, we have $0 = T_nA_2u = R_nu + AT_nu$. Finally, we have

$$\begin{aligned}\|T_ny\|_{V'} &\leqslant \|R_nu\|_{V'} + \|AT_nu\|_{V'} \leqslant \|R_n\|_{\mathscr{L}(V,V')}\|u\|_V + \|A\|_{\mathscr{L}(V,V')}\|T_n\|_{\mathscr{L}(V)}\|u\|_V \\ &\leqslant \big(\|R_n\|_{\mathscr{L}(V,V')}\|A^{-1}\|_{\mathscr{L}(V',V)} + \|A\|_{\mathscr{L}(V,V')}\|T_n\|_{\mathscr{L}(V)}\big)\,\|v\|_{V'},\end{aligned}$$

so that $T_n : V' \to V'$ is also bounded.

The Banach-Steinhaus theorem implies that the sequence of linear operators (T_n) is bounded in $\mathscr{L}(H)$, $\mathscr{L}(V)$, and $\mathscr{L}(V')$, i.e.

$$\sup_{n\in\mathbb{N}}\big\|T_n\big\|_{\mathscr{L}(H)} + \sup_{n\in\mathbb{N}}\big\|T_n\big\|_{\mathscr{L}(V)} + \sup_{n\in\mathbb{N}}\big\|T_n\big\|_{\mathscr{L}(V')} < \infty.$$

The continuity property of the adjoint family (T_n^*) established next plays an important role in the proof of the Itô-type formula for the square of the H-norm in §5.

Lemma 2.1. *The sequence of adjoint operators $(T_n^*)_{n\in\mathbb{N}} \subset \mathscr{L}(H)$ is contained in $\mathscr{L}(V)$ and converges to the identity in $\mathscr{L}_s(H)$.*

Proof. By the continuity of (T_n) in $\mathscr{L}_s(H)$ one has, for every $x, y \in H$,

$$\langle T_n^*x, y\rangle = \langle x, T_ny\rangle \to \langle x, y\rangle,$$

hence T_n^*x converges weakly to x in H for every $x \in H$. Furthermore, for any $x \in V$ and $y \in H$, one has

$$\langle T_n^*x, y\rangle = \langle x, T_ny\rangle \leqslant \big\|x\big\|_V\big\|T_ny\big\|_{V'} \leqslant N\big\|x\big\|_V\big\|y\big\|_{V'},$$

where $N := \sup_{n\in\mathbb{N}}\|T_n\|_{\mathscr{L}(V')}$. Since H is densely and continuously embedded in V', this readily implies that $T_n^*x \in V'' \simeq V$ and

$$\big\|T_n^*x\big\|_V \leqslant N\big\|x\big\|_V \qquad \forall x \in V, \quad \forall n \in \mathbb{N}.$$

Since V is reflexive, for any sequence $(n') \subset \mathbb{N}$, there exist $z \in V$ and a subsequence $(n'') \subset (n')$, possibly depending on x, and such that $T_{n''}^*x$ converges weakly in V to z as $n'' \to \infty$. Since V is compactly embedded in H, $T_{n''}^*x$ converges strongly to z in H. Recalling that T_n^*x converges weakly to x as $n \to \infty$, hence that so does $T_{n''}^*x$, we infer that $z = x$, i.e., $T_{n''}^*x$ converges strongly to x in H. By a standard result of classical analysis, this yields the convergence of T_n^*x to x in H, that is, along the original sequence, which is independent of $x \in V$. The result can finally be extended to $x \in H$ by a density argument: let $(x_k) \subset V$ be a sequence converging to x in H. The triangle inequality yields

$$\begin{aligned}\|T_n^* x - x\| &\leqslant \|T_n^* x - T_n^* x_k\| + \|T_n^* x_k - x_k\| + \|x_k - x\| \\ &\leqslant \big(1 + \sup_n \|T_n\|_{\mathscr{L}(H)}\big)\|x_k - x\| + \|T_n^* x_k - x_k\|,\end{aligned}$$

from which one easily concludes. □

Remark 2.2. In general, the adjunction map $T \mapsto T^*$ for linear bounded operators on a Hilbert space is continuous with respect to the uniform and the weak operator topology, but not with respect to the strong operator topology. The previous lemma thus identifies a (very!) special subset of linear bounded operators for which the adjunction map is continuous also with respect to the strong operator topology.

Let us now discuss some consequences of assumption (J). For every $(\omega, t) \in \Omega \times [0, T]$, let $j^*(\omega, t, \cdot)$ denote the convex conjugate of $j(\omega, t, \cdot)$, defined as

$$j^*(\omega, t, y) = \sup_{x \in \mathbb{R}} \big(xy - j(\omega, t, x)\big).$$

The measurability and continuity hypotheses on j imply that j and j^* are normal integrands, or, equivalently, that their epigraphs are progressively Effros-measurable (see, e.g., [11, 27]). More precisely, let us recall that, given a function $\phi : \Omega \times [0, T] \times \mathbb{R} \to \mathbb{R}$, its epigraph at $(\omega, t) \in \Omega \times [0, T]$ is given by

$$\operatorname{epi} \phi(\omega, t) := \big\{(x, y) \in \mathbb{R}^2 : \phi(\omega, t, x) \leqslant y\big\}.$$

The progressive Effros-measurability of the epigraph of ϕ is then defined as the progressive measurability of the set

$$\big\{(\omega, t) \in \Omega \times [0, T] : \operatorname{epi} \phi(\omega, t) \cap E \neq \varnothing\big\}$$

for every open $E \subset \mathbb{R}^2$.

Moreover, if j is a normal integrand, then β is also progressively Effros-measurable (see *op. cit*), which in turn implies that the resolvent $(I + \lambda\beta)^{-1}$ and the Yosida approximation β_λ of β, both real-valued functions on $\Omega \times [0, T] \times \mathbb{R}$, are measurable with respect to the product of the progressive σ-algebra and the Borel σ-algebra (see, e,g., [16, Proposition 3.12]).

Assumption (c) can be interpreted by saying that, for any fixed (ω, t), the rates of growth of j at plus and minus infinity are comparable. For instance, this is satisfied if $j(\omega, t, \cdot)$ is even for every (ω, t).

Assumption (d) implies that $j^*(\omega, t, \cdot)$ is superlinear at infinity, uniformly with respect to (ω, t), i.e. that

$$\lim_{|y| \to +\infty} \frac{j^*(\omega, t, y)}{|y|} = +\infty \quad \text{uniformly in } (\omega, t) \in \Omega \times [0, T].$$

Lastly, taking $z = 0$ in the definition of β as subdifferential of j, assumption (e) implies that, for all $(\omega, t) \in \Omega \times [0, T]$, $j(\omega, t, x) \leqslant yx$ for all $y \in \beta(\omega, t, x)$, that is, $j(\omega, t, \cdot)$ is bounded on bounded sets uniformly over $\Omega \times [0, T]$.

The above measurability conditions are obviously satisfied if β is non-random and time-independent, i.e. if β is an everywhere defined maximal monotone graph in $\mathbb{R} \times \mathbb{R}$. Moreover, in this case the convex function $j : \mathbb{R} \to \mathbb{R}_+$ such that $\partial j = \beta$ and $j(0) = 0$ is uniquely determined, and $\mathsf{D}(\beta) = \mathbb{R}$ implies that j^* is superlinear at infinity.

The boundedness assumption (e) is the natural generalization of the analogous ones commonly used for time-dependent maximal monotone graphs (see, e.g., [3, p. 4]).

Note that all the conditions assumed to hold for every $(\omega, t) \in \Omega \times [0, T]$ could have been assumed for almost every $(\omega, t) \in \Omega \times [0, T]$ instead. Indeed, in such a case, if $E \subset \Omega \times [0, T]$ has measure 0 and all hypotheses hold outside E, then one can consider the restriction of j to the complement of E instead of j.

3 Main result

The concept of solution we are going to work with is as follows. We recall that $T \in \mathbb{R}_+$ is an arbitrary but fixed time horizon.

Definition 3.1. *Let $\tau \leqslant T$ be a stopping time. A strong solution on $[\![0,\tau]\!]$ to* (1.1) *is a pair (X,ξ), where X is an adapted càdlàg H-valued process and ξ is an adapted $L^1(D)$-valued process, such that*

(a) $\mathbb{1}_{[\![0,\tau]\!]}X \in L^1(0,T;V)$ and $\mathbb{1}_{[\![0,\tau]\!]}\xi \in L^1([0,T]\times D)$ $\mathbb{P}$-a.s., with $\xi \in \beta(\cdot,X)$ a.e. in $[\![0,\tau[\![\times D$;

(b) $\mathbb{1}_{[\![0,\tau]\!]}B(\cdot,X)$ is integrable with respect to Z;

(c) one has, as an identity in $V' \cap L^1(D)$,

$$X^\tau + \int_0^{\cdot\wedge\tau} AX(s)\,ds + \int_0^{\cdot\wedge\tau} \xi(s)\,ds = X_0 + \big(\mathbb{1}_{[\![0,\tau]\!]}B(\cdot,X)\big)\cdot Z.$$

A strong solution on $[0,T]$ will simply be called a strong solution.

The main results of the paper are collected in the following theorem. These ensure that (1.1) admits a strong solution, which is unique within a natural class of processes, and depends continuously on the initial datum.

Theorem 3.2. *Equation* (1.1) *admits a strong solution (X,ξ), with X optional, and it is the only one such that*

$$\sup_{t\leqslant T}\|X(t)\|^2 + \int_0^T \|X(s)\|_V^2\,ds + \int_0^T\int_D \xi(s)X(s)\,dx\,ds < \infty \qquad \mathbb{P}\text{-a.s.}$$

Moreover, the solution map $X_0 \mapsto X$ is continuous from $L^0(\Omega;H)$ to $L^0(\Omega;D([0,T];H)\cap L^2(0,T;V))$, where $D([0,T];H)$ is endowed with the topology generated by the supremum norm.

Note that since $\xi \in \beta(\cdot,X)$ we have $|\xi X| = \xi X = j(\cdot,X) + j^*(\cdot,\xi) \geqslant 0$, so that Theorem 3.2 ensures that

$$\xi X = j(\cdot,X) + j^*(\cdot,\xi) \in L^1((0,T)\times D) \quad \mathbb{P}\text{-a.s.}$$

4 Preliminaries and auxiliary results

We recall those results from the approach to stochastic integration developed by Métivier and Pellaumail that we need, referring to [23, 24] for details. We also prove two additional lemmata pertaining to this theory that are indispensable for the proofs in the following sections.

Moreover, we provide a sufficient condition for a process to be weakly càdlàg and a generalized version of the uniform integrability criterion by de la Vallée Poussin.

4.1 Stochastic integration with respect to Hilbert-space-valued semimartingales

Let G be a separable Hilbert space. An $\mathscr{L}(K,G)$-valued process Y is elementary if there exist $n \in \mathbb{N}$, sequences (s_k), $(t_k) \subseteq \mathbb{R}_+$, $(F_k) \subset \mathscr{F}$, and $(u_k) \subset \mathscr{L}(K,G)$, $k = 1,\dots,n$, with $s_k \leqslant t_k$ and $F_k \in \mathscr{F}_{s_k}$, such that

$$Y = \sum_{k=1}^n \mathbb{1}_{]s_k,t_k]\times F_k} u_k.$$

Then the stochastic integral of Y with respect to Z is defined as

$$\big(Y\cdot Z\big)_t := \sum_{k=1}^n \mathbb{1}_{F_k} u_k\big(Z_{t_k\wedge t} - Z_{s_k\wedge t}\big) \qquad \forall t \in \mathbb{R}_+.$$

Definition 4.1. *A positive increasing adapted process C is called a control process for Z if, for every separable Hilbert space G, for every elementary $\mathscr{L}(K,G)$-valued process Y, and for every stopping time τ, one has*

$$\mathbb{E}\sup_{t<\tau}\big\|(Y\cdot Z)_t\big\|_G^2 \leqslant \mathbb{E}\, C_{\tau-}\int_{]0,\tau[}\big\|Y(s)\big\|^2_{\mathscr{L}(K,G)}\,dC(s).$$

It turns out that an adapted càdlàg K-valued process is a semimartingale if and only if it admits a control process. In particular, the set of control processes for a semimartingale Z, that we shall denote by $\mathscr{C}(Z)$, is not empty. One can also show (see [23, Theorems. 23.9–23.14]) that, writing $Z = M + V$, with M locally square integrable local martingale and V a finite-variation process, a control process is given by

$$C = 8\big(\langle M,M\rangle + [\check{M},\check{M}]\big) + 2\big(2\vee|V|\big),$$

where $\langle M,M\rangle$ is the predictable quadratic variation of M, $|V|$ is the variation of V, and $[\check{M},\check{M}]$ is the quadratic variation of the pure-jump martingale part of M, in the sense of [23, Definition 19.3].

We need to introduce some notation: for any control process C and any strongly measurable adapted process Y with values in $\mathscr{L}(K,G)$, let us define the process $\lambda^C(Y)$ as

$$\lambda^C_t(Y) := C_t\int_0^t\big\|Y(s)\big\|^2_{\mathscr{L}(K,G)}\,dC(s) \qquad \forall t\in\mathbb{R}_+.$$

For any stopping time τ, let us define the measure m^Z_τ on the predictable σ-algebra as

$$m^Z_\tau : P \longmapsto \mathbb{E}\, C_{\tau-}\big(\mathbb{1}_P\cdot C\big)_{\tau-},$$

and note that m^Z_τ is finite if $\mathbb{E}|C_{\tau-}|^2<\infty$. The space of strongly predictable processes Y with values in $\mathscr{L}(K,G)$ such that $\mathbb{E}\,\lambda^C_{\tau-}(Y)$ is finite coincides with the Bochner L^2 space with respect to the measure m^Z_τ and values in $\mathscr{L}(K,G)$, with norm

$$\big\|Y\big\|_{L^2(m^Z_\tau)} = \big(\mathbb{E}\,\lambda^C_{\tau-}(Y)\big)^{1/2} = \Big(\mathbb{E}\, C_{\tau-}\big(\|Y\|^2\cdot C\big)_{\tau-}\Big)^{1/2},$$

where the norm of Y is taken in $\mathscr{L}(K,G)$ (see [23, § 24.1, Lemmata 1–3, and § 26.1]). Denoting the Banach space of adapted càdlàg processes S with values in G such that $\mathbb{E}\,S^{*2}<\infty$ by $\mathbb{S}^2$, with norm $\|S\|_{S^2} := (\mathbb{E}\,S^{*2})^{1/2}$, the inequality in the definition of control process can thus be written as

$$\big\|(Y\cdot Z)^{\tau-}\big\|_{\mathbb{S}^2} \leqslant \big\|Y\big\|_{L^2(m^Z_\tau)}.$$

The first step in the construction of the stochastic integral for more general integrands is as follows: suppose that there exists a stopping time τ such that $\mathbb{E}|C_{\tau-}|^2<\infty$, so that m^Z_τ is a finite measure and the vector space of elementary processes is dense in $L^2(m^Z_\tau)$. Then the mapping $Y\mapsto(Y\cdot Z)^{\tau-}$, initially defined on elementary processes, admits a unique extension to a linear continuous map from $L^2(m^Z_\tau)$ to $\mathbb{S}^2$. As a second step, assume that C is a control process for Z and Y is a process with values in $\mathscr{L}(K,G)$ such that the process $\lambda^C(Y)$ is finite, and introduce the sequence of stopping times (τ_n) defined as

$$\tau_n := \inf\big\{t\geqslant 0 :\, C_t\wedge\lambda^C_t(Y)\geqslant n\big\},$$

so that $\mathbb{E}|C_{\tau_n-}|^2<\infty$ as well as $\mathbb{E}\,\lambda^C_{\tau_n-}(Y)<\infty$, i.e. $Y\in L^2(m^Z_{\tau_n})$. Then, by the previous step, one has $(Y\cdot Z)^{\tau_n-}\in\mathbb{S}^2$ for all $n\in\mathbb{N}$. Since τ_n increases to ∞ as $n\to\infty$ and it is not difficult to show that $(Y\cdot Z)^{\tau_n-}=(Y\cdot Z)^{\tau_m-}$ on $[\![0,\tau_n\wedge\tau_m[\![$ for all n, $m\in\mathbb{N}$, one has a well-defined process $Y\cdot Z$. One then shows that such a process does not depend on the sequence (τ_n). However, it may still depend on the control process C. A final step shows that if Y admits two control processes C_1 and C_2 such that the processes $\lambda^{C_1}(Y)$ and $\lambda^{C_2}(Y)$ are finite, then the stochastic integrals constructed in the two possible ways coincide. The following definition is therefore meaningful.

Definition 4.2. *A strongly predictable $\mathscr{L}(K,G)$-valued process Y is integrable with respect to Z if there exists a control process C for Z such that the process $\lambda^C(Y)$ is finite.*

We shall occasionally use the symbol $\mathscr{S}_C(Z)$ to denote the set of strongly predictable $\mathscr{L}(K,G)$-valued processes Y such that the process $\lambda^C(Y)$ is finite.

Note that the construction of $Y \cdot Z$ implies that the inequality in the definition of control processes can be extended as follows: for every $C \in \mathscr{C}(Z)$, $Y \in \mathscr{S}_C(Z)$, and stopping time τ, one has

$$\big\|(Y \cdot Z)^{\tau-}\big\|_{\mathbb{S}^2} \leqslant \Big(\mathbb{E}\, C_{\tau-}\big(\|Y\|^2 \cdot C\big)_{\tau-}\Big)^{1/2}.$$

We shall need a further maximal inequality for stochastic integrals with respect to a semimartingale, whose proof relies on the following deep inequality (see [14, Lemma 1.3]).

Lemma 4.3. *Let X be a positive real-valued measurable process and A an increasing predictable process such that, for every finite stopping time σ,*

$$\mathbb{E}\,\mathbb{1}_{\{\sigma>0\}} X(\sigma) \leqslant a\,\mathbb{E}\,\mathbb{1}_{\{\sigma>0\}} A(\sigma)$$

for a constant $a > 0$. Then for every concave function $F : \mathbb{R}_+ \to \mathbb{R}$ and every finite stopping time τ one has

$$\mathbb{E}\,\mathbb{1}_{\{\tau>0\}} F(X(\tau)) \leqslant (a+1)\,\mathbb{E}\,\mathbb{1}_{\{\tau>0\}} F(A(\tau)).$$

Let C be a control process for Z and $Y \in \mathscr{S}_C(Z)$, so that

$$\mathbb{E}\big(Y \cdot Z\big)^{*2}_{\sigma-} \leqslant \mathbb{E}\, C_{\sigma-}\big(\|Y\|^2 \cdot C\big)_{\sigma-}.$$

Since the process $C_-\big(\|Y\|^2 \cdot C\big)_-$ is left-continuous, hence predictable, the previous lemma yields, taking $F(r) = \sqrt{r}$, $r \geqslant 0$,

$$\mathbb{E}\big(Y \cdot Z\big)^{*}_{\tau-} \leqslant 2\,\mathbb{E}\big(C_{\tau-}\big(\|Y\|^2 \cdot C\big)_{\tau-}\big)^{1/2}.$$

The following elementary lemma is essential in the last section.

Lemma 4.4. *Let C be a control process for the semimartingale Z and τ a stopping time. Then $C^{\tau-}$ is a control process for the semimartingale $Z^{\tau-}$.*

Proof. For every elementary $\mathscr{L}(K,G)$-valued process Y and every stopping time σ one has $Y \cdot Z^{\tau-} = (Y \cdot Z)^{\tau-}$, hence also

$$\big(Y \cdot Z^{\tau-}\big)^*_{\sigma-} = \big(Y \cdot Z\big)^*_{(\sigma\wedge\tau)-},$$

which in turns implies

$$\mathbb{E}\big(Y \cdot Z^{\tau-}\big)^{*2}_{\sigma-} = \mathbb{E}\big(Y \cdot Z\big)^{*2}_{(\sigma\wedge\tau)-} \leqslant \mathbb{E}\, C_{(\sigma\wedge\tau)-}\big(\|Y\|^2 \cdot C\big)_{(\sigma\wedge\tau)-},$$

where $C_{(\sigma\wedge\tau)-} = C^{\tau-}_{\sigma-}$ and $\big(\|Y\|^2 \cdot C\big)_{(\sigma\wedge\tau)-} = \big(\|Y\|^2 \cdot C^{\tau-}\big)_{\sigma-}$. □

We also recall the following version of the dominated convergence theorem for stochastic integrals with respect to semimartingales (cf. [23, Theorem 26.3]).

Proposition 4.5. *Let $(X_n)_{n\in\mathbb{N}}$, X be predictable $\mathscr{L}(K,H)$-valued processes such that $X_n \to X$ in $\mathscr{L}(K,H)$ a.e. in $\Omega \times [0,T]$. If there exists a control process C for Z and $\phi \in \mathscr{S}_C(Z)$ such that*

$$\big\|X_n\big\|_{\mathscr{L}(K,H)} \leqslant \big\|\phi\big\|_{\mathscr{L}(K,H)} \qquad \forall n \in \mathbb{N},$$

then $X_n \in \mathscr{S}_C(Z)$ for every $n \in \mathbb{N}$, $X \in \mathscr{S}_C(Z)$, and

$$\big(X_n \cdot Z - X \cdot Z\big)^*_t \longrightarrow 0$$

in probability for every $t \in [0,T]$.

Finally, we recall, for the reader's convenience, the following stochastic version of Gronwall's lemma (cf. [23, Lemma 29.1]).

Lemma 4.6. *Let A be an adapted, right-continuous, increasing, positive process defined on a stochastic interval $[\![0,\tau[\![$, with $\ell := \sup_{t<\tau} A(t) \in \mathbb{R}_+$. Let also ϕ be a real, increasing, adapted process such that, for every stopping time $\sigma \leqslant \tau$,*

$$\mathbb{E}\,\phi(\sigma-) \leqslant a + b\,\mathbb{E}\int_0^{\sigma-} \phi(s-)\,dA(s)$$

for certain constants $a, b \in \mathbb{R}$. Then,

$$\mathbb{E}\,\phi(\tau-) \leqslant a \sum_{k=0}^{[2b\ell]} (2b\ell)^k.$$

4.2 Weak right-continuity of vector-valued functions

Throughout this section E and F denote two Banach spaces, with E reflexive, densely and continuously embedded in F. A classical result by Strauss (see [29]) states that

$$L^\infty(0,T;E) \cap C_w([0,T];F) = C_w([0,T];E).$$

We are going to show that the result continues to hold replacing the spaces of weakly continuous functions by spaces of weakly càdlàg functions.

Lemma 4.7. *One has*

$$L^\infty(0,T;E) \cap D_w([0,T];F) = D_w([0,T];E).$$

Proof. The inclusion of the space on the right-hand side in the space on the left-hand side is evident. Let $u \in L^\infty(0,T;E) \cap D_w([0,T];F)$. Since $\{T\}$ is negligible with respect to the Lebesgue measure on $[0,T]$, it is not restrictive to suppose that $u(T) \in E$ (otherwise, we shall modify the value of u in T, obtaining a version of u which is still in $L^\infty(0,T;E) \cap D_w([0,T];F)$). We first show that, in order for u to belong to $D_w([0,T];E)$, it suffices to prove that there exists a constant M such that $\|u(t)\|_E \leqslant M$ for every $t \in [0,T]$.
Step 1. Assuming that $u([0,T])$ is bounded in E, let $t \in [0,T)$ and $(t_n) \subset [t,T)$ be a sequence converging to t. Then $u(t_n) \to u(t)$ weakly in F by assumption, and, since E is reflexive, there exists a subsequence $(t_{n'})$ and $v \in E$ such that $u(t_{n'}) \to v$ weakly in E. Therefore $v = u(t)$ and $u(t_n) \to u(t)$ weakly in E, i.e. u is weakly càd with values in E. A completely analogous (in fact easier) argument shows that u is also weakly làg with values in E.
Step 2. Let (ρ_n) be a sequence of mollifiers in $\mathbb{R}$ whose support is contained in $[-\frac{2}{n}, 0]$. Denoting the extension of u to zero outside $[0,T]$ by the same symbol, it follows from $u \in L^\infty(\mathbb{R};E)$ that $u_n := \rho_n * u \in C(\mathbb{R};E)$. In particular, Minkowski's inequality yields

$$\big\|u_n(t)\big\|_E \leqslant \int_{\mathbb{R}} |\rho_n(s)| \big\|u(t-s)\big\|_E\,ds \leqslant \big\|u\big\|_{L^\infty(0,T;E)} =: M$$

for all $t \in \mathbb{R}$ and $n \in \mathbb{N}$. Let $t_0 \in [0,T)$ be arbitrary but fixed. By reflexivity of E, there exist $v \in E$ and a subsequence of $(u_n(t_0))$, denoted by the same symbol for simplicity, such that $u_n(t_0) \to v$ weakly in E. Moreover, for any $\varphi \in F'$,

$$\langle \varphi, u_n \rangle_F = \langle \varphi, \rho_n * u \rangle_F = \rho_n * \langle \varphi, u \rangle_F,$$

where $f := \langle \varphi, u \rangle_F \in D([0,T])$ by assumption. In particular, f is right-continuous at t_0, i.e. for any $\delta > 0$ there exists $N \in \mathbb{N}$ such that $|f(t_0 - s) - f(t_0)| < \delta$ for all $s \in [-2/N, 0]$. Since the support of ρ_n is contained in $[-2/n, 0]$, for $n > N$ we have

$$\begin{aligned}\big|\langle\varphi, u_n(t_0)\rangle_F - \langle\varphi, u(t_0)\rangle_F\big| &\leqslant \int_{\mathbb{R}} \rho_n(s)\big|f(t_0-s) - f(t_0)\big|\,ds\\ &\leqslant \delta \int_{\mathbb{R}} \rho_n(s) = \delta,\end{aligned}$$

i.e. $\langle\varphi, u_n(t)\rangle_F \to \langle\varphi, u(t)\rangle_F$ as $n \to \infty$. Since this holds for any $\varphi \in F'$, we infer that $u_n(t_0) \to u(t_0)$ weakly in F. Moreover, as $(u_n(t_0))$ is bounded in E and E is reflexive, we easily deduce that $u_n(t_0) \to u(t_0)$ weakly in E, thus also, by weak lower semicontinuity of the norm, that

$$\big\|u(t_0)\big\|_E \leqslant \liminf_{n\to\infty}\big\|u_n(t_0)\big\|_E \leqslant M.$$

Since $t_0 \in [0, T)$ was arbitrary, this implies that $\|u(t)\|_E \leqslant M$ for all $t \in [0, T)$. Moreover, since $u(T) \in E$, we have that $u([0, T])$ is bounded in E, as required. □

4.3 A criterion for uniform integrability

We shall need a slightly generalized version of the de la Vallée-Poussin criterion for uniform integrability. For the purposes of this paragraph only, $(E, \mathscr{E}, \mu)$ will denote a finite measure space, and m will stand for the product measure of $\mathbb{P}$, the Lebesgue measure, and μ on $\Omega \times [0, T] \times E$. For compactness of notation, we set

$$L^p(m) := L^p(\Omega \times [0,T] \times E, \mathscr{F} \otimes \mathscr{B}([0,T]) \otimes \mathscr{E}, m)$$

for any $p \in [0, \infty]$.

Lemma 4.8. *Let $F : \Omega \times [0,T] \times \mathbb{R} \to [0, +\infty]$ be proper, convex and lower semicontinuous in the third variable, measurable in the first two, and such that*

$$\lim_{|x|\to+\infty} \frac{F(\omega,t,x)}{|x|} = +\infty \qquad \textit{uniformly in } (\omega,t) \in \Omega \times [0,T].$$

If $\mathcal{G} \subseteq L^0(m)$ is such that there exists a constant C for which

$$\|F(\cdot,\cdot,g)\|_{L^1(m)} < C \qquad \forall g \in \mathcal{G},$$

then $\mathcal{G}$ is uniformly integrable in $\Omega \times [0,T] \times E$.

Proof. We need to show that $\mathcal{G}$ is bounded in $L^1(m)$ and that for every $\varepsilon > 0$ there exist δ such that, for any measurable set A with $m(A) < \delta$, one has

$$\int_A |g|\,dm < \varepsilon \qquad \forall g \in \mathcal{G}.$$

Let $M > 0$ be a constant. By assumption there exists R such that $x \in \mathbb{R}$ with $|x| > R$ implies $|F(\omega,t,x)| > M|x|$ for every $(\omega,t) \in \Omega \times [0,T]$. Then one has, for any $g \in \mathcal{G}$,

$$\begin{aligned}\int_A |g|\,dm &= \int_{A\cap\{|g|\leqslant R\}} |g|\,dm + \int_{A\cap\{|g|>R\}} |g|\,dm\\ &\leqslant R\,m(A) + \frac{1}{M}\int F(\cdot,\cdot,g)\,dm\\ &\leqslant R\,m(A) + \frac{C}{M}.\end{aligned}$$

Choosing $A = \Omega \times [0,T] \times E$ it immediately follows that $\mathcal{G}$ is bounded in $L^1(m)$. Moreover, for every $\varepsilon > 0$ there exists M such that $\frac{C}{M} < \frac{\varepsilon}{2}$, hence $\delta := \frac{\varepsilon}{2R}$ satisfies the condition we are looking for. □

The same argument shows, keeping $\omega \in \Omega$ fixed, that if there exists a finite positive random variable $C : \Omega \to \mathbb{R}_+$ such that, for $\mathbb{P}$-a.e. $\omega \in \Omega$,

$$\|F(\omega, \cdot, g)\|_{L^1([0,T]\times E)} < C(\omega) \qquad \forall g \in \mathcal{G},$$

then $\mathcal{G}(\omega, \cdot)$ is uniformly integrable in $(0,T) \times E$ for $\mathbb{P}$-a.e. $\omega \in \Omega$.

5 The Itô formula

In this section we prove an Itô-type formula for the square of the H-norm: this can be seen as an integration-by-parts formula in a generalized setting. We point out that the framework that we consider here is is "unusual", as we work with processes with components in V' and $L^1(D)$ simultaneously, for which Itô's formula is not available using existing techniques. Let us recall also that the quadratic variation of Z is defined as the process

$$[Z, Z] := \|Z\|^2 - \|Z_0\|^2 - 2Z_- \cdot Z.$$

In the sequel we shall denote $[0,T] \times D$ by D_T.

Proposition 5.1. *Let C be a control process for Z, $G \in \mathscr{S}_C(Z)$, $Y_0 \in L^0(\Omega, \mathscr{F}_0, \mathbb{P}; H)$, and the adapted processes*

$$\begin{aligned}
&Y \in L^0(\Omega; L^\infty(0,T;H)) \cap L^0(\Omega; L^2(0,T;V)) \\
&v \in L^0(\Omega; L^2(0,T;V')), \\
&g \in L^0(\Omega; L^1(0,T;L^1(D)))
\end{aligned}$$

be such that

$$Y + \int_0^\cdot v(s)\,ds + \int_0^\cdot g(s)\,ds = Y_0 + G \cdot Z. \tag{5.1}$$

Furthermore, assume that there exists a real number $a > 0$ such that

$$j(\cdot, aY) + j^*(\cdot, ag) \in L^0(\Omega; L^1(D_T)).$$

Then

$$\begin{aligned}
\frac{1}{2}\|Y\|^2 + \int_0^\cdot \langle v(s), Y(s)\rangle\,ds + \int_0^\cdot\!\!\int_D g(s)Y(s)\,dx\,ds \\
= \frac{1}{2}\|Y_0\|^2 + \frac{1}{2}\big[G\cdot Z, G \cdot Z\big] + (Y_- G)\cdot Z.
\end{aligned}$$

Proof. Let us first show that the stochastic integral $(Y_-G)\cdot Z$ is well defined: it follows from (5.1) that Y is strongly càdlàg in V_0'. Since $Y \in L^\infty(0,T;H)$, Lemma 4.7 implies that Y is weakly càdlàg in H, i.e. that, for any $h \in H$, $\langle Y, h\rangle$ is càdlàg, hence that $\langle Y_-, h\rangle$ is left-continuous, in particular predictable, or, equivalently, that Y_- is weakly predictable. However, since H is separable, Pettis' theorem implies that Y_- is predictable. Moreover, one has

$$\begin{aligned}
\lambda_T^C(Y_-G) &= C(T)\big(\big\|Y_-G\big\|^2_{\mathscr{L}(K,\mathbb{R})}\cdot C\big)_T \\
&\leqslant C(T)\sup_{t<T}\big\|Y(t)\big\|^2\big(\big\|G\big\|^2_{\mathscr{L}(K,H)}\cdot C\big)_T = \sup_{t<T}\big\|Y(t)\big\|^2\lambda_T^C(G) < +\infty.
\end{aligned}$$

Denoting the action of the operator T_n by a superscript n, we have

$$Y^n + \int_0^\cdot v^n(s)\,ds + \int_0^\cdot g^n(s)\,ds = Y_0^n + G^n \cdot Z,$$

as the Bochner integral as well as the stochastic integral commute with linear continuous operators. Since all integrands on the left-hand side are H-valued processes, the integration-by-parts formula for H-valued semimartingales yields (cf. [23, §25])

$$\begin{aligned}\frac{1}{2}\|Y^n\|^2 + \int_0^\cdot \langle v^n(s), Y^n(s)\rangle\,ds + \int_0^\cdot\!\!\int_D g^n(s)Y^n(s)\,dx\,ds \\ = \frac{1}{2}\|Y_0^n\|^2 + \frac{1}{2}\big[G^n\cdot Z, G^n\cdot Z\big] + (Y_-^n G^n)\cdot Z.\end{aligned}$$

We are now going to pass to the limit as $n\to\infty$ in this identity. The continuity of (T_n) in $\mathscr{L}_s(H)$ immediately yields

$$\begin{aligned}\|Y_0^n\|^2 &\longrightarrow \|Y_0\|^2,\\ \|Y^n(t)\|^2 &\longrightarrow \|Y(t)\|^2 \qquad \forall t\in[0,T],\\ G^n &\longrightarrow G \qquad \text{in } \mathscr{L}_s(K,H) \text{ a.e. in } \Omega\times[0,T].\end{aligned}$$

Similarly, since (T_n) is also continuous in the strong operator topology of V, V', and $L^1(D)$, the dominated convergence theorem readily implies that

$$\begin{aligned}Y^n &\longrightarrow Y \quad \text{in } L^2(0,T;V),\\ v^n &\longrightarrow v \quad \text{in } L^2(0,T;V'),\\ g^n &\longrightarrow g \quad \text{in } L^1(D_T).\end{aligned}$$

In particular, passing to a subsequence if necessary, this implies that $g^nY^n\to gY$ almost everywhere in D_T. Therefore, if we show that (g^nY^n) is uniformly integrable on D_T, we can conclude by Vitali's theorem that the latter convergence continues to hold also in $L^1(D_T)$. Thanks to the assumptions on the behavior at infinity of j, the sub-Markovianity of T_n, and the generalized Jensen inequality for positive operators (cf. [10]), we have

$$\begin{aligned}\pm a^2 g^nY^n \leqslant j(\cdot,\pm aY^n) + j^*(\cdot,ag^n) &\lesssim 1 + j(\cdot,aY^n) + j^*(\cdot,g^n)\\ &\leqslant 1 + T_n\big(j(\cdot,aY) + j^*(\cdot,g)\big),\end{aligned}$$

where

$$T_n\big(j(\cdot,aY) + j^*(\cdot,g)\big) \longrightarrow j(\cdot,aY) + j^*(\cdot,g) \qquad \text{in } L^1(D_T)$$

as $n\to\infty$, because the right-hand side belongs to $L^1(D_T)$ a.s. by assumption. In particular, $T_n\big(j(\cdot,aY) + j^*(\cdot,g)\big)$ is uniformly integrable on D_T, and so is (g^nY^n) by comparison. This implies, as explained above, that

$$\int_0^\cdot\!\!\int_D g^n(s)Y^n(s)\,dx\,ds \longrightarrow \int_0^\cdot\!\!\int_D g(s)Y(s)\,dx\,ds.$$

Let us now consider the quadratic variation term. By definition we have

$$\big[G^n\cdot Z, G^n\cdot Z\big] = \big\|G^n\cdot Z\big\|^2 - 2((G^n\cdot Z)_-)\cdot(G^n\cdot Z),$$

where the stochastic integral on the right-hand side can be written as $\tilde{G}^n\cdot Z$, with

$$\tilde{G}^n:\Omega\times[0,T]\to\mathscr{L}(K,\mathbb{R}), \qquad \tilde{G}^n(\omega,t)k := \langle (G^n\cdot Z)(\omega,s-), G^n(\omega,s)k\rangle, \quad k\in K.$$

Noting that $G^n\cdot Z = T_n(G\cdot Z)$, it is immediate that $(G^n\cdot Z)_t\to(G\cdot Z)_t$ for all $t\in[0,T]$ $\mathbb{P}$-a.s. as $n\to\infty$. Moreover, setting

$$\tilde{G}:\Omega\times[0,T]\to\mathscr{L}(K,\mathbb{R}), \qquad \tilde{G}(\omega,s)k := \langle (G\cdot Z)(\omega,s-), G(\omega,s)k\rangle, \quad k\in K,$$

one has

$$\begin{aligned}\tilde{G}^n k - \tilde{G}k &= \bigl\langle (G^n\cdot Z)_-, G^n k\bigr\rangle - \bigl\langle (G\cdot Z)_-, Gk\bigr\rangle\\ &= \bigl\langle T_n^*T_n(G\cdot Z)_- - (G\cdot Z)_-, Gk\bigr\rangle\\ &\leqslant \bigl\|T_n^*T_n(G\cdot Z)_- - (G\cdot Z)_-\bigr\|\bigl\|G\bigr\|_{\mathscr{L}(K,H)}\|k\|_K,\end{aligned}$$

where

$$\begin{aligned}\bigl\|T_n^*T_n(G\cdot Z)_- - (G\cdot Z)_-\bigr\| &\leqslant \bigl\|T_n^*T_n(G\cdot Z)_- - T_n^*(G\cdot Z)_-\bigr\|\\ &\quad + \|T_n^*(G\cdot Z)_- - (G\cdot Z)_-\|\\ &\leqslant \sup_{n\in\mathbb{N}}\bigl\|T_n^*\bigr\|_{\mathscr{L}(H)}\bigl\|T_n(G\cdot Z)_- - (G\cdot Z)_-\bigr\|\\ &\quad + \bigl\|T_n^*(G\cdot Z)_- - (G\cdot Z)_-\bigr\|,\end{aligned}$$

and the right-hand side converges to zero pointwise in time $\mathbb{P}$-a.s. because both T_n and its adjoint converge to the identity operator in $\mathscr{L}_s(H)$. Therefore $\tilde{G}^n$ converges to $\tilde{G}$ in $\mathscr{L}(K,\mathbb{R})$ a.e. in $\Omega\times[0,T]$, and it follows by Proposition 4.5 that

$$\bigl[G^n\cdot Z, G^n\cdot Z\bigr]_t \longrightarrow \bigl[G\cdot Z, G\cdot Z\bigr]_t \qquad \forall t\in[0,T] \quad \mathbb{P}\text{-a.s.}$$

Lastly, let us consider the convergence of the term $(Y_-^nG^n)\cdot Z$. Note that the $\mathscr{L}(K,\mathbb{R})$-valued processes $Y_-^nG^n$ and Y_-G are defined as

$$(Y_-^nG^n): k\longmapsto \bigl\langle T_nY_-, T_nGk\bigr\rangle, \qquad (Y_-G): k\longmapsto \bigl\langle Y_-, Gk\bigr\rangle,$$

so that

$$\begin{aligned}\bigl|(Y_-^nG^n - Y_-G)k\bigr| &= \bigl\langle T_nY_-, T_nGk\bigr\rangle - \bigl\langle Y_-, Gk\bigr\rangle\\ &= \bigl\langle T_n^*T_nY_- - Y_-, Gk\bigr\rangle\\ &\leqslant \bigl\|T_n^*T_nY_- - Y_-\bigr\|\bigl\|G\bigr\|_{\mathscr{L}(K,H)}\|k\|_K,\end{aligned}$$

which in turn yields

$$\bigl\|(Y_-^nG^n - Y_-G)\bigr\|_{\mathscr{L}(K,\mathbb{R})} \leqslant \bigl\|G\bigr\|_{\mathscr{L}(K,H)}\bigl\|T_n^*T_nY_- - Y_-\bigr\| \qquad \text{a.e. in } \Omega\times(0,T].$$

Recalling that $(T_n)_n$ is uniformly bounded in $\mathscr{L}(H)$, hence so is $(T_n^*)_n$, it follows that

$$\begin{aligned}\bigl\|T_n^*T_nY_- - Y_-\bigr\| &\leqslant \bigl\|T_n^*T_nY_- - T_n^*Y_-\bigr\| + \bigl\|T_n^*Y_- - Y_-\bigr\|\\ &\leqslant \sup_{n\in\mathbb{N}}\bigl\|T_n^*\bigr\|_{\mathscr{L}(H)}\bigl\|T_nY_- - Y_-\bigr\| + \bigl\|T_n^*Y_- - Y_-\bigr\|,\end{aligned}$$

where the right-hand side converges to zero a.e. in $\Omega\times(0,T]$ thanks to the assumptions on (T_n) and to Lemma 2.1. Therefore $Y_-^nG^n$ converges to Y_-G in $\mathscr{L}(K,\mathbb{R})$ a.e. in $\Omega\times(0,T]$, so that Proposition 4.5 allows us to conclude that $(Y_-^nG^n)\cdot Z$ converges to $(Y_-G)\cdot Z$ in probability uniformly in time. □

6 Well-posedness with additive noise

The goal of this section is to establish a well-posedness result for the following version of (1.1) with additive noise:

$$dX(t) + AX(t)\,dt + \beta(t,X(t))\,dt \ni G(t)\,dZ(t), \qquad X(0)=X_0, \tag{6.1}$$

where G is a strongly predictable $\mathscr{L}(K,H)$-valued process integrable with respect to Z. This is an essential step towards the proof of the main results in the next section.

We begin with an existence result.

Theorem 6.1. *Let C be a control process for Z and $G \in \mathscr{S}_C(Z)$ such that $\mathbb{E}\,\lambda^C_{T-}(G) < \infty$ and assume that $X_0 \in L^2(\Omega; H)$. Then (6.1) admits a strong solution.*

The main idea of the proof is to regularize both A and β in (6.1), so that the regularized equation admits a (unique) strong solution in the classical sense, to obtain uniform estimates on such solutions, and finally to pass to the limit using compactness and monotonicity arguments.

For any $\lambda \in]0,1[$, let $\beta_\lambda : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}$ and $A_\lambda \in \mathscr{L}(H)$ be the Yosida approximations of $r \mapsto \beta(\cdot,\cdot,r)$ and of A_2, respectively (see [2] for references). Recall that A_2 denotes the part of A in H and that, setting $J_\lambda := (I + \lambda A_2)^{-1}$, by definition of A_λ we have that $A_\lambda = AJ_\lambda$.

Let us consider the regularized equation

$$dX_\lambda(t) + A_\lambda X_\lambda(t)\,dt + \beta_\lambda(t, X_\lambda(t))\,dt = G(t)\,dZ(t), \qquad X_\lambda(0) = X_0. \tag{6.2}$$

Since $A_\lambda + \beta_\lambda$ is Lipschitz continuous (uniformly over $\Omega \times [0,T]$), the equation admits a unique strong solution X_λ in the classical sense, i.e. X_λ is an adapted càdlàg H-valued process, with

$$\mathbb{E}\sup_{t \leqslant T} \|X_\lambda(t)\|^2 < +\infty,$$

such that

$$X_\lambda + \int_0^\cdot A_\lambda X_\lambda(s)\,ds + \int_0^\cdot \beta_\lambda(s, X_\lambda(s))\,ds = X_0 + G \cdot Z$$

(see [23, Thm. 34.7–35.2]).

We are now going to establish a priori estimates on (X_λ) and functionals thereof.

Lemma 6.2. *There exists a constant $N > 0$ such that, for every $\lambda \in]0,1[$,*

$$\mathbb{E}\sup_{t<T} \|X_\lambda(t)\|^2 + \mathbb{E}\|J_\lambda X_\lambda\|^2_{L^2(0,T;V)} + \mathbb{E}\|\beta_\lambda(\cdot, X_\lambda)X_\lambda\|_{L^1(D_T)} < N.$$

Proof. The integration-by-parts formula for H-valued processes yields

$$\begin{aligned}\frac{1}{2}\|X_\lambda\|^2 &+ \int_0^\cdot \bigl\langle A_\lambda X_\lambda(s), X_\lambda(s)\bigr\rangle\,ds + \int_0^\cdot\!\!\int_D \beta(s, X_\lambda(s))X_\lambda(s)\,ds\\ &= \frac{1}{2}\|X_0\|^2 + \frac{1}{2}\bigl[G\cdot Z, G\cdot Z\bigr] + (X_{\lambda-}G)\cdot Z,\end{aligned}$$

where $X_{\lambda-}$ denotes the process $(X_\lambda)_-$. Taking the supremum in time over $[0,T[$, recalling the identity

$$\bigl\langle A_\lambda X_\lambda, X_\lambda\bigr\rangle = \bigl\langle AJ_\lambda X_\lambda, J_\lambda X_\lambda\bigr\rangle + \lambda\bigl\|A_\lambda X_\lambda\bigr\|^2,$$

one has, by coercivity of A,

$$\begin{aligned}\bigl(X_\lambda\bigr)^{*2}_{T-} &+ 2c\int_0^T \bigl\|J_\lambda X_\lambda(s)\bigr\|^2_V\,ds + 2\int_0^T\!\!\int_D \beta_\lambda(s, X_\lambda(s))X_\lambda(s)\,dx\,ds\\ &\leqslant \|X_0\|^2 + \bigl[G\cdot Z, G\cdot Z\bigr]_{T-} + 2\bigl((X_{\lambda-}G)\cdot Z\bigr)^*_{T-}.\end{aligned}$$

We are going to estimate the last two terms on the right-hand side of the last inequality. By definition of quadratic variation we have

$$\begin{aligned}[G\cdot Z, G\cdot Z] &= \|G\cdot Z\|^2 - 2(G\cdot Z)_-\cdot(G\cdot Z)\\ &= \|G\cdot Z\|^2 - 2\tilde{G}\cdot Z,\end{aligned}$$

where $\tilde{G} : \Omega \times [0,T] \to \mathscr{L}(K,\mathbb{R}) \simeq K$ is defined as $K \ni k \mapsto \bigl\langle (G\cdot Z)_-, Gk\bigr\rangle$. By definition of control process and by the second inequality for stochastic integrals in §4.1 we thus have

$$\begin{aligned}\mathbb{E}[G\cdot Z, G\cdot Z]_{T-} &\leqslant \mathbb{E}\bigl(G\cdot Z\bigr)^{*2}_{T-} + 2\,\mathbb{E}\bigl(\tilde{G}\cdot Z\bigr)^*_{T-}\\ &\leqslant \mathbb{E}\,\lambda^C_{T-}(G) + 4\,\mathbb{E}\,\lambda^C_{T-}(\tilde{G})^{1/2},\end{aligned}$$

where, by elementary inequalities,

$$\begin{aligned}\mathbb{E}\,\lambda^C_{T-}(\tilde{G})^{1/2} &= \mathbb{E}\,C(T-)^{1/2}\big(\|\tilde{G}\|^2_K\cdot C\big)^{1/2}_{T-}\\ &\leqslant \mathbb{E}\,C(T-)^{1/2}\bigg(\int_0^{T-}\|(G\cdot Z)_-\|^2\|G\|^2_{\mathscr{L}(K,H)}\,dC\bigg)^{1/2}\\ &\leqslant \mathbb{E}\,\lambda^C_{T-}(G)^{1/2}\,\big(G\cdot Z\big)^*_{T-}\\ &\leqslant \frac{1}{2}\,\mathbb{E}\,\lambda^C_{T-}(G)+\frac{1}{2}\,\mathbb{E}\big(G\cdot Z\big)^{*2}_{T-}\leqslant \mathbb{E}\,\lambda^C_{T-}(G),\end{aligned}$$

so that $\mathbb{E}[G\cdot Z, G\cdot Z]_{T-}\leqslant 5\,\mathbb{E}\,\lambda^C_{T-}(G)$. Similarly, one has

$$\begin{aligned}\mathbb{E}\big((X_{\lambda-}G)\cdot Z\big)^*_{T-} &\leqslant 2\,\mathbb{E}\,C(T-)^{1/2}\bigg(\int_0^{T-}\|X_\lambda\|^2\|G\|^2_{\mathscr{L}(K,H)}\,dC\bigg)^{1/2}\\ &\leqslant 2\,\mathbb{E}\,\lambda^C_{T-}(G)^{1/2}\,\big(X_\lambda\big)^*_{T-}\\ &\leqslant \frac{1}{4}\,\mathbb{E}\big(X_\lambda\big)^{*2}_{T-}+4\,\mathbb{E}\,\lambda^C_{T-}(G),\end{aligned}$$

therefore also

$$\begin{aligned}\mathbb{E}\big(X_\lambda\big)^{*2}_{T-}+\mathbb{E}\int_0^T\big\|J_\lambda X_\lambda(s)\big\|^2_V\,ds+\mathbb{E}\int_0^T\!\!\int_D\beta_\lambda(s,X_\lambda(s))X_\lambda(s)\,dx\,ds\\ \lesssim \mathbb{E}\|X_0\|^2+\mathbb{E}\,\lambda^C_{T-}(G),\end{aligned}$$

uniformly over $\lambda\in\,]0,1[$, as the implicit constant depends only on c, the coercivity constant of A. We conclude noting that $\beta_\lambda(\cdot,X_\lambda)X_\lambda\geqslant 0$ by monotonicity of β_λ. □

We are going to establish an existence and uniqueness result for (6.1) under the additional assumption that

$$G:\Omega\times[0,T]\to\mathscr{L}(K,V_0). \tag{6.3}$$

This is only a technical "temporary" assumption that will be dispensed of in the proof of Theorem 6.1.

Proposition 6.3. *Assume that the hypotheses of Theorem 6.1 hold and that G satisfies* (6.3)*. Then* (6.1) *admits a unique strong solution.*

For the proof we need further a priori estimates on the solution to the regularized equation (6.2).

Lemma 6.4. *Let G satisfy* (6.3)*. There exists $\Omega'\in\mathscr{F}$ with $\mathbb{P}(\Omega')=1$ such that, for every $\omega\in\Omega'$, the following properties hold:*

(a) $\big(X_\lambda(\omega)\big)$ is bounded in $L^\infty(0,T;H)$;

(b) $\big(J_\lambda X_\lambda(\omega)\big)$ is bounded in $L^2(0,T;V)$;

(c) $\big(\lambda^{1/2}A_\lambda X_\lambda(\omega)\big)$ is bounded in $L^2(0,T;H)$;

(d) $\big(\beta_\lambda(\cdot,X_\lambda(\omega))X_\lambda(\omega)\big)$ is bounded in $L^1([0,T]\times D)$.

Proof. Thanks to assumption (6.3), there exists $\Omega'\in\mathscr{F}$, with $\mathbb{P}(\Omega')=1$, such that

$$(G\cdot Z)(\omega)\in L^\infty(0,T;V_0)\qquad\forall\omega\in\Omega'.$$

Let $\omega \in \Omega'$ be arbitrary but fixed, so that indication of the explicit dependence on ω of the various processes involved will be suppressed for compactness of notation. By inspection of (6.2) it follows that $X_\lambda - G \cdot Z \in H^1(0,T;V')$, so that we can write

$$\frac{d}{dt}(X_\lambda - G \cdot Z) + A_\lambda X_\lambda + \beta_\lambda(\cdot, X_\lambda) = 0$$

as an identity in V' which holds for a.a. $t \in]0,T[$. The (deterministic) integration-by-parts formula then yields

$$\begin{aligned}\frac{1}{2}\big\|X_\lambda - G \cdot Z\big\|^2 &+ \int_0^\cdot \big\langle A_\lambda X_\lambda(s), X_\lambda(s) - G \cdot Z(s)\big\rangle\, ds \\ &+ \int_0^\cdot\!\!\int_D \beta_\lambda(s, X_\lambda(s))(X_\lambda(s) - G \cdot Z(s))\, dx\, ds = \frac{1}{2}\|X_0\|^2,\end{aligned}$$

where (i) by the triangle inequality and the elementary inequality $(a-b)^2/2 \geqslant \frac{1}{4}a^2 - \frac{1}{2}b^2$, a, $b \in \mathbb{R}$, one has

$$\frac{1}{2}\big\|X_\lambda - G \cdot Z\big\|^2 \geqslant \frac{1}{2}\big(\|X_\lambda\| - \|G \cdot Z\|\big)^2 \geqslant \frac{1}{4}\|X_\lambda\|^2 - \frac{1}{2}\|G \cdot Z\|^2;$$

(ii) one has, for any $h \in H$, $\langle A_\lambda h, h\rangle = \langle A J_\lambda h, J_\lambda h\rangle + \lambda\|A_\lambda h\|^2$, so that, by coercivity of A and Young's inequality in the form $ab \leqslant \varepsilon a^2 + b^2/\varepsilon$, $a, b \in \mathbb{R}$, $\varepsilon > 0$, it follows that

$$\begin{aligned}\big\langle A_\lambda X_\lambda, X_\lambda - G \cdot Z\big\rangle &= \big\langle A_\lambda X_\lambda, X_\lambda\big\rangle - \big\langle A J_\lambda X_\lambda, G \cdot Z\big\rangle \\ &\geqslant c\big\|J_\lambda X_\lambda\big\|_V^2 + \lambda\big\|A_\lambda X_\lambda\big\|^2 \\ &\quad - \varepsilon\big\|A\big\|^2_{\mathscr{L}(V,V')}\big\|J_\lambda X_\lambda\big\|_V^2 + \frac{1}{\varepsilon}\big\|G \cdot Z\big\|_V^2;\end{aligned}$$

(iii) one has, for any $x \in \mathbb{R}$, slightly simplifying notation,

$$\begin{aligned}\beta_\lambda(x)x &= \beta_\lambda(x)(I + \lambda\beta)^{-1}(x) + \beta_\lambda(x)\big(x - (I + \lambda\beta)^{-1}(x)\big) \\ &= \beta_\lambda(x)(I + \lambda\beta)^{-1}(x) + \lambda\big|\beta_\lambda(x)\big|^2,\end{aligned}$$

hence also, recalling that $\beta_\lambda \in \beta \circ (I + \lambda\beta)^{-1}$ and that, for any $a, b \in \mathbb{R}$, $ab = j(a) + j^*(b)$ if and only if $b \in \partial j(a) = \beta(a)$,

$$\beta_\lambda(X_\lambda)X_\lambda \geqslant j\big((I + \lambda\beta)^{-1}(X_\lambda)\big) + j^*\big(\beta_\lambda(X_\lambda)\big) \geqslant j^*\big(\beta_\lambda(X_\lambda)\big);$$

(iv) Young's inequality in the form

$$ab \leqslant j^*(\varepsilon a) + j(b/\varepsilon) \leqslant \varepsilon j^*(a) + j(b/\varepsilon), \qquad a, b \in \mathbb{R},\ 0 < \varepsilon < 1,$$

implies

$$-\beta_\lambda(\cdot, X_\lambda)(G \cdot Z) \geqslant -\varepsilon j^*\big(\cdot, \beta_\lambda(\cdot, X_\lambda)\big) - j\big(\cdot, (G \cdot Z)/\varepsilon\big).$$

Choosing $\varepsilon < 1$, it follows from (i)–(iv) that

$$\begin{aligned}&\frac{1}{4}\big\|X_\lambda\big\|^2 + c\int_0^\cdot \big\|J_\lambda X_\lambda(s)\big\|_V^2\, ds + \lambda\int_0^\cdot \big\|A_\lambda X_\lambda(s)\big\|^2\, ds \\ &\qquad + \int_0^\cdot\!\!\int_D j^*\big(s, \beta_\lambda(s, X_\lambda(s))\big)\, dx\, ds \\ &\leqslant \frac{1}{2}\|X_0\|^2 + \frac{1}{2}\big\|G \cdot Z\big\|^2 \\ &\qquad + \varepsilon\big\|A\big\|_{\mathscr{L}(V,V')}\int_0^\cdot \big\|J_\lambda X_\lambda(s)\big\|_V^2\, ds + \frac{1}{\varepsilon}\int_0^\cdot \big\|(G \cdot Z)_s\big\|_V^2\, ds \\ &\qquad + \varepsilon\int_0^\cdot\!\!\int_D j^*\big(s, \beta_\lambda(s, X_\lambda(s))\big)\, dx\, ds + \int_0^\cdot\!\!\int_D j\big(s, (G \cdot Z)_s/\varepsilon\big)\, dx\, ds.\end{aligned}$$

First rearranging terms and choosing ε sufficiently small, then taking the essential supremum in time, one gets

$$\begin{aligned}\|X_\lambda\|^2_{L^\infty(0,T;H)} + \|J_\lambda X_\lambda\|^2_{L^2(0,T;V)} + \lambda\|A_\lambda X_\lambda\|^2_{L^2(0,T;H)} + \|j^*(\cdot,\beta_\lambda(\cdot,X_\lambda))\|_{L^1(D_T)}\\ \lesssim \|X_0\|^2 + \|G\cdot Z\|^2_{L^2(0,T;V)} + \int_{D_T} j\big(s,(G\cdot Z)_s/\varepsilon\big)\,dx\,ds,\end{aligned}$$

where the right-hand side is finite because $G\cdot Z \in L^\infty(0,T;V_0)$. In fact, recalling that V_0 is continuously embedded in V, this immediately implies that $G\cdot Z\in L^2(0,T;V)$; moreover, there exists $D'_T \subset D_T$, with $D_T\setminus D'_T$ of measure zero, such that the restriction of $G\cdot Z$ to D'_T is bounded. The finiteness of the last term on the right-hand side then follows by the boundedness on bounded sets of $y\mapsto j(\omega,t,y)$ uniformly over $(\omega,t)\in\Omega\times[0,T]$. □

The pathwise boundedness properties just proved entail several compactness properties in suitable topologies.

Lemma 6.5. *Let G satisfy (6.3). There exists $\Omega'\in\mathscr{F}$ with $\mathbb{P}(\Omega')=1$ such that, for every $\omega\in\Omega'$, there exist a subsequence $\lambda'=\lambda'(\omega)$ of λ and*

$$X(\omega)\in L^\infty(0,T;H)\cap L^2(0,T;V), \qquad \xi(\omega)\in L^1([0,T]\times D)$$

such that

$$\begin{aligned} X_{\lambda'}(\omega,\cdot) &\longrightarrow X(\omega,\cdot) && \text{weakly* in } L^\infty(0,T;H),\\ X_{\lambda'}(\omega,\cdot) &\longrightarrow X(\omega,\cdot) && \text{in } L^2(0,T;H),\\ J_{\lambda'}X_{\lambda'}(\omega,\cdot) &\longrightarrow X(\omega,\cdot) && \text{weakly in } L^2(0,T;V),\\ \beta_{\lambda'}(\cdot,X_{\lambda'}(\omega,\cdot)) &\longrightarrow \xi(\omega,\cdot) && \text{weakly in } L^1([0,T]\times D).\end{aligned}$$

Proof. Let Ω' be as in Lemma 6.4 and $\omega\in\Omega'$ arbitrary but fixed (whose indication will still be omitted). Since (X_λ) is bounded in $L^\infty(0,T;H)$, hence also in $L^2(0,T;H)$, there exist $X\in L^\infty(0,T;H)$ and a subsequence λ', depending on ω, such that $X_{\lambda'}$ converges weakly* to X in $L^\infty(0,T;H)$ and weakly in $L^2(0,T;H)$. The boundedness of $(J_\lambda X_\lambda)$ in $L^2(0,T;V)$ implies that there exists $\bar X\in L^2(0,T;V)$ such that $J_{\lambda'}X_{\lambda'}$ converges weakly to $\bar X$ in $L^2(0,T;V)$. Boundedness of $(\sqrt{\lambda}A_\lambda X_\lambda)$ in $L^2(0,T;H)$ implies that $\lambda A_\lambda X_\lambda$ converges to zero in $L^2(0,T;H)$. Writing

$$J_\lambda X_\lambda = X_\lambda - \lambda A_\lambda X_\lambda,$$

one immediately infers that $J_{\lambda'}X_{\lambda'}$ converges weakly to X in $L^2(0,T;H)$. Since it also converges weakly to $\bar X$ in $L^2(0,T;V)$, it follows that $\bar X = X$.

The same argument used in part (iii) of the proof of Lemma 6.4 yields

$$j^*(t,\beta_\lambda(t,X_\lambda)) \leqslant \beta_\lambda(t,X_\lambda)X_\lambda,$$

where the right-hand side, as a family indexed by λ, is bounded in $L^1(D_T)$. The generalized de la Vallée-Poussin criterion of Lemma 4.8 then ensures that $(\beta_\lambda(\cdot,X_\lambda))$ is uniformly integrable in D_T and hence relatively weakly compact in $L^1(D_T)$ by the Dunford-Pettis theorem, i.e. there exists $\xi\in L^1(D_T)$ such that $\beta_{\lambda'}(\cdot,X_{\lambda'})$ converges weakly to ξ in $L^1(D_T)$.

As a last step, we are going to show that $X_{\lambda'}$ converges to X in the norm topology of $L^2(0,T;H)$, rather than just in its weak topology. Writing the regularized equation as in Lemma 6.4, we have

$$\frac{d}{dt}(X_\lambda - G\cdot Z) + A_\lambda X_\lambda + \beta_\lambda(\cdot,X_\lambda) = 0,$$

where $A_\lambda X_\lambda = AJ_\lambda X_\lambda$ is bounded in $L^2(0,T;V')$ and $\beta_\lambda(\cdot,X_\lambda)$ is bounded in $L^1(D_T)$. Therefore $\frac{d}{dt}(X_\lambda - G\cdot Z)$ is bounded in $L^1(0,T;V'_0)$, and Simon's compactness criterion (see [28, Corollary 4, p. 85]) implies that $(X_\lambda - G\cdot Z)$ is relatively compact in $L^2(0,T;H)$. Since $G\cdot Z\in L^2(0,T;H)$ is independent of λ, the same conclusion holds for (X_λ) and by uniqueness of the weak limit in $L^2(0,T;H)$ it immediately follows that X_λ converges to X in $L^2(0,T;H)$. □

The last lemma provides us with a pair (X,ξ) of (potentially non-measurable) processes that serves as candidate solution to (6.1).

Proof of Proposition 6.3. We split the proof in several steps. We use the same symbols used in the proofs of the previous lemmata, without recalling their definitions explicitly.

STEP 1. We are going to pass to the limit on each trajectory $\omega \in \Omega'$ in the regularized equation

$$X_\lambda + \int_0^\cdot A_\lambda X_\lambda(s)\,ds + \int_0^\cdot \beta_\lambda(s, X_\lambda(s))\,ds = X_0 + G\cdot Z$$

along the subsequence λ'. Let then ω be fixed and let us omit its explicit indication. By Lemma 6.5 and the linearity of A, one has

$$\int_0^t A_{\lambda'} X_{\lambda'}(s)\,ds \longrightarrow \int_0^t AX(s)\,ds \qquad \text{weakly in } V',$$

hence also weakly in V_0', for every $t\in[0,T]$. Indeed, for any $\varphi \in V$ the map $\psi := s \mapsto 1\!\!1_{[0,t]}(s)\varphi$ belongs to $L^2(0,T;V)$ and

$$\begin{aligned}\Big\langle \varphi, \int_0^t A_{\lambda'} X_{\lambda'}(s)\,ds \Big\rangle &= \int_0^T \big\langle AJ_{\lambda'} X_{\lambda'}(s), \psi(s)\big\rangle\,ds\\ &\to \int_0^T \big\langle AX(s), \psi(s)\big\rangle\,ds = \Big\langle \varphi, \int_0^t AX(s)\,ds\Big\rangle.\end{aligned}$$

The same argument yields, choosing $\varphi \in L^\infty(D)$ or $\varphi\in V_0$, that

$$\int_0^t \beta_{\lambda'}(s, X_{\lambda'}(s))\,ds \longrightarrow \int_0^t \xi(s)\,ds$$

weakly in $L^1(D)$ and weakly in V_0' for all $t\in[0,T]$. Therefore, for every $t\in[0,T]$, there exists $\tilde X(t)\in V_0'$ such that $X_{\lambda'}(t)$ converges to $\tilde X(t)$ weakly in V_0'. From this it easily follows that $X_{\lambda'}$ converges to $\tilde X$ weakly* in $L^\infty(0,T;V_0')$. In fact, for any $\psi\in L^1(0,T;V_0)$, one has $\big\langle X_{\lambda'}(s),\psi(s)\big\rangle \to \big\langle \tilde X(s), \psi(s)\big\rangle$ for a.a. $s\in[0,T]$, and

$$\big|\big\langle X_{\lambda'}(s), \psi(s)\big\rangle\big| \lesssim \big\|X_\lambda\big\|_{L^\infty(0,T;H)} \big\|\psi(s)\big\|_{V_0},$$

where the right-hand side, as a function of s, belongs to $L^1(0,T)$. Then

$$\int_0^T \big\langle X_{\lambda'}(s), \psi(s)\big\rangle\,ds \longrightarrow \int_0^T \big\langle \tilde X(s), \psi(s)\big\rangle\,ds$$

by the dominated convergence theorem. However, since $X_{\lambda'}$ converges to X weakly* in $L^\infty(0,T;H)$, we infer that $X=\tilde X$ in $L^\infty(0,T;H)$. Therefore, taking the limit along λ', we get

$$X + \int_0^\cdot AX(s)\,ds + \int_0^\cdot \xi(s)\,ds = X_0 + G\cdot Z \qquad \text{in } V_0'.$$

This in turn implies that X is càdlàg in V_0', and since it also belongs to $L^\infty(0,T;H)$, it follows by Lemma 4.7 that X is weakly càdlàg in H.

STEP 2. We are going to prove that $j(\cdot,X)+j^*(\cdot,\xi)\in L^1(D_T)$ and $\xi\in\beta(\cdot,X)$ a.e. in D_T. Since $\beta_{\lambda'}(X_{\lambda'})$ converges weakly to ξ in $L^1(D_T)$, the weak lower semicontinuity of convex integrals (see, e.g., [7, Theorem 2.3, p. 18]) immediately yields

$$\int_{D_T} j^*(\xi)\,dx\,dt \leqslant \liminf_{\lambda'\to 0} \int_{D_T} j^*(\beta_{\lambda'}(X_{\lambda'}))\,dx\,dt$$

where the right-hand side is finite by Lemma 6.4 (here and below we do not explicitly denote the dependence of j and related maps on ω and t). Writing

$$\lambda\beta_\lambda(X_\lambda) = X_\lambda - (I+\lambda\beta)^{-1}X_\lambda,$$

the weak convergence of $\beta_\lambda(X_\lambda)$ is $L^1(D_T)$ implies its boundedness, hence the left-hand side of the previous identity converges to zero in $L^1(D_T)$ along λ'. Moreover, as $X_{\lambda'}$ converges to X in $L^2(0,T;H)$, it follows that $(I+\lambda'\beta)^{-1}X_{\lambda'}$ converges to X in $L^1(D_T)$. Therefore, again by lower semicontinuity of convex integrals,

$$\int_{D_T} j(X)\,dx\,dt \leqslant \liminf_{\lambda'\to 0} \int_{D_T} j\big((I+\lambda'\beta)^{-1}(X_{\lambda'})\big)\,dx\,dt, \tag{6.4}$$

where the right-hand side is finite because the integrand is bounded by $\beta_{\lambda'}(X_{\lambda'})X_{\lambda'}$ (see part (iii) of the proof of Lemma 6.4).

Let j_λ be the Moreau-Yosida regularization of j, i.e.

$$\begin{aligned} j_\lambda : \Omega\times[0,T]\times\mathbb{R} &\longrightarrow [0,+\infty[\\ (\omega,t,r) &\longmapsto \inf_{s\in\mathbb{R}}\Big(\frac{1}{2\lambda}|r-s|^2 + j(\omega,t,s)\Big). \end{aligned}$$

Recall that, for every $(\omega,t)\in\Omega\times[0,T]$, $j_\lambda(\omega,t,\cdot)$ is a convex differentiable function, with derivative equal to $\beta_\lambda(\omega,t,\cdot)$, that converges pointwise to $j(\omega,t,\cdot)$ from below. By definition of subdifferential one has, for any measurable set $E\subset D_T$,

$$\int_E \beta_\lambda(\cdot,X_\lambda)(X_\lambda - z)\,dx\,dt \geqslant \int_E j_\lambda(\cdot,X_\lambda)\,dx\,dt - \int_E j_\lambda(\cdot,z)\,dx\,dt \qquad \forall z\in L^\infty(E).$$

Since $X_{\lambda'}\to X_{\lambda'}$ in $L^2(0,T;H)$, there exists a subsequence of λ', denoted by same symbol for simplicity, such that $X_{\lambda'}\to X$ a.e. in D_T. Therefore, thanks to the Severini-Egorov theorem, for every $\eta>0$ there exists $E_\eta\subseteq D_T$, with $|D_T\setminus E_\eta|\leqslant\eta$, such that $X_{\lambda'}\to X$ uniformly on E_η. Choosing $E=E_\eta$ and passing to the limit along λ' in the last inequality yields

$$\int_{E_\eta}(X-z)\xi\,dx\,dt \geqslant \liminf_{\lambda'\to 0}\int_{E_\eta} j_{\lambda'}(X_{\lambda'})\,dx\,dt - \int_{E_\eta} j(z)\,dx\,dt \qquad \forall z\in L^\infty(E_\eta)$$

because $\beta_{\lambda'}(X_{\lambda'})$ converges weakly to ξ in $L^1(D_T)$ and $X_{\lambda'}$ converges to X uniformly on D_T, and $j_\lambda\leqslant j$. Moreover, by a well-known identity satisfied by the Moreau-Yosida regularization, one has

$$j_\lambda(X_\lambda) = j\big((I+\lambda\beta)^{-1}X_\lambda\big) + \frac{1}{2}\lambda\big(X_\lambda - (I+\lambda\beta)^{-1}X_\lambda\big)^2.$$

Since (X_λ) is bounded in $L^2(D_T)$ and $(I+\lambda\beta)^{-1}$ is a contraction on $\mathbb{R}$, it is easily seen that

$$\lambda'\int_{D_T}\big(X_{\lambda'} - (I+\lambda'\beta)^{-1}X_{\lambda'}\big)^2\,dx\,dt \longrightarrow 0$$

as $\lambda'\to 0$. By (6.4) it then follows

$$\int_{E_\eta}(X-z)\xi\,dx\,dt \geqslant \int_{E_\eta}\big(j(X)-j(z)\big)\,dx\,dt \qquad \forall z\in L^\infty(E_\eta).$$

By a suitable choice of z, this implies

$$(X-z)\xi \geqslant j(X)-j(z) \qquad \text{a.e. in } E_\eta \quad \forall z\in\mathbb{R}$$

(cf. [21] for a detailed argument in a slightly simpler setting), and hence that $\xi\in\partial j(X)=\beta(X)$ a.e. in E_η. Since η is arbitrary, it follows that $\xi\in\beta(X)$ a.e. in D_T.

STEP 3. We are now going to show that the solution pair (X,ξ) constructed in step 1 is unique. In particular, we claim that if there exist

$$X_i \in L^\infty(0,T;H) \cap L^2(0,T;V), \quad \xi_i \in L^1(D_T), \qquad i=1,2,$$

with $\xi_i \in \beta(\cdot, X_i)$ a.e. in D_T and $j(\cdot, X_i) + j^*(\cdot, \xi_i) \in L^1(D_T)$ such that

$$X_i + \int_0^\cdot AX_i(s)\,ds + \int_0^\cdot \xi_i(s)\,ds = X_0 + G\cdot Z,$$

then $(X_1,\xi_1) = (X_2,\xi_2)$. In fact, setting $X := X_1 - X_2$ and $\xi := \xi_1 - \xi_2$, one has

$$X + \int_0^\cdot AX(s)\,ds + \int_0^\cdot \xi(s)\,ds = 0,$$

where $X\xi$ belongs to $L^1(D_T)$: in fact, $X\xi \geqslant 0$ by monotonicity of β and, thanks to the convexity of j and j^* and to the hypothesis on their behavior at infinity, one has

$$\begin{aligned}\frac{1}{4}X\xi &\leqslant j(X/2) + j^*(\xi/2) = j\big(X_1/2 - X_2/2\big) + j^*\big(\xi_1/2 - \xi_2/2\big)\\ &\lesssim 1 + j(X_1) + j(X_2) + j^*(\xi_1) + j^*(\xi_2) \in L^1(D_T).\end{aligned}$$

By an argument completely analogous to the one used in the proof of Proposition 5.1 (in fact easier), one obtains

$$\|X\|^2 + \int_0^\cdot\!\!\int_D X(s)\xi(s)\,dx\,ds \leqslant 0.$$

Since the integrand in the previous identity is positive, it follows that $X=0$, which in turn implies that $\int_0^t \xi(s)\,ds = 0$ for all $t \in [0,T]$, hence also that $\xi = 0$, thus proving the claim.

STEP 4. The uniqueness result proved in the previous step allows us to show that the collection of pairs (X,ξ) indexed by $\omega \in \Omega'$ constructed in step 1 is in fact an optional process with values in $H \times L^1(D)$. This is far from obvious, mainly because X and ξ have been constructed, for each $\omega \in \Omega'$, as limits along subsequences λ' that depend themselves on ω. The crucial observation, which is an immediate consequence of the previous steps, is the following: from any subsequence of λ one can extract a further subsequence λ' (depending on ω) such that the convergences of Lemma 6.5 hold; but since the limits are unique, a classical result of elementary analysis ensures that the convergences hold along the original sequence λ, which is independent of ω. As X_λ converges to X in $L^2(0,T;H)$ $\mathbb{P}$-almost surely and (X_λ) is bounded in $L^2(\Omega;L^2(0,T;H))$, one has, passing to a subsequence if necessary, that X_λ converges to X weakly in $L^2(\Omega\times[0,T];H)$. Since (X_λ) is also bounded in $L^2(\Omega\times[0,T];V)$, it follows that X_λ, again passing to a subsequence if necessary, converges weakly to X in the latter space as well. Therefore there exists a sequence in the convex envelope of (X_λ) that converges strongly to X in $L^2(\Omega\times[0,T];V)$: since X_λ is adapted and càdlàg with values in H, hence optional, for every $\lambda > 0$, X is an H-valued optional process. Completely analogously, X is a (measurable) adapted V-valued process. In order to establish measurability properties of ξ, we need a more involved argument. Setting $\xi_\lambda := \beta_\lambda(\cdot, X_\lambda)$ for convenience, let $\phi \in L^\infty(D_T)$ and define

$$\Xi_\lambda := \int_{D_T} \xi_\lambda \phi\,dx\,dt, \qquad \Xi := \int_{D_T} \xi\phi\,dx\,dt,$$

so that Ξ_λ converges to Ξ $\mathbb{P}$-a.s. Jensen's inequality and part (iii) in the proof of Lemma 6.2 imply

$$j^*(\cdot,\Xi_\lambda) \lesssim_{|D_T|,\phi} \int_{D_T} j^*(\cdot,\xi_\lambda)\,dx\,dt \leqslant \int_{D_T} \xi_\lambda X_\lambda\,dx\,dt,$$

where the right-hand side, as a family indexed by λ, is bounded in $L^1(\Omega)$ by Lemma 6.2. Lemma 4.8 then implies that (Ξ_λ) is uniformly integrable in Ω and hence, by Vitali's theorem, that Ξ_λ converges to Ξ in $L^1(\Omega)$. Again the estimate $j^*(\cdot,\xi_\lambda) \leqslant \xi_\lambda X_\lambda$ implies, recalling that the right-hand side, as a family indexed by λ, is bounded in $L^1(\Omega \times D_T)$, that (ξ_λ) is uniformly integrable in $\Omega \times D_T$, hence relatively weakly compact as well, so that, by the Dunford-Pettis theorem, there exists $\tilde{\xi} \in L^1(\Omega \times D_T)$ such that ξ_λ converges weakly to $\tilde{\xi}$ in $L^1(\Omega \times [0,T]; L^1(D))$, from which it follows, by a reasoning already used, that $\tilde{\xi}$ is an optional $L^1(D)$-valued process. For every λ and $F \in \mathscr{F}$ one has, setting $h := \mathbb{1}_F \in L^\infty(\Omega)$,

$$\mathbb{E}\, h\Xi_\lambda = \int_{\Omega\times D_T} \xi_\lambda \phi h \, dx\, dt\, d\mathbb{P},$$

hence, passing to the limit as $\lambda \to 0$,

$$\mathbb{E}\, h\Xi = \int_{D_T} \bigl(\mathbb{E}\, h\xi\bigr)\phi \, dx\, dt = \int_{D_T} \bigl(\mathbb{E}\, h\tilde{\xi}\bigr)\phi \, dx\, dt.$$

Therefore $\mathbb{E}\,\mathbb{1}_F\xi = \mathbb{E}\,\mathbb{1}_F\tilde{\xi}$ in $L^1(D_T)$ for every $F \in \mathscr{F}$, i.e. $\xi = \tilde{\xi}$ in $L^1(D_T)$ $\mathbb{P}$-a.s.

Step 5. With the measurability properties of the processes X and ξ available, we can establish estimates of their moments. In fact, by the weak convergences of Lemma 6.5 and the estimates of Lemma 6.2, thanks to the weak and weak* lower semicontinuity of the norms, and to Fatou's lemma, it follows, writing $\xi_\lambda := \beta_\lambda(\cdot, X_\lambda)$, that

$$\begin{aligned}
\mathbb{E}\bigl\|X\bigr\|^2_{L^\infty(0,T;H)} &\leqslant \mathbb{E}\liminf_{\lambda\to 0}\bigl\|X_\lambda\bigr\|^2_{L^\infty(0,T;H)} &&\leqslant \liminf_{\lambda\to 0}\mathbb{E}\bigl\|X_\lambda\bigr\|^2_{L^\infty(0,T;H)},\\
\mathbb{E}\bigl\|X\bigr\|^2_{L^2(0,T;V)} &\leqslant \mathbb{E}\liminf_{\lambda\to 0}\bigl\|J_\lambda X_\lambda\bigr\|^2_{L^2(0,T;V)} &&\leqslant \liminf_{\lambda\to 0}\mathbb{E}\bigl\|J_\lambda X_\lambda\bigr\|^2_{L^2(0,T;V)},\\
\mathbb{E}\bigl\|\xi\bigr\|_{L^1(D_T)} &\leqslant \mathbb{E}\liminf_{\lambda\to 0}\bigl\|\xi_\lambda\bigr\|_{L^1(D_T)} &&\leqslant \liminf_{\lambda\to 0}\mathbb{E}\bigl\|\xi_\lambda\bigr\|_{L^1(D_T)},
\end{aligned}$$

where the right-hand sides are all finite. Similarly, the lower semicontinuity inequality

$$\int_{D_T} \bigl(j(\cdot,X) + j^*(\cdot,\xi)\bigr)\, dx\, dt \leqslant \liminf_{\lambda\to 0} \int_{D_T} \bigl(j(\cdot,X_\lambda) + j^*(\cdot,\xi_\lambda)\bigr)\, dx\, dt$$

yields, taking expectations on both sides and invoking Fatou's lemma,

$$\begin{aligned}
\mathbb{E}\int_{D_T} \bigl(j(\cdot,X) + j^*(\cdot,\xi)\bigr)\, dx\, dt &\leqslant \mathbb{E}\liminf_{\lambda\to 0} \int_{D_T} \bigl(j(\cdot,X_\lambda) + j^*(\cdot,\xi_\lambda)\bigr)\, dx\, dt\\
&\leqslant \liminf_{\lambda\to 0}\mathbb{E} \int_{D_T} \bigl(j(\cdot,X_\lambda) + j^*(\cdot,\xi_\lambda)\bigr)\, dx\, dt\\
&\leqslant \liminf_{\lambda\to 0}\bigl\|\xi_\lambda X_\lambda\bigr\|_{L^1(\Omega\times D_T)},
\end{aligned}$$

where the last term on the right-hand side is finite by Lemma 6.2.

Step 6. To conclude, let us show that the trajectories of X are càdlàg in H. Proposition 5.1 yields

$$\begin{aligned}
\|X\|^2 + 2\int_0^\cdot \langle AX(s), X(s)\rangle\, ds + 2\int_0^\cdot\int_D \xi(s)X(s)\, dx\, ds&\\
= \|X_0\|^2 + [G\cdot Z, G\cdot Z] + 2(X_-G)\cdot Z&,
\end{aligned} \tag{6.5}$$

where, by Fubini's theorem,

$$\int_D \xi X\, dx \leqslant \int_D j(\cdot,X)\, dx + \int_D j^*(\cdot,\xi)\, dx \in L^1(0,T),$$

thus also, taking into account that $X \in L^2(0,T;V)$ and $AX \in L^2(0,T;V')$,

$$\int_0^\cdot \langle AX(s), X(s)\rangle\, ds + \int_0^\cdot\int_D \xi(s)X(s)\, dx\, ds \in C([0,T]).$$

Furthermore, the last term on the right-hand side of (6.5) is càdlàg, being a stochastic integral with respect to a semimartingale. Recalling the definition of quadratic variation, the same reasoning applies to the second term on the right-hand side of (6.5). We deduce by inspection of (6.5) that the real-valued process $\|X\|^2$ is càdlàg. Since X is also weakly càdlàg in H (see step 1) and H is reflexive, we infer that the trajectories of X are also strongly càdlàg in H. In fact, let $t \in [0,T[$ and (t_n) a sequence converging to t from the right. Then $X(t_n) \to X(t)$ weakly in H and $\|X(t_n)\| \to \|X(t)\|$ imply that $X(t_n) \to X(t)$ in H. Similarly, if $t \in]0,T]$ and (t_n) is a sequence converging to t from the left, $X(t_n) \to X(t-)$ weakly in H and $\|X(t_n)\| \to \|X(t-)\|$ yield $X(t_n) \to X(t-)$ in H. □

In order to prove well-posedness of (6.1) without the extra regularity assumption (6.3) on the coefficient G, we prove continuity, in a suitable sense, of the map $(X_0, G) \mapsto X$.

Proposition 6.6. *Let (X_i, ξ_i), $i = 1, 2$, be strong solutions to (6.1) with initial conditions $X_{0i} \in L^2(\Omega; H)$ and coefficients $G_i \in \mathscr{S}_C(Z)$, respectively, where $C \in \mathscr{C}(Z)$ and $\mathbb{E}\,\lambda^C_{T-}(G_i) < \infty$. Then*

$$\begin{aligned}\mathbb{E}\big(X_1 - X_2\big)^{*2}_{T-} + \mathbb{E}\int_0^T \big\|X_1(t) - X_2(t)\big\|^2_V \, dt \\ \lesssim \mathbb{E}\big\|X_{01} - X_{02}\big\|^2 + \mathbb{E}\,\lambda^C_{T-}(G_1 - G_2),\end{aligned}$$

where the implicit constant depends only on the coercivity constant of A.

Proof. Setting

$$\begin{aligned} X &:= X_1 - X_2, & \xi &:= \xi_1 - \xi_2, \\ X_0 &:= X_{01} - X_{02}, & G &:= G_1 - G_2, \end{aligned}$$

one has

$$X + \int_0^\cdot AX(s)\,ds + \int_0^\cdot \xi(s)\,ds = X_0 + G \cdot Z.$$

In analogy to a reasoning already used, the hypotheses on j imply that

$$\frac{1}{4} X\xi \leqslant j(X/2) + j^*(\xi/2) \lesssim 1 + j(X_1) + j(X_2) + j^*(\xi_1) + j^*(\xi_2) \in L^1(D_T),$$

so that Proposition 5.1 yields

$$\begin{aligned}\frac{1}{2}\|X\|^2 + \int_0^\cdot \langle AX(s), X(s)\rangle\,ds + \int_0^\cdot\!\!\int_D \xi(s)X(s)\,dx\,ds \\ = \frac{1}{2}\|X_0\|^2 + \frac{1}{2}[G\cdot Z, G\cdot Z] + (X_- G)\cdot Z.\end{aligned}$$

Proceeding exactly as in the proof of Lemma 6.2, one has

$$\mathbb{E}[G\cdot Z, G\cdot Z]_{T-} \leqslant 5\,\mathbb{E}\,\lambda^C_{T-}(G)$$

and

$$\mathbb{E}\big((X_- G)\cdot Z\big)^*_{T-} \leqslant \frac{1}{4}\,\mathbb{E}\,X^{*2}_{T-} + 4\,\mathbb{E}\,\lambda^C_{T-}(G),$$

which immediately yield the claim by monotonicity and coercivity of A, and monotonicity of β. □

We are now in the position to prove Theorem 6.1.

Proof. Let us set, for every $n \in \mathbb{N}$, $G^n := T_n G$. Then G^n takes values in $\mathscr{L}(K, V_0)$ and

$$\|G^n\|_{\mathscr{L}(K,V_0)} \leqslant \|T_n\|_{\mathscr{L}(H,V_0)} \|G\|_{\mathscr{L}(K,H)},$$

so that G^n satisfies (6.3) for every $n \in \mathbb{N}$. Moreover, by the uniform boundedness of (T_n) in $\mathscr{L}(H)$, one has

$$\|G^n\|_{\mathscr{L}(K,H)} \leqslant \sup_{n\in\mathbb{N}} \|T_n\|_{\mathscr{L}(H)} \|G\|_{\mathscr{L}(K,H)},$$

so that, setting

$$\bar{C} := \sup_{n\in\mathbb{N}} \|T_n\|_{\mathscr{L}(H)} C \in \mathscr{C}(Z),$$

it follows that $G^n \in \mathscr{S}_{\bar{C}}(Z)$ for every $n \in \mathbb{N}$. Proposition 6.3 then ensures the existence and uniqueness of a strong solution (X^n, ξ^n) to (1.1) with data (X_0, G^n) for every $n \in \mathbb{N}$, i.e. such that

$$X^n + \int_0^\cdot AX^n(s)\,ds + \int_0^\cdot \xi_n(s)\,ds = X_0 + G^n \cdot Z. \tag{6.6}$$

Furthermore, by inspection of the proof of Lemma 6.2 it follows that

$$\mathbb{E}\|X^n\|^2_{L^\infty(0,T;H)} + \mathbb{E}\|X^n\|^2_{L^2(0,T;V)} + \mathbb{E}\|\xi^n X^n\|_{L^1(D_T)} \lesssim \mathbb{E}\|X_0\|^2 + \mathbb{E}\,\lambda^{\bar{C}}_{T-}(G^n),$$

where the implicit constant is independent of n. In particular, since

$$\begin{aligned}\lambda^{\bar{C}}_{T-}(G^n) &= \bar{C}(T-) \int_0^{T-} \|G^n(s)\|^2_{\mathscr{L}(K,H)}\,d\bar{C}(s)\\ &\leqslant \sup_{n\in\mathbb{N}} \|T_n\|^2_{\mathscr{L}(H)} \lambda^{\bar{C}}_{T-}(G) \in L^1(\Omega),\end{aligned}$$

there exists a constant N, independent of n, such that

$$\mathbb{E}\|X^n\|^2_{L^\infty(0,T;H)} + \mathbb{E}\|X^n\|^2_{L^2(0,T;V)} + \mathbb{E}\|\xi^n X^n\|_{L^1(D_T)} < N.$$

Moreover, since $\bar{C}$ does not depend on n, Proposition 6.6 implies that

$$\mathbb{E}\|X^{n_1} - X^{n_2}\|^2_{L^\infty(0,T;H)\cap L^2(0,T;V)} \lesssim \mathbb{E}\,\lambda^{\bar{C}}_{T-}\bigl(G^{n_1} - G^{n_2}\bigr) \qquad \forall n_1, n_2 \in \mathbb{N}.$$

By the properties of $(T_n)_n$ and the dominated convergence theorem, the right-hand side converges to zero as $n_1, n_2 \to \infty$, hence the sequence (X^n) is Cauchy in the space $L^2(\Omega; L^\infty(0,T;H)) \cap L^2(\Omega; L^2(0,T;V))$. As $X^n\xi^n = j(\cdot, X^n) + j^*(\cdot, \xi^n)$ and j is positive, $\bigl(j^*(\cdot,\xi^n)\bigr)$ is bounded in $L^1(\Omega \times (0,T) \times D)$, hence, taking Lemma 4.8 into account and arguing as in the proof of Lemma 6.5, it is easily seen that the sequence (ξ^n) is relatively compact in $L^1(\Omega \times [0,T] \times D)$. Therefore, passing to a subsequence if necessary,

$$\begin{aligned} X^n &\longrightarrow X \quad \text{in } L^2(\Omega; L^\infty(0,T;H)) \cap L^2(\Omega; L^2(0,T;V)),\\ \xi^n &\longrightarrow \xi \quad \text{weakly in } L^1(\Omega \times D_T).\end{aligned}$$

The first convergence implies that

$$\int_0^\cdot AX^n(s)\,ds \longrightarrow \int_0^\cdot AX(s)\,ds \qquad \text{in } L^2(\Omega; C([0,T];V')),$$

and that $\bigl(X^n - X\bigr)^*_{T-} \to 0$ in $L^2(\Omega)$, because X^n has càdlàg trajectories for each $n \in \mathbb{N}$ thanks to Proposition 6.3. In particular, X has càdlàg trajectories as well. The uniform boundedness of (T_n) in $\mathscr{L}(H)$ and the dominated convergence theorem for stochastic integrals yield

$$\bigl(G^n \cdot Z - G \cdot Z\bigr)^*_{T-} \longrightarrow 0 \qquad \text{in } L^2(\Omega).$$

From $\Delta X^n(T) = \Delta(G^n \cdot Z)_T = G^n_T \Delta Z_T$ and the above uniform convergences up to $T-$ it immediately follows that

$$\big(X^n - X\big)^*_T \longrightarrow 0, \qquad \big(G^n \cdot Z - G \cdot Z\big)^*_T \longrightarrow 0$$

in $L^0(\Omega)$ as $n \to \infty$. Let $\phi \in V_0$ and $F \in \mathscr{F}$. Recalling that $\xi^n \to \xi$ weakly in $L^1(\Omega \times D_T)$, one has

$$\mathbb{E}\,\mathbb{1}_F \Big\langle \phi, \int_0^t \xi^n(s)\,ds \Big\rangle \longrightarrow \mathbb{E}\,\mathbb{1}_F \Big\langle \phi, \int_0^t \xi(s)\,ds \Big\rangle \qquad \forall t \in [0,T].$$

Taking the duality product of both sides of (6.6) with $\phi \in V_0$ and multiplying by $\mathbb{1}_F$, one readily infers, passing to the limit as $n \to \infty$ and taking into account that φ and F are arbitrary, that

$$X(t) + \int_0^t AX(s)\,ds + \int_0^t \xi(s)\,ds = X_0 + \big(G \cdot Z\big)_t \qquad \forall t \in [0,T]$$

as an identity in V_0'. Since both sides of the equality are immediately seen to be càdlàg (with values in V_0'), it follows that equality holds in V_0' also in the sense of indistinguishability, not only in the sense of modifications. By comparison, the identity also holds in $V' \cap L^1(D)$. Moreover, arguing as in step 2 of the proof of Proposition 6.3, we deduce that $\xi \in \beta(X)$ a.e. in $\Omega \times (0,T) \times D$. The uniqueness of (X, ξ) follows by an argument completely analogous to the one used in step 3 of the proof of Proposition 6.3, appealing to the integration-by-parts formula of Proposition 5.1. □

Suitably localized versions of the previous results hold.

Proposition 6.7. *Let $\tau \neq 0$ be a stopping time with $\tau \leqslant T$, C a control process for Z, and G a strongly predictable process such that $\mathbb{E}\,\lambda^C_{\tau-}(G) < \infty$. If $X_0 \in L^2(\Omega; H)$, then* (6.1) *admits a unique strong solution on $[\![0,\tau]\!]$.*

Proof. Let us consider the equation

$$d\tilde{X} + A\tilde{X}\,dt + \beta(\tilde{X})\,dt \ni G\,dZ^{\tau-}, \qquad X(0) = X_0, \tag{6.7}$$

where $Z^{\tau-}$ is a semimartingale with control process $C^{\tau-}$ (see Lemma 4.4). Since

$$\lambda^{C^{\tau-}}_{T-}(G) = C^{\tau-}_{T-}\big(\|G\|^2 \cdot C^{\tau-}\big)_{T-} = C_{\tau-}\big(\|G\|^2 \cdot C\big)_{\tau-} = \lambda^C_{\tau-}(G)$$

where the expectation of the last term is finite by assumption, equation (6.7) admits a unique strong solution $(\tilde{X}, \tilde{\xi})$. In particular,

$$\tilde{X} + \int_0^\cdot A\tilde{X}(s)\,ds + \int_0^\cdot \tilde{\xi}(s)\,ds = X_0 + G \cdot Z^{\tau-},$$

which implies that $\Delta\tilde{X}_\tau = 0$, because the Lebesgue integrals and the stochastic integral have no jump at τ. Setting $X = \tilde{X}$ on $[\![0,\tau[\![$ and $X_\tau := X_{\tau-} + G_\tau \Delta Z_\tau$, and $\xi := \mathbb{1}_{[\![0,\tau]\!]}\tilde{\xi}$, we are left with

$$X^\tau + \int_0^{\cdot\wedge\tau} AX(s)\,ds + \int_0^{\cdot\wedge\tau} \xi(s)\,ds = X_0 + (G\mathbb{1}_{[\![0,\tau]\!]}) \cdot Z,$$

i.e. (X,ξ) is a strong solution on $[\![0,\tau]\!]$ to (6.1). Since $\tilde{\xi} \in \beta(\tilde{X})$ a.e in $\Omega \times (0,T) \times D$, we have in particular that $\xi \in \beta(X)$ a.e. in $[\![0,\tau[\![\times D$. To prove uniqueness it suffices to note that a strong solution (X,ξ) on $[\![0,\tau]\!]$ to (6.1) coincides on $[\![0,\tau[\![$ with the restriction to $[\![0,\tau[\![$ of the unique strong solution $(\tilde{X},\tilde{\xi})$ to (6.7). Uniqueness on the closed stochastic interval $[\![0,\tau]\!]$ follows by the definition of X_τ. □

As an immediate consequence of the uniqueness argument just used, one obtains that (strong) solutions on closed stochastic intervals form a direct system, in the following sense: if (X,ξ) is a solution on $[\![0,\tau]\!]$ to (6.1) and σ is a stopping time with $\sigma \leqslant \tau$, it is easily seen that (X^σ,ξ^σ) is a solution on $[\![0,\sigma]\!]$ to (6.1). Such a solution, by the reasoning of the previous remark, is the unique solution on $[\![0,\sigma]\!]$. This also implies that, given (X_1,ξ_1) solution on $[\![0,\tau_1]\!]$ and (X_2,ξ_2) solution on $[\![0,\tau_2]\!]$, one can construct a solution (X,ξ) on $[\![0,\tau_1 \vee \tau_2]\!]$ setting

$$(X,\xi) := \begin{cases} (X_1,\xi_1) & \text{on } [\![0,\tau_1]\!], \\ (X_2,\xi_2) & \text{on } [\![0,\tau_2]\!]. \end{cases}$$

Proposition 6.8. *Let (X_i,ξ_i) be strong solutions on $[\![0,\tau_i]\!]$, $i=1,2$, to (6.1) with initial conditions $X_{0i} \in L^2(\Omega;H)$ and coefficients $G_i \in \mathscr{S}_C(Z)$, respectively, where C is a control process for the semimartingale Z and $\mathbb{E}\,\lambda^C_{\tau_i-}(G_i) < \infty$. Setting $\tau := \tau_1 \wedge \tau_2$, one has*

$$\begin{aligned} &\mathbb{E}\big(X_1 - X_2\big)^{*2}_{\tau-} + \mathbb{E}\int_0^\tau \big\|X_1(t) - X_2(t)\big\|_V^2\,dt \\ &\qquad \lesssim \mathbb{E}\big\|X_{01} - X_{02}\big\|^2 + \mathbb{E}\,\lambda^C_{\tau-}(G_1 - G_2). \end{aligned}$$

Proof. By the above discussion about strong solutions on closed stochastic intervals forming a direct system, it is immediately seen that (X_1,ξ_1) and (X_2,ξ_2) are strong solutions on $[\![0,\tau]\!]$, as well as that $(X_i^{\tau-},\xi_i^{\tau-}) = (\tilde{X}_i^{\tau-},\tilde{\xi}_i^{\tau-})$, where $(\tilde{X}_i,\tilde{\xi}_i)$ is the unique strong solution to

$$d\tilde{X}_i + A\tilde{X}_i\,dt + \beta(\cdot,\tilde{X}_i)\,dt \ni G_i\,dZ^{\tau-}, \qquad \tilde{X}_i(0) = X_{0i}.$$

Since $\mathbb{E}\,\lambda^{C^{\tau-}}_{T-}\big(G_1 - G_2\big) = \mathbb{E}\,\lambda^C_{\tau-}(G_1 - G_2)$ and $C^{\tau-}$ is a control process for $Z^{\tau-}$ by Lemma 4.4, Proposition 6.6 yields

$$\begin{aligned} &\mathbb{E}\big(X_1 - X_2\big)^{*2}_{\tau-} + \mathbb{E}\int_0^\tau \big\|X_1(t) - X_2(t)\big\|_V^2\,dt \\ &\qquad \lesssim \mathbb{E}\big\|X_{01} - X_{02}\big\|^2 + \mathbb{E}\,\lambda^C_{\tau-}(G_1 - G_2). \end{aligned}$$ □

7 Well-posedness with multiplicative noise

This section is devoted to the proof of Theorem 3.2. We begin showing that strong solutions on closed stochastic intervals exist.

Proposition 7.1. *There exists a stopping time $\tau \neq 0$ and a strong solution on $[\![0,\tau]\!]$ to* (1.1).

Proof. Let $\alpha \in \,]0,1[$ be a constant to be chosen later, C a control process for Z, and τ^0 the stopping time defined as

$$\tau^0 := \inf\big\{t \in [0,T] : C_t(L_t - L_0) \geqslant \alpha\big\} \wedge T.$$

Note that τ^0 is well-defined and not identically 0 as the process $C(L - L_0)$ starts from 0 and is right-continuous. Let $R \in \mathbb{R}_+$ be such that the event $\{\|X_0\| \leqslant R\}$ has strictly positive probability, and set $\tau := \tau^0 \mathbb{1}_F$. Since $F \in \mathscr{F}_0$, it is easily seen that τ is a stopping time. Let $\mathbb{S}^2(T-)$ denote the vector space of adapted càdlàg processes $Y : \Omega \times [0,T[\, \to H$ such that

$$\|Y\|_2 := \big(\mathbb{E}\,Y^{*2}_{T-}\big)^{1/2} < \infty.$$

It is not difficult to see that $\mathbb{S}^2(T-)$, endowed with the norm $\|\cdot\|_2$, is a Banach space. For every $Y \in \mathbb{S}^2(T-)$ one has

$$\begin{aligned} \lambda^{C^{\tau-}}_{T-}\big(B(Y)\big) &= C^{\tau-}_{T-}\int_0^{T-} \big\|[B(Y)](s)\big\|^2_{\mathscr{L}(K,H)}\,dC^{\tau-}_s \\ &\leqslant C_{\tau-}(L_{\tau-} - L_0)\big(1 + Y^{*2}_{T-}\big) \leqslant \alpha\big(1 + Y^{*2}_{T-}\big) \in L^1(\Omega), \end{aligned}$$

so that the equation

$$d\tilde{X}(t) + A\tilde{X}(t)\,dt + \beta(\tilde{X}(t))\,dt \ni B(Y)\,dZ_t^{\tau-}, \qquad X(0) = X_0,$$

admits a unique strong solution $(\tilde{X}, \tilde{\xi})$ by Theorem 6.1 (by the definition of the stopping time τ, the latter result is indeed applicable). In particular, the map $Y \mapsto \tilde{X}$ is a homomorphism of $\mathbb{S}^2(T-)$. Moreover, for any Y_1, $Y_2 \in \mathbb{S}^2(T-)$, Proposition 6.6 yields, with obvious meaning of the notation,

$$\big\|\tilde{X}_1 - \tilde{X}_2\big\|_2^2 + \mathbb{E}\big\|\tilde{X}_1 - \tilde{X}_2\big\|^2_{L^2(0,T;V)} \lesssim \mathbb{E}\,\lambda_{T-}^{C^{\tau-}}\big(B(Y_1) - B(Y_2)\big),$$

where, by the Lipschitz assumption on B,

$$\begin{aligned}
&\lambda_{T-}^{C^{\tau-}}\big(B(Y_1) - B(Y_2)\big)\\
&\qquad = C_{T-}^{\tau-}\int_0^{T-} \big\|[B(Y_1)](s) - [B(Y_2)](s)\big\|^2_{\mathscr{L}(K,H)}\,dC_s^{\tau-}\\
&\qquad \leqslant C(\tau-)\int_0^{T-} \big(Y_1 - Y_2\big)^{*2}_{s-}\,dL^{\tau-}(s) \leqslant \alpha\big(Y_1 - Y_2\big)^{*2}_{T-},
\end{aligned}$$

which implies

$$\big\|\tilde{X}_1 - \tilde{X}_2\big\|_2^2 + \mathbb{E}\big\|\tilde{X}_1 - \tilde{X}_2\big\|^2_{L^2(0,T;V)} \lesssim \alpha\big\|Y_1 - Y_2\big\|_2^2.$$

Choosing α small enough, $Y \mapsto \tilde{X}$ is a contraction of $\mathbb{S}^2(T-)$, hence it admits a unique fixed point $\tilde{X} \in \mathbb{S}^2(T-)$ (the abuse of notation is harmless). Setting $X := X_0 1\!\!1_{\{\tau=0\}} + \tilde{X}$ in $[\![0,\tau[\![$, $X_\tau := X_{\tau-} + [B(\tilde{X})]_\tau \Delta Z_\tau$, and $\xi := \tilde{\xi}1\!\!1_{[\![0,\tau]\!]}$, it is immediately seen that (X,ξ) is a strong solution on $[\![0,\tau]\!]$ to (1.1). □

Once existence of solutions on stochastic intervals is established, we establish their uniqueness in a local sense.

Lemma 7.2. *Let (X_1, ξ_1) and (X_2, ξ_2) be strong solutions to (1.1) on $[\![0,\tau_1]\!]$ and $[\![0,\tau_2]\!]$, respectively. Then, setting $\tau := \tau_1 \wedge \tau_2$, one has $X_1 = X_2$ and $\xi_1 = \xi_2$ on $[\![0,\tau]\!]$.*

Proof. Setting $X := X_1 - X_2$ and $\xi := \xi_1 - \xi_2$, one has

$$X^\tau + \int_0^\cdot 1\!\!1_{[\![0,\tau]\!]} AX(s)\,ds + \int_0^\cdot 1\!\!1_{[\![0,\tau]\!]}\xi(s)\,ds = \big(1\!\!1_{[\![0,\tau]\!]}(B(X_1) - B(X_2)\big)\cdot Z, \tag{7.1}$$

where $B(X_1) \in \mathscr{S}_{C_1}(Z)$, $B(X_2) \in \mathscr{S}_{C_2}(Z)$, with C_1 and C_2 control processes for Z. Recalling that $C := C_1 + C_2$ is a control process for Z, let us set, for every $k \in \mathbb{N}$,

$$\tau_k^0 := \inf\big\{t \in [0,T] : C(t)(L(t) - L(0)) \geqslant k\big\} \wedge \tau$$

and $\tau_k := \tau_k^0 1\!\!1_{F_k}$, where F_k is the event $\{\|X_0\| \leqslant k\}$. By the hypotheses on B it follows that

$$\begin{aligned}
\lambda^C_{\tau_k-}\big(B(X_i)\big) &= C_{\tau_k-}\int_0^{\tau_k-} \big\|[B(X_i)](s)\big\|^2_{\mathscr{L}(K,H)}\,dC_s\\
&\leqslant C_{\tau_k-}\int_0^{\tau_k-} \big(1 + (X_i)^{*2}_{s-}\big)\,dL_s\\
&\leqslant C_{\tau_k-}(L_{\tau_k-} - L_0)\big(1 + (X_i)^{*2}_{\tau_k-}\big)\\
&\leqslant k\big(1 + (X_i)^{*2}_{\tau_k-}\big) \in L^1(\Omega).
\end{aligned}$$

Hence, for every stopping time $\sigma \leqslant \tau_k$, Proposition 6.6 yields

$$\mathbb{E}\,X^{*2}_{\sigma-} + \mathbb{E}\int_0^\sigma \|X(s)\|_V^2\,ds \lesssim \mathbb{E}\,\lambda^C_{\sigma-}\big(B(X_1) - B(X_2)\big),$$

thus also, by the Lipschitz continuity of B,

$$\mathbb{E}\big(X_1 - X_2\big)^{*2}_{\sigma-} \lesssim k\,\mathbb{E}\big((X_1 - X_2)^{*2} \cdot L\big)_{\sigma-},$$

which implies, by Lemma 4.6, that $\mathbb{E}\big(X_1 - X_2\big)^{*2}_{\tau_k-} = 0$ for every $k \in \mathbb{N}$. Since τ_k tends monotonically to τ as $k \to \infty$, it immediately follows that $X_1 = X_2$ on $[\![0,\tau[\![$. This implies that $B(X_1) = B(X_2)$ on $[\![0,\tau]\!]$, hence the jumps at τ of X_1 and X_2 are both equal to $[B(X_1)]_\tau \Delta Z_\tau$, so that $X_1 = X_2$ on $[\![0,\tau]\!]$. Finally, by comparison in (7.1), one gets $\int_0^\cdot \xi(s)\,ds = 0$, which implies also $\xi_1 = \xi_2$. □

Let us now come to the core of the proof of Theorem 3.2. The idea is simply to iterate the construction of Proposition 7.1, to obtain a solution on a sequence of stochastic intervals $[\![\tau_n, \tau_{n+1}]\!]$, $n \in \mathbb{N}$, and to show that $\mathbb{P}(\tau_n < T)$ tends to zero as $n \to \infty$. Calling τ_1 the stopping time given by Proposition 7.1, let us define the increasing sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ defined as

$$\tau_{n+1} := \begin{cases} \tau_n, & \text{if } \|X(\tau_n)\| > n,\\ \inf\big\{t \geqslant \tau_n :\ C_t(L_t - L_{\tau_n}) > \alpha\big\} \wedge T, & \text{if } \|X(\tau_n)\| \leqslant n, \end{cases}$$

where α is a constant as chosen in the proof of Proposition 7.1. Note that τ_{n+1} is indeed a stopping time because the event $\{\|X(\tau_n)\| > n\}$ belongs to $\mathscr{F}_{\tau_n}$. Proposition 7.1 yields the existence of a strong solution on $[\![\tau_n, \tau_{n+1}]\!]$ to equation (1.1) started at τ_n. A standard patching argument shows that one thus obtains a strong solution (X_n, ξ_n) on $[\![0, \tau_n]\!]$ for every $n \in \mathbb{N}$.

We are going to show that $\mathbb{P}(\lim_n \tau_n < T) = 0$. Assume, by contradiction, that $\mathbb{P}(\lim_n \tau_n < T) > 0$. One can rule out that $\tau_{n+1} \neq \tau_n$ occurs only a finite number of times. In fact, if it were the case, then there would exist $\bar{n} \in \mathbb{N}$ such that $\|X(\tau_{\bar{n}})\|$ is larger than every natural number on an event of positive probability. This is impossible, because X_{τ_n} is a well-defined H-valued random variable for all $n \in \mathbb{N}$. This implies that, on an event F of strictly positive probability, $L_{\tau_{n+1}} - L_{\tau_n} > 0$ for every n belonging to an infinite subset $\mathbb{N}'$ of $\mathbb{N}$. Since C is increasing, one has

$$L_{\tau_{n+1}} - L_{\tau_n} > \frac{\alpha}{C_{\tau_n}} \geqslant \frac{\alpha}{C_T} \qquad \forall n \in \mathbb{N}',$$

hence denoting the variation of L by $|L|$ and recalling that L is also increasing,

$$|L| \geqslant \sum_{n\in\mathbb{N}'} \big|L_{\tau_{n+1}} - L_{\tau_n}\big| = \infty \qquad \text{on } F.$$

This contradicts the hypotheses on L, therefore $\tau_n \to T$ $\mathbb{P}$-a.s. as $n \to \infty$. The solution constructed above is thus defined on the whole interval $[0,T]$. Furthermore, such a solution is also unique, thanks to Lemma 7.2.

An argument entirely analogous to the one used in the proof of Lemma 7.2 yields, bearing in mind the definition of τ_n,

$$\mathbb{E}\,X^{*2}_{\tau_n-} + \mathbb{E}\int_0^{\tau_n} \|X(s)\|_V^2\,ds + \mathbb{E}\int_0^{\tau_n}\!\!\int_D \xi(s)X(s)\,dx\,ds \lesssim n^2 \qquad \forall n \in \mathbb{N},$$

hence, in particular,

$$X^{*2}_{\tau_n-} + \int_0^{\tau_n} \|X(s)\|_V^2\,ds + \int_0^{\tau_n}\!\!\int_D \xi(s)X(s)\,dx\,ds$$

is finite $\mathbb{P}$-a.s. for all $n \in \mathbb{N}$. Since $X^*_{\tau_n} \leqslant X^*_{\tau_n-} + \|\Delta X(\tau_n)\|$ and, for all ω in an event of probability one, there exists $\bar{n}$ such that $\tau_n(\omega) = T$ for all $n \geqslant \bar{n}$, it follows that

$$X^{*2}_T + \int_0^T \|X(s)\|_V^2\,ds + \int_0^T\!\!\int_D \xi(s)X(s)\,dx\,ds < \infty$$

with probability one.

Let us now turn to the continuity with respect to the initial datum. Let (X_{0n}) be a sequence of $\mathscr{F}_0$-measurable random variables such that $X_{0n} \to X_0$ in probability, and let X_n be the unique solution to (1.1) with initial datum X_{0n}. Then there exists a subsequence $(X_{0n'})$ converging to X_0 $\mathbb{P}$-almost surely. Setting

$$S_k := \bigcap_{n' \geqslant k} \bigl\{ \|X_{0n'} - X_0\| \leqslant 1 \bigr\},$$

it is clear that (S_k) is an increasing sequence of elements of $\mathscr{F}_0$ whose limit as $k \to \infty$ is an event of probability one. In fact,

$$\mathbb{P}(S_k) = \mathbb{P}\bigl(\|X_{0n'} - X_0\| \leqslant 1 \ \forall n' \geqslant k\bigr),$$

which converges to one as $k \to \infty$ by definition of almost sure convergence. Moreover, $(X_{0n} - X_0)\mathbb{1}_{S_k}$ obviously converges to zero in probability as $n \to \infty$ for every k, and

$$\bigl\|(X_{0n} - X_0)\mathbb{1}_{S_k}\bigr\| \leqslant 1 \qquad \forall n \geqslant k.$$

Therefore, by the dominated convergence theorem, $(X_{0n} - X_0)\mathbb{1}_{S_k}$ converges to zero in $L^2(\Omega; H)$ as $n \to \infty$ for each k. Let (τ_k) be an increasing sequence of stopping times converging to T, for instance as the one constructed above, and define a new sequence of stopping times (σ_k) as $\sigma_k := \tau_k \mathbb{1}_{S_k}$. Then a (by now) familiar reasoning using Itô's formula for the square of the norm, stopping at σ_k-, and applying the stochastic Gronwall lemma, much as in the proof of Lemma 7.2, yields

$$\mathbb{E}\bigl(X - X_n\bigr)^{*2}_{\sigma_k-} + \mathbb{E}\int_0^{\sigma_k} \|(X - X_n)(s)\|_V^2 \, ds \lesssim \mathbb{E}\bigl\|X_0 - X_{0n}\bigr\|^2 \mathbb{1}_{S_k},$$

where the right-hand side converges to zero as $n \to \infty$ for every k. We have thus shown that X_n converges to X prelocally in $\mathbb{S}^2(T)$. Since T was arbitrary and all results continue to hold if T is replaced by, e.g., $T+1$, X^n converges to X prelocally also in $\mathbb{S}^2((T+1))$, which implies that $(X_n - X)^*_T$ converges to zero in probability (see, e.g., [26, p. 261]). The proof of Theorem 3.2 is thus completed.

References

[1] S. Albeverio, H. Kawabi, and M. Röckner, *Strong uniqueness for both Dirichlet operators and stochastic dynamics to Gibbs measures on a path space with exponential interactions*, J. Funct. Anal. **262** (2012), no. 2, 602–638. MR 2854715

[2] V. Barbu, *Nonlinear differential equations of monotone types in Banach spaces*, Springer, New York, 2010. MR 2582280

[3] ———, *A variational approach to stochastic nonlinear parabolic problems*, J. Math. Anal. Appl. **384** (2011), no. 1, 2–15. MR 2822846

[4] Stefano Bonaccorsi, Carlo Marinelli, and Giacomo Ziglio, *Stochastic FitzHugh-Nagumo equations on networks with impulsive noise*, Electron. J. Probab. **13** (2008), no. 49, 1362–1379. MR 2438810 (2010d:60143)

[5] Z. Brzeźniak, Wei Liu, and Jiahui Zhu, *Strong solutions for SPDE with locally monotone coefficients driven by Lévy noise*, Nonlinear Anal. Real World Appl. **17** (2014), 283–310. MR 3158475

[6] B. Gess and J. M. Tölle, *Multi-valued, singular stochastic evolution inclusions*, J. Math. Pures Appl. (9) **101** (2014), no. 6, 789–827. MR 3205643

[7] M. Giaquinta, *Multiple integrals in the calculus of variations and nonlinear elliptic systems*, Princeton University Press, Princeton, NJ, 1983. MR 717034

[8] I. Gyöngy, *On stochastic equations with respect to semimartingales. III*, Stochastics **7** (1982), no. 4, 231–254.

[9] I. Gyöngy and N. V. Krylov, *On stochastics equations with respect to semimartingales II. Itô formula in Banach spaces*, Stochastics **6** (1981/82), no. 3-4, 153–173. MR MR665398 (84m:60070a)

[10] M. Haase, *Convexity inequalities for positive operators*, Positivity **11** (2007), no. 1, 57–68. MR 2297322 (2008d:39034)

[11] Ch. Hess, *On the measurability of the conjugate and the subdifferential of a normal integrand*, J. Convex Anal. **2** (1995), no. 1-2, 153–165. MR 1363366

[12] J. Jacod, *Une condition d'existence et d'unicité pour les solutions fortes d'équations différentielles stochastiques*, Stochastics **4** (1980/81), no. 1, 23–38. MR 587427 (82j:60105)

[13] N. V. Krylov and B. L. Rozovskiĭ, *Stochastic evolution equations*, Current problems in mathematics, Vol. 14 (Russian), Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Informatsii, Moscow, 1979, pp. 71–147, 256. MR MR570795 (81m:60116)

[14] E. Lenglart, D. Lépingle, and M. Pratelli, *Présentation unifiée de certaines inégalités de la théorie des martingales*, Séminaire de Probabilités, XIV (Paris, 1978/1979), Lecture Notes in Math., vol. 784, Springer, Berlin, 1980, pp. 26–52. MR 580107 (82d:60087)

[15] W. Liu and M. Röckner, *Stochastic partial differential equations: an introduction*, Springer, Cham, 2015. MR 3410409

[16] W. Liu and M. Stephan, *Yosida approximations for multivalued stochastic partial differential equations driven by Lévy noise on a Gelfand triple*, J. Math. Anal. Appl. **410** (2014), no. 1, 158–178. MR 3109829

[17] C. Marinelli, *Local well-posedness of Musiela's SPDE with Lévy noise*, Math. Finance **20** (2010), no. 3, 341–363. MR 2667893

[18] ———, *Well-posedness for a class of dissipative stochastic evolution equations with Wiener and Poisson noise*, Seminar on Stochastic Analysis, Random Fields and Applications VII, Birkhäuser/Springer, Basel, 2013, pp. 187–196. MR 3380100

[19] C. Marinelli and M. Röckner, *Well-posedness and asymptotic behavior for stochastic reaction-diffusion equations with multiplicative Poisson noise*, Electron. J. Probab. **15** (2010), no. 49, 1528–1555. MR 2727320

[20] C. Marinelli and L. Scarpa, *Refined existence and regularity results for a class of semilinear dissipative SPDEs*, arXiv:1711.11091.

[21] ———, *Strong solutions to SPDEs with monotone drift in divergence form*, Stoch. Partial Differ. Equ. Anal. Comput. **6** (2018), no. 3, 364–396. MR 3844654

[22] ———, *A variational approach to dissipative SPDEs with singular drift*, Ann. Probab. **46** (2018), no. 3, 1455–1497. MR 3785593

[23] M. Métivier, *Semimartingales*, Walter de Gruyter & Co., Berlin, 1982. MR MR688144 (84i:60002)

[24] M. Métivier and J. Pellaumail, *Stochastic integration*, Academic Press, New York, 1980. MR MR578177 (82b:60060)

[25] E. Pardoux, *Equations aux derivées partielles stochastiques nonlinéaires monotones*, Ph.D. thesis, Université Paris XI, 1975.

[26] Ph. E. Protter, *Stochastic integration and differential equations*, second ed., Springer-Verlag, Berlin, 2004. MR MR2020294 (2005k:60008)

[27] R. T. Rockafellar, *Integral functionals, normal integrands and measurable selections*, Nonlinear operators and the calculus of variations (Summer School, Univ. Libre Bruxelles, Brussels, 1975), Lecture Notes in Math., vol. 543, Springer, Berlin, 1976, pp. 157–207. MR 0512209

[28] J. Simon, *Compact sets in the space* $L^p(0, T; B)$, Ann. Mat. Pura Appl. (4) **146** (1987), 65–96. MR 916688 (89c:46055)

[29] W. A. Strauss, *On continuity of functions with values in various Banach spaces*, Pacific J. Math. **19** (1966), 543–551. MR 0205121 (34 #4956)