A variational approach to dissipative SPDEs with singular drift

Carlo Marinelli* Luca Scarpa†

March 24, 2017

Abstract

We prove global well-posedness for a class of dissipative semilinear stochastic evolution equations with singular drift and multiplicative Wiener noise. In particular, the nonlinear term in the drift is the superposition operator associated to a maximal monotone graph everywhere defined on the real line, on which neither continuity nor growth assumptions are imposed. The hypotheses on the diffusion coefficient are also very general, in the sense that the noise does not need to take values in spaces of continuous, or bounded, functions in space and time. Our approach combines variational techniques with a priori estimates, both pathwise and in expectation, on solutions to regularized equations.

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

Key words and phrases: stochastic evolution equations, singular drift, variational approach, well-posedness, multiplicative noise, monotone operators.

1 Introduction

Our aim is to establish existence and uniqueness of solutions, and their continuous dependence on the initial datum, to the following semilinear stochastic evolution equation on $L^2(D)$, with $D \subset \mathbb{R}^n$ a bounded domain:

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

where A is a linear maximal monotone operator on $L^2(D)$ associated to a coercive Markovian bilinear form, β is a maximal monotone graph in $\mathbb{R} \times \mathbb{R}$ defined everywhere, W is a cylindrical Wiener process on a separable Hilbert space U, and B takes values in the space of Hilbert-Schmidt operators from U to $L^2(D)$ and satisfies suitable Lipschitz continuity assumptions. Precise assumptions on the data of the problem and on the definition of solution are given in Section 2 below. Since any increasing function $\beta_0 : \mathbb{R} \to \mathbb{R}$ can be extended in a canonical way to a maximal monotone graph of $\mathbb{R} \times \mathbb{R}$ by "filling the gaps" (i.e., setting $\beta(x) := [\beta_0(x^-), \beta_0(x^+)]$ for all $x \in \mathbb{R}$, where $\beta(x^-)$

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

†Department of Mathematics, University College London, Gower Street, London WC1E 6BT, United Kingdom. E-mail: luca.scarpa.15@ucl.ac.uk

and $\beta(x^+)$ denote the limit from the left and from the right of β_0 at x, respectively), Equation (1.1) can be interpreted as a formulation of the stochastic evolution equation

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

Semilinear equations with singular and rapidly growing drift appear, for instance, in mathematical models of Euclidean quantum field theory (see, e.g., [1] for an equation with exponentially growing drift), and, most importantly for us, cannot be directly treated with the existing methods, hence are interesting from a purely mathematical perspective as well. In particular, the variational approach (see [24, 33]) works only assuming that β satisfies suitable polynomial growth conditions depending on the dimension n of the underlying Euclidean space (see also [28, pp. 137-ff.] for improved sufficient conditions, still dependent on the dimension), whereas most available results relying on the semigroup approach require just polynomial growth, although usually compensated by rather stringent hypotheses on the noise (see, e.g., [15, 16]). Under natural assumptions on the noise, well-posedness in L^p spaces is proven, with different methods, in [25], under the further assumption that β is locally Lipschitz continuous, and in [30]. A common basis for both works is the semigroup approach on UMD Banach spaces. A special mention deserves the short note [6], where the author considers problem (1.1) with $A = -\Delta$ and B independent of X, and proves existence of a pathwise solution[1] assuming that the solution Z to the equation with $\beta \equiv 0$ (i.e., the stochastic convolution) is jointly continuous in space and time. Furthermore, assuming that

$$\mathbb{E}\int_0^T\int_D j(Z) < \infty,$$

where j is a primitive of β, he obtains that the pathwise solution may admit a version that can be considered as a generalized mild solution to (1.1). This is the only result we are aware of about existence of solutions to stochastic semilinear parabolic equations *without* growth assumptions on the drift in any dimension. It is well known that a well-posedness theory for stochastic evolution equations on a Hilbert space H of the type

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

with A an arbitrary (nonlinear) maximal monotone operator, is, in full generality, not yet available, even if B does not depend on u and is a fixed non-random operator. However, a satisfactory treatment in the finite-dimensional case has been given by Pardoux and Răşcanu in [34, §4.2], where the authors consider stochastic differential equations in $\mathbb{R}^n$ of the type

$$dX_t + A(X_t)\,dt + F(t, X_t)\,dt \ni G(t, X_t)\,dB_t,$$

where A is a (multivalued) maximal monotone operator whose domain has non-emtpy interior, B is a k-dimensional Wiener process, G satisfies standard Lipschitz continuity assumptions, and $F(t,\cdot)$ is continuous and monotone (not necessarily Lipschitz continuous). While the assumptions on A are not restrictive in finite dimensions, unbounded linear operators generating contraction semigroups in infinite-dimensional spaces, as in our case, have dense domain, whose interior is hence empty.

[1] To avoid misunderstandings, we should clarify once and for all that with this expression we do *not* refer to a solution in the sense of rough paths, but simply "with ω fixed".

On the other hand, in the deterministic setting complete results have long been known for equations of the type

$$\frac{du}{dt} + Au \ni f, \qquad u(0) = u_0,$$

even in the much more general setting where A is a (multivalued) m-accretive operator on a Banach space E and $f \in L^1(0, T; E)$ (see, e.g., [5, 13]). Although a solution to the general stochastic problem does not currently seem within reach, significant results have been obtained in special cases: apart of the above-mentioned works on semilinear equations, well-posedness for the stochastic porous media equation under fairly general assumptions is known (see [7], where the same hypotheses on β imposed here are used and the noise is assumed to satisfy suitable boundedness conditions, and [8] for an extension to jump noise). Moreover, the variational theory by Pardoux, Krylov and Rozovskiĭ is essentially as complete as the corresponding deterministic theory. As mentioned above, however, large classes of maximal monotone operators on $H = L^2(D)$ cannot be cast in the variational framework.

The main contribution of this work is a well-posedness result for (1.1) under the most general conditions known so far, to the best of our knowledge. These conditions are quite sharp for A, but not for β. In particular, the conditions on A are close to those needed to show that $A + \beta(\cdot)$ is maximal monotone on $L^2(D)$, but the hypothesis that β is finite on the whole real line is not needed in the deterministic theory. Finally, the conditions on B are the natural ones to have function-valued noise, and are in this sense as general as possible. Equations with white noise in space and time, that have received much attention lately, are not within the scope of our approach (nor of others, most likely, under such general conditions on β).

In forthcoming work we shall extend our well-posedness results to equations where A is a nonlinear operator satisfying suitable Leray-Lions conditions (thus including the p-Laplacian, for instance), as well as to equations driven by discontinuous noise.

Let us now briefly outline the structure of the paper and the main ideas of the proof. Section 2 contains the statement of the main well-posedness result, and in Section 3 we discuss the hypotheses on the drift and diffusion coefficients, providing corresponding examples. After collecting useful preliminaries in Section 4, we consider in Section 5 a version of equation (1.1) with additive noise satisfying a strong boundedness assumption. Using the Yosida regularization of β, we obtain a family of approximating equations with Lipschitz coefficients, which can be treated by the standard variational theory. The solutions to such equations are shown to satisfy suitable uniform estimates, both pathwise and in expectation. Such estimates allow us to obtain key regularity and integrability properties for the solution to the equation with additive bounded noise. A crucial role is played by Simon's compactness criterion, which is applied pathwise, and by compactness criteria in L^1 spaces, applied both pathwise and in expectation. It is, in essence, precisely this interplay between pathwise and "averaged" arguments that permits to avoid many restrictive hypotheses of the existing literature. An abstract version of Jensen's inequality for positive operators, combined with the lower semicontinuity of convex integrals, is also an essential tool. In Section 6 we prove well-posedness for equations with additive noise removing the boundedness assumption of the previous section. This is accomplished by a further regularization scheme, this time on the diffusion operator B, and by a priori estimates for solutions to the regularized equations. A key role is played again by a combination of estimates and passages to the limit both pathwise and in expectation. We also prove continuity of the solution map with respect to the initial datum and the diffusion coefficient, by means of Itô's formula and regularizations, for which smoothing properties of the resolvent of A are essential. Finally, in Section 7 we obtain well-posedness in the general case by a fixed-point argument, using the Lipschitz continuity of B only. Introducing weighted spaces of stochastic processes, we obtain directly global well-posedness, thus avoiding a tedious construction by "patching" local solutions.

Some tools and reasonings used in this work are obviously not new: weak compactness arguments in L^1, for instance, are extensively used in the literature on partial differential equations (see, e.g., [10, 12] and references therein), as well as, to a lesser extent, in the stochastic setting (cf. [6, 7, 31]). However, even where similarities are present, our arguments are considerably streamlined and more general. The pathwise application of Simon's compactness criterion, made possible by a construction based on the variational framework, seems to be new, at least in the context of stochastic evolution equations. It is in fact somewhat surprising that the variational setting, which notoriously fails when dealing with semilinear equations, is at a basis of an approach that leads to well-posedness of those same equations, even with singular and rapidly increasing drift.

Acknowledgments. The authors are partially supported by a grant of The Royal Society. The first-named author is very grateful to Prof. S. Albeverio for the warm hospitality and the excellent working conditions at the Interdisziplinäres Zentrum für Komplexe Systeme, University of Bonn, where parts of this work were written. Two anonymous referees provided useful comments and suggestions that led to a better presentation of our results.

2 Main result

In this section, after fixing notation and conventions used throughout the paper, we state our main result.

2.1 Notation

All functional spaces will be defined on a smooth bounded domain $D \subset \mathbb{R}^n$. We shall denote $L^2(D)$ by H and its inner product by $\langle \cdot, \cdot \rangle$. The domain and the range of a generic map G will be denoted by $\mathsf{D}(G)$ and $\mathsf{R}(G)$, respectively. If E and F are subsets of a topological space, we shall write $E \hookrightarrow F$ to mean that E is continuously embedded in F, i.e. that E is a subset of F and that the injection $i : E \to F$ is continuous. Let E, F be Banach spaces. The space of linear continuous operators from E to F is denoted 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, that is, $T_n \to T$ in $\mathscr{L}_s(E, F)$ if $T_n u \to Tu$ in F for all $u \in E$. If $F = \mathbb{R}$, $\mathscr{L}(E, \mathbb{R})$ is the dual space E^*. If E and F are Hilbert spaces, we shall denote the space of Hilbert-Schmidt operators from E to F by $\mathscr{L}^2(E, F)$.

We shall occasionally use the symbols $\rightharpoonup$ and $\overset{*}{\rightharpoonup}$ to denote convergence in the weak and weak* topology of Banach spaces, respectively, while the symbol $\to$ is reserved for convergence in the norm topology.

All random quantities will be defined on a fixed probability space $(\Omega, \mathscr{F}, \mathbb{P})$ endowed with a right-continuous and saturated filtration $\mathbb{F} := (\mathscr{F}_t)_{t\in[0,T]}$, where T is a positive number. All expressions involving random quantities are meant to hold $\mathbb{P}$-almost surely, unless otherwise stated. With W we shall denote a cylindrical Wiener process on a separable Hilbert space U, that may coincide with H, but does not have to. We shall use the standard notation of stochastic calculus, such as $K \cdot W$ to mean the stochastic integral of K with respect to W, and, for a process X taking values in a normed space E, $X_t^* := \operatorname{ess\,sup}_{s\in[0,t]} \|X(s)\|_E$.

Let E be a separable Banach space. Given a measure space $(Y, \mathscr{A}, \mu)$ and $p \in [1, \infty]$, we shall denote the space of strongly measurable functions from $\phi : Y \to E$ such that $\|\phi\|_E \in L^p(Y)$ by $L^p(Y; E)$. Moreover, we shall write $L^2(\Omega; L^\infty(0, T; E))$ to denote the space of $\mathscr{F} \otimes \mathscr{B}([0, T])$-measurable processes $\phi : \Omega \times [0, T] \to E$ such that

$$\|\phi\|_{L^2(\Omega;L^\infty(0,T;E))} := \Big(\mathbb{E} \operatorname*{ess\,sup}_{t\in[0,T]} \|\phi(t)\|_E^2\Big)^{1/2} < \infty.$$

Given an interval $I \subseteq \mathbb{R}$, the space of continuous and of weakly continuous functions from I to E will be denoted by $C(I; E)$ and $C_w(I; E)$, respectively.

We shall write $a \lesssim b$ to mean that there exists a constant N such that $a \leq Nb$. If such a constant depends on certain parameters of interest, we shall put these in parentheses or write them as subscripts.

2.2 Assumptions

The following assumptions on the data of the problem are assumed to be in force throughout and will not always be recalled explicitly.

Assumption A. Let V be Hilbert space that is densely, continuously, and compactly embedded in H. The linear operator A belongs to $\mathscr{L}(V, V^*)$ and satisfies the following properties:

(i) there exists $C > 0$ such that

$$\langle Av, v\rangle \geq C\|v\|_V^2 \qquad \forall v \in V;$$

(ii) the part of A in H admits a unique m-accretive extension A_1 in $L^1(D)$;

(iii) the resolvent $\big((I + \lambda A_1)^{-1}\big)_{\lambda>0}$ is sub-Markovian;

(iv) there exists $m \in \mathbb{N}$ such that

$$\big\|(I + A_1)^{-m}\big\|_{\mathscr{L}(L^1(D),L^\infty(D))} < \infty.$$

Here we have used $\langle\cdot, \cdot\rangle$ also to denote the duality pairing of V and V^*, which is compatible with the scalar product in H. In fact, identifying H with its dual, one has the so-called Gel'fand triple

$$V \hookrightarrow H \hookrightarrow V^*,$$

where both embeddings are dense (see, e.g., [27, §2.9]). Moreover, we recall that the part of A in H is the operator A_2 on H defined as $\mathsf{D}(A_2) := \{x \in V : Au \in H\}$ and

$A_2x := Ax$ for all $x \in \mathsf{D}(A_2)$. If one identifies the operators with their graphs, this is equivalent to setting $A_2 := A \cap (V \times H)$. We shall often refer to condition (i) as the coercivity of A. The sub-Markovianity condition (iii) amounts to saying that, for all functions $f \in L^1(D)$ such that $0 \leq f \leq 1$, one has

$$0 \leq (I + A_1)^{-1} f \leq 1.$$

In other words, $(I + A_1)^{-1}$ is positivity preserving and contracting in $L^\infty(D)$.

From Section 5 onwards, we shall often use the symbol A to denote also A_1 and A_2.

Let us observe that if A is the negative Laplacian with Dirichlet boundary conditions, all hypotheses are met. Much wider classes of operators satisfying hypotheses (i)-(iv) will be given below.

Assumption B. β is a maximal monotone graph of $\mathbb{R} \times \mathbb{R}$ such that $\mathsf{D}(\beta) = \mathbb{R}$, $0 \in \beta(0)$, and its potential j is even.

We recall that the potential j of β is the convex, proper, lower semicontinuous function $j : \mathbb{R} \to \mathbb{R}_+$, with $j(0) = 0$, such that $\partial j = \beta$, where ∂ stands for the subdifferential in the sense of convex analysis.[2]

Assumption C. The diffusion coefficient

$$B : \Omega \times [0, T] \times H \to \mathscr{L}^2(U, H)$$

is Lipschitz continuous and grows linearly in its third argument, uniformly over $\Omega \times [0, T]$, i.e., there exist constants L_B, N_B such that

$$\begin{aligned} \big\| B(\omega, t, x) - B(\omega, t, y) \big\|_{\mathscr{L}^2(U,H)} &\leq L_B \|x - y\|_H, \\ \big\| B(\omega, t, x) \big\|_{\mathscr{L}^2(U,H)} &\leq N_B \big(1 + \|x\|_H\big) \end{aligned}$$

for all $\omega \in \Omega$, $t \in [0, T]$, and x, $y \in H$. Moreover, $B(\cdot, \cdot, x)$ is progressively measurable for all $x \in H$, i.e., for all $t \in [0, T]$, the map $(\omega, s) \mapsto B(\omega, s, x)$ from $\Omega \times [0, t]$, endowed with the σ-algebra $\mathscr{F}_t \otimes \mathscr{B}([0, t])$, to $\mathscr{L}^2(U, H)$, endowed with its Borel σ-algebra, is strongly measurable. We recall that, since U and H are separable, the space of Hilbert-Schmidt operators $\mathscr{L}^2(U, H)$ is itself a separable Hilbert space, hence strong and weak measurability coincide. Whenever we deal with maps with values in separable Banach spaces, since strong and weak measurability coincide, we shall drop the qualifier "strong".

2.3 The well-posedness result

Definition 2.1. *Let X_0 be an H-valued $\mathscr{F}_0$-measurable random variable. A* strong solution *to the stochastic equation* (1.1) *is a pair* (X, ξ) *satisfying the following properties:*

(i) *X is a measurable adapted V-valued process such that $AX \in L^1(0, T; V^*)$ and $B(\cdot, X) \in L^2(0, T; \mathscr{L}^2(U, H))$;*

(ii) *ξ is a measurable adapted $L^1(D)$-valued process such that $\xi \in L^1(0, T; L^1(D))$ and $\xi \in \beta(X)$ almost everywhere in $(0, T) \times D$;*

[2] See §4.1 below for a summary of the notions of convex analysis and of the theory of nonlinear monotone operators used throughout.

(iii) *one has, as an equality in* $L^1(D) \cap V^*$,

$$X(t) + \int_0^t AX(s)\,ds + \int_0^t \xi(s)\,ds = X_0 + \int_0^t B(s, X(s))\,dW(s)$$

for all $t \in [0, T]$.

Note that $L^1(D) \cap V^*$ is not empty because D has finite Lebesgue measure, hence, for instance, H is contained in both spaces.

Let us denote by $\mathscr{J}$ the set of pairs (ϕ, ζ), where ϕ and ζ are measurable adapted processes with values in H and $L^1(D)$, respectively, such that

$$\begin{aligned} \phi &\in L^2(\Omega; L^\infty(0,T;H)) \cap L^2(\Omega; L^2(0,T;V)), \\ \zeta &\in L^1(\Omega \times [0,T] \times D), \\ j(\phi) + j^*(\zeta) &\in L^1(\Omega \times [0,T] \times D). \end{aligned}$$

We shall say that (1.1) is well posed in $\mathscr{J}$ if there exists a unique process in $\mathscr{J}$ which is a strong solution and such that the solution map $X_0 \mapsto X$ is continuous from $L^2(\Omega; H)$ to $L^2(\Omega; L^\infty(0,T;H)) \cap L^2(\Omega; L^2(0,T;V))$.

The central result of this work is the following.

Theorem 2.2. *Let* $X_0 \in L^2(\Omega, \mathscr{F}_0, \mathbb{P}; H)$*. Then* (1.1) *is well-posed in* $\mathscr{J}$*. Moreover, the solution map* $X_0 \mapsto X$ *is Lipschitz continuous and the paths of* X *are weakly continuous with values in* H*.*

Let us stress the fact that the more general problem of unconditional well-posedness (i.e. without the extra condition that strong solutions belong to $\mathscr{J}$) remains open and is beyond the scope of the techniques used in this work. In particular, we can only prove uniqueness of solutions within $\mathscr{J}$.

3 Examples and remarks

Some comments and examples on the assumptions on the data of the problem are in order. In particular, the hypotheses on A deserve special attention. The coercivity condition $\langle Av, v \rangle \geq C\|v\|_V^2$ for all $v \in V$ is equivalent to $A \in \mathscr{L}(V, V^*)$ being determined by a bounded V-elliptic[3] bilinear form $\mathscr{E} : V \times V \to \mathbb{R}$, i.e. such that

$$|\mathscr{E}(u,v)| \lesssim \|u\|_V \|v\|_V, \qquad \mathscr{E}(v,v) \geq C\|v\|_V^2 \qquad \forall u, v \in V.$$

This is an immediate consequence of the Lax-Milgram theorem, which also implies that A is an isomorphism between V and V^* (see, e.g., [4, §5.2] or [32, Lemma 1.3]).

The bilinear form $\mathscr{E}$ can also be seen as a closed unbounded form on H with domain V. This defines a (unique) linear m-accretive operator A_2 on H, that is nothing else than

[3]We prefer this terminology, taken from [27], over the currently more common "V-coercive", to avoid possible confusion with related terminology used in the theory of Dirichlet forms, where coercivity is meant in a somewhat different sense (cf. [29, Definition 2.4, p. 16]).

the part of A in H (see, e.g., [4, §5.3] or [32, p. 34]). Conversely, given a positive closed bilinear form $\mathscr{E}$ on H with dense domain $\mathsf{D}(\mathscr{E})$ satisfying the *strong sector condition*[4]

$$|\mathscr{E}(u,v)| \lesssim \mathscr{E}(u,u)^{1/2}\mathscr{E}(v,v)^{1/2} \qquad \forall u,v \in \mathsf{D}(\mathscr{E}),$$

and such that $\mathscr{E}(u,u) > 0$ for all $u \in \mathsf{D}(\mathscr{E})$, $u \neq 0$, setting $V := \mathsf{D}(\mathscr{E})$ with inner product given by the symmetric part $\mathscr{E}^s$ of $\mathscr{E}$, that is

$$\mathscr{E}^s(u,v) := \frac{1}{2}\big(\mathscr{E}(u,v) + \mathscr{E}(v,u)\big), \qquad u,v \in \mathsf{D}(\mathscr{E}),$$

there is a unique linear operator $A \in \mathscr{L}(V,V^*)$ such that $\mathscr{E}(u,v) = \langle Au, v\rangle$ for all $u,v \in V$. This amounts to trivial verifications, since, obviously, $\mathscr{E}(u,u) = \mathscr{E}^s(u,u)$ for all $u \in \mathsf{D}(\mathscr{E})$. As a particular case, let A' be a linear positive self-adjoint (unbounded) operator H such that $\langle A'u, u\rangle > 0$ for all $u \in \mathsf{D}(A)$, $u \neq 0$. Then A' admits a square root $\sqrt{A'}$, which is in turn a linear positive self-adjoint operator on H. One can then define the Hilbert space $V := \mathsf{D}(\sqrt{A'})$, endowed with the inner product

$$\langle u,v\rangle_V := \big\langle \sqrt{A'}u, \sqrt{A'}v\big\rangle,$$

and the symmetric bounded bilinear form $\mathscr{E} : V \times V \to \mathbb{R}$,

$$\mathscr{E}(u,v) := \big\langle \sqrt{A'}u, \sqrt{A'}v\big\rangle, \qquad u,v \in V,$$

which is obviously V-elliptic. By a theorem of Kato ([23, Theorem 2.23, p. 331]), there is in fact a bijective correspondence between linear positive self-adjoint operators on H and positive densely-defined closed symmetric bilinear forms. More generally, if A' is a linear (unbounded) m-accretive operator on H such that

$$\big|\langle A'u, v\rangle\big| \lesssim \langle A'u, u\rangle^{1/2}\langle A'v, v\rangle^{1/2} \qquad \forall u,v \in \mathsf{D}(A'),$$

and $\langle A'u,u\rangle > 0$ for all $u \in \mathsf{D}(A')$, $u \neq 0$, then there exists a (unique) closed V-elliptic bilinear form $\mathscr{E}$ that determines an operator $A \in \mathscr{L}(V,V^*)$, with $V := \mathsf{D}(\mathscr{E})$ and $\langle\cdot,\cdot\rangle_V := \mathscr{E}^s$, such that A' is the part on H of A. This follows, for instance, by [29, p. 27].

Note, however, that in the previous examples V may *not* be continuously embedded in H, unless $\mathscr{E}$ satisfies a Poincaré inequality, i.e. $\|u\|_H^2 \lesssim \mathscr{E}(u,u)$ for all $u \in \mathsf{D}(\mathscr{E})$ (as is the case, for instance, for the Dirichlet Laplacian). This limitation is resolved by the following important observation: all our well-posedness result continues to hold if we assume, in place of hypothesis (i), the following weaker one:

(i') there exist constants $C_1 > 0$, $C_2 \in \mathbb{R}$ such that

$$\langle Av, v\rangle \geq C_1\|v\|_V^2 - C_2\|v\|_H^2 \qquad \forall v \in V,$$

which is clearly equivalent to assuming that $\tilde{A} := A + C_2 I$ is V-elliptic. Under this assumption, equation (1.1) can equivalently be written as

$$dX(t) + \tilde{A}X(t)\,dt + \beta(X(t))\,dt = C_2X(t)\,dt + B(t,X(t))\,dW(t).$$

[4] Throughout this section we shall follow the terminology on Dirichlet forms of [29].

The only added complication in the proofs to follow would be the appearance of functional spaces with an exponential weight in time, very much as in the proof of Proposition 6.2 below. An analogous argument, in a slightly different context, is developed in detail in [30]. This seemingly trivial observation allows to considerably extend the class of operators A that can be treated. For instance, one has the following criterion.

Lemma 3.1. *A coercive closed form $\mathscr{E}$ on H uniquely determines an operator A satisfying* (i').

Proof. The hypothesis of the Lemma means that $\mathscr{E}$ is a densely defined bilinear form such that its symmetric part $\mathscr{E}^s$ is closed and $\mathscr{E}$ satisfies the *weak sector condition*

$$\big|\mathscr{E}_1(u,v)\big| \lesssim \mathscr{E}_1(u,u)^{1/2}\mathscr{E}_1(v,v)^{1/2} \qquad \forall u,v \in \mathsf{D}(\mathscr{E}),$$

where $\mathscr{E}_1 := \mathscr{E} + I$. In other words, $\mathscr{E}$ satisfies the weak sector condition if the shifted form $\mathscr{E} + I$ satisfies the strong sector condition. Therefore, adapting in the obvious way an argument used above, it is enough to take $V := \mathsf{D}(\mathscr{E})$ with inner product $\langle\cdot,\cdot\rangle_V := \langle\cdot,\cdot\rangle_H + \mathscr{E}^s$ to obtain that the generator A_2 of $\mathscr{E}$ can be (uniquely) extended to an operator $A \in \mathscr{L}(V, V^*)$ satisfying (i') with $C_1 = C_2 = 1$. □

Note that in all the above constructions one has $V \hookrightarrow H$ densely and continuously (under appropriate assumptions), but the embedding is *not* necessarily compact. The latter condition has to be proved depending on the situation at hand. For a general compactness criterion in terms of ultracontractivity properties, see Proposition 3.3 below.

As regards condition (ii), the simplest sufficient condition ensuring that A_2 admits an m-accretive extension A_1 in $L^1(D)$ is that $-A_2$ is the generator of a symmetric Markovian semigroup of contractions S_2 on H, or, equivalently, that A_2 is positive self-adjoint with a Markovian resolvent. In fact, this implies that, for any $p \in [1,\infty[$, there exists a (unique) symmetric Markovian semigroup of contractions S_p on $L^p(D)$ such that all S_p, $1 \leq p < \infty$, are consistent, hence the corresponding negative generators A_p coincide on the intersections of their domains (see, e.g., [18, Theorem 1.4.1]). In the general case, i.e. if A_2 is not self-adjoint, the same conclusion remains true if the semigroup S_2 and its adjoint S_2^* are both sub-Markovian, or, equivalently, if S_2 is sub-Markovian and L^1-contracting (cf. [4, Lemma 10.13 and Theorem 10.15] or [32, Corollary 2.16]). In particular, if A_2 is the generator of a Dirichlet form on H, these conclusions hold. Moreover, since the resolvent of A_1 is sub-Markovian if and only if the resolvent of A_2 is sub-Markovian, we obtain the following complement to the previous Lemma.

Lemma 3.2. *A Dirichlet form $\mathscr{E}$ on H uniquely determines an operator A satisfying* (i')*,* (ii)*, and* (iii)*.*

Without assuming that S_2^* is sub-Markovian (which is the case, for instance, if A is determined by a semi-Dirichlet form on H, so that (i') and (iii) only are satisfied), we note that $D(A_2)$ is dense in $L^1(D)$, and the image of $I + A_2$ is dense in $L^1(D)$: the former assertion follows by $D(A_2) \subset L^2(D)$ densely and $L^2(D) \subset L^1(D)$ densely and continuously. Moreover, since A_2 generates a contraction semigroup in $L^2(D)$, the Lumer-Phillips theorem (see, e.g., [19, p. 83]) implies that $\mathsf{R}(I + A_2) = L^2(D)$, hence $\mathsf{R}(I + A_2)$ is dense in $L^1(D)$. The Lumer-Phillips theorem again guarantees that the closure of A_2 in $L^1(D)$ is m-accretive if A_2 is accretive in $L^1(D)$. The latter property is often not difficult to verify in concrete examples.

The most delicate condition is (iv), i.e. the ultracontractivity of suitable powers of the resolvent of A_1. If A_2 is self-adjoint, a simple duality arguments shows that, for any $t \geq 0$,

$$\|S_2(t)\|_{\mathscr{L}(L^1,L^\infty)} \leq \|S_2(t/2)\|^2_{\mathscr{L}(L^2,L^\infty)}.$$

Sufficient conditions for $S_2(t)$ to be bounded from $L^2(D)$ to $L^\infty(D)$ are known in terms, for instance, of logarithmic Sobolev inequalities, Sobolev inequalities, and Nash inequalities (see, e.g., [18, Chapter 2] and [32, Chapter 6]). The non-symmetric case is more difficult, but ultracontractivity estimates are known in many special cases, such as in the examples that we are going to discuss next. Ultracontractivity estimates for powers of the resolvent can then be obtained from estimates for the semigroup, as explained below. The following result (probably known, but for which we could not find a reference) shows that hypothesis (iv) guarantees that the embedding $\mathsf{D}(\mathscr{E}) \hookrightarrow H$ is compact, thus answering a question left open above.

Proposition 3.3. *Let A_2 be the generator of a closed coercive form $\mathscr{E}$ in H. If there exists $m \in \mathbb{N}$ such that the m-th power of the resolvent of A_2 is bounded from $L^2(D)$ to $L^\infty(D)$, then $\mathsf{D}(\mathscr{E})$ is compactly embedded in H.*

Proof. Let $(u_k)_k$ be a bounded sequence in $\mathsf{D}(\mathscr{E})$, i.e., there exists a constant N such that

$$\|u_k\|^2_H + \mathscr{E}^s(u_k,u_k) < N \qquad \forall k \in \mathbb{N}.$$

In particular, there exists a subsequence of k, denoted by the same symbol, such that u_k converges weakly to u in H as $k \to \infty$. The goal is to show that the convergence is in fact strong. Since $\mathsf{D}(A_2^m) \subset L^\infty(D)$ by assumption, it follows by a result of Arendt and Bukhvalov, see [3, Theorem 4.16(b)], that the resolvent $J_\lambda := (I+\lambda A_2)^{-1}$ is a compact operator on H for all $\lambda > 0$. The triangle inequality yields

$$\|u_k - u\| \leq \|u_k - J_\lambda u_k\| + \|J_\lambda u_k - J_\lambda u\| + \|J_\lambda u - u\|,$$

where the second term on the right-hand side converges to zero as $k \to \infty$ by compactness of J_λ. Moreover, since $J_\lambda \to I$ in $\mathscr{L}_s(H,H)$ as $\lambda \to 0$, the third term on the right-hand side can be made arbitrarily small. Therefore we only have to bound the first term on the right-hand side: note that $I - J_\lambda = \lambda A_\lambda$, where A_λ, $\lambda > 0$, stands for the Yosida approximation of A_2, hence $\|u_k - J_\lambda u_k\| = \lambda \|A_\lambda u_k\|$, and

$$\begin{aligned}\langle A_\lambda u_k, u_k\rangle &= \langle A_\lambda u_k, u_k - J_\lambda u_k + J_\lambda u_k\rangle = \lambda\|A_\lambda u_k\|^2 + \langle A_\lambda u_k, J_\lambda u_k\rangle \\ &\geq \lambda \|A_\lambda u_k\|^2,\end{aligned}$$

where we have used, in the last step, the identity $A_\lambda = A_2 J_\lambda$ and the monotonicity of A_2. Since, by [29, Lemma 2.11(iii), p. 20], one has

$$\bigl|\mathscr{E}_1^{(\lambda)}(u,v)\bigr| \lesssim \mathscr{E}_1(u,u)^{1/2}\mathscr{E}_1^{(\lambda)}(v,v)^{1/2} \qquad \forall u \in \mathsf{D}(\mathscr{E}),\ v \in H,$$

where $\mathscr{E}^{(\lambda)}(u,v) := \langle A_\lambda u, v\rangle$, $u, v \in H$, and the implicit constant depends only on $\mathscr{E}$, it follows that

$$\mathscr{E}_1^{(\lambda)}(u,u) \lesssim \mathscr{E}_1(u,u) \qquad \forall u \in \mathsf{D}(\mathscr{E}),$$

hence

$$\|u_k - J_\lambda u_k\|^2 = \lambda^2 \|A_\lambda u_k\|^2 \le \lambda \langle A_\lambda u_k, u_k\rangle = \lambda \mathscr{E}_1^{(\lambda)}(u_k, u_k) \lesssim \lambda \mathscr{E}_1(u_k, u_k).$$

By the assumptions on the sequence (u_k),

$$\mathscr{E}_1(u_k, u_k) = \|u_k\|^2 + \mathscr{E}(u_k, u_k) = \|u_k\|^2 + \mathscr{E}^s(u_k, u_k)$$

is bounded uniformely over k, hence $\|u_k - J_\lambda u_k\|^2$ can be made arbitrarily small as well, thus proving the claim. □

Let us now consider some concrete examples: we first consider the case of A being a suitable "realization" of a second-order differential operator, and then of a nonlocal operator.

Example 3.4 (Symmetric divergence-form operators). Consider the bilinear form $\mathscr{E}$ on $V := H_0^1(D)$ defined by

$$\mathscr{E}(u, v) := \big\langle a\nabla u, \nabla v\big\rangle = \sum_{j,k=1}^{n} a_{jk} \partial_j u \partial_k v,$$

where $a = (a_{jk})$ with $a_{jk} \in L^\infty(D)$ for all j, k, and $a_{jk} = a_{kj}$. The (formal) differential operator associated to $\mathscr{E}$ is

$$A_0 u := -\operatorname{div}\big(a\nabla u\big), \qquad u \in C_c^\infty(D),$$

where $C_c^\infty(D)$ stands for the set of infinitely differentiable functions with compact support contained in D. The form $\mathscr{E}$ is V-elliptic if there exists $C > 0$ such that $\langle a\xi, \xi\rangle \ge C|\xi|^2$ for all $\xi \in \mathbb{R}^n$. Moreover, if there exists a positive function $\mu \in C(D)$ such that $\langle a\xi, \xi\rangle \le \mu(\xi)|\xi|^2$ for all $\xi \in D$, then A_2 has sub-Markovian resolvent (details can be found, e.g., in [18, Chapter 1] and, in much more generality, in [29, Chapter II]). Ultracontractivity estimates follow as a special case of the corresponding estimates for non-symmetric forms treated next.

Example 3.5 (Non-symmetric divergence-form operators with lower-order terms). Consider the differential operator on smooth functions

$$\begin{aligned} A_0 u &:= -\operatorname{div}(a\nabla u) + b \cdot \nabla u - \operatorname{div}(cu) + a_0 u \\ &= -\sum_{j,k=1}^{n} \partial_j (a_{jk} \partial_k u) + \sum_{j=1}^{n} \big(b_j \partial_j u - \partial_j (c_j u)\big) + a_0 u, \end{aligned}$$

where a_{jk}, b_j, c_j, $a_0 \in L^\infty(D)$, and the associated (non-symmetric) bilinear form $\mathscr{E}$ on $V := H_0^1(D)$ is defined as

$$\begin{aligned} \mathscr{E}(u, v) &= \langle a\nabla u, \nabla v\rangle + \langle b \cdot \nabla u, v\rangle + \langle u, c \cdot \nabla v\rangle + \langle a_0 u, v\rangle \\ &= \int_D \Big(\sum_{jk} a_{jk} \partial_j u \, \partial_k v + \sum_j \big(b_j \partial_j u \, v + c_j u \partial_j v\big) + a_0 uv\Big). \end{aligned}$$

The bilinear form $\mathscr{E}$ is continuous, as it easily follows from the boundedness of its coefficients. If there exists a constant $C > 0$ such that $\langle a\xi, \xi\rangle \geq C|\xi|^2$, then $\mathscr{E}$ is not V-elliptic, but satisfies the weaker estimate

$$\mathscr{E}(u,u) \geq C_1\|u\|_V^2 - C_2\|u\|_H^2 \qquad \forall u \in V,$$

where $C_1 > 0$ and $C_2 \in \mathbb{R}$ (see, e.g., [4, §11.2] or [32, p. 100]), i.e. the corresponding operator A satisfies (i'), but not (i). Using the Poincaré inequality, it is not difficult to show that $\mathscr{E}$ is V-elliptic if the diameter of D is small enough (see [17, pp. 385–387]). If we furthermore assume that $a_0 - \operatorname{div} c \geq 0$ (in the sense of distributions), then the semigroup S_2 is sub-Markovian, and so is also the resolvent of A_2. Similarly, if $a_0 - \operatorname{div} b \geq 0$,[5] then the semigroup S_2 is L^1-contracting (these results can be found, for instance, in [4, Proposition 11.14], or deduced from [32, §4.3]). As already mentioned above, this implies that S_2 can be extended to a consistent family of semigroups S_p for all $p \in [1,\infty[$. Finally, let us discuss ultracontractivity: if $\mathscr{E}$ is V-elliptic, and S_2 as well as S_2^* are sub-Markovian, then a reasoning based on the Nash inequality

$$\big\|u\big\|_{L^2}^{2+4/n} \leq N\big\|u\big\|_{H_0^1}^2\big\|u\big\|_{L^1}^{4/n} \qquad \forall u \in H_0^1,$$

implies the estimate

$$\big\|S_2(t)\big\|_{\mathscr{L}(L^1,L^\infty)} \leq N_1 t^{-n/2},$$

where $N_1 := \big(Nn/(2\alpha)\big)^{n/2}$. For a proof, see, e.g., [2, Theorem 12.3.2] or [32, p. 159]. The Laplace transform representation of the resolvent yields

$$(I+\lambda A_1)^{-m} = \frac{\lambda^m}{(m-1)!}\int_0^\infty t^{m-1}e^{-\lambda t}S(t)\,dt$$

(see, e.g., [4, p. 17] or [35, p. 21]), hence

$$\big\|(I+\lambda A_1)^{-m}\big\|_{\mathscr{L}(L^1,L^\infty)} \lesssim \frac{\lambda^m}{(m-1)!}\int_0^\infty t^{m-1-n/2}e^{-\lambda t}\,dt.$$

Thus it suffices to choose m large enough to infer the ultracontractivity of the m-th power of the resolvent.

Example 3.6 (Fractional Laplacian). Let Δ be the Dirichlet Laplacian on H. Since it is a positive self-adjoint operator, it follows that, for any $\alpha \in \,]0,1[$, $(-\Delta)^\alpha$ is itself a positive self-adjoint (densely defined) operator on H. Furthermore, the bilinear form

$$\mathscr{E}(u,v) := \big\langle(-\Delta)^\alpha u, v\big\rangle = \big\langle(-\Delta)^{\alpha/2}u, (-\Delta)^{\alpha/2}v\big\rangle, \qquad u,v \in \mathsf{D}\big((-\Delta)^{\alpha/2}\big),$$

is a symmetric Dirichlet form on H, which, as already seen, uniquely determines an operator A satisfying conditions (i'), (ii), and (iii): in particular, $V = \mathsf{D}\big((-\Delta)^{\alpha/2}\big)$, equipped with the scalar product $\langle\cdot,\cdot\rangle_V := \langle\cdot,\cdot\rangle + \mathscr{E}$, and A is just the extension of $(-\Delta)^\alpha$, generator of $\mathscr{E}$, to V. In order to prove (iv), we are going to use again an

[5] These two conditions involving a_0 and the divergence of b, c, are not restrictive, as they are close to necessary to ensure that the bilinear form $\mathscr{E}$ is positive. This can be seen by a simple computation based on integration by parts, cf. [29, p. 48].

argument based on the Nash inequality, which is however more involved as before. In particular, since $-\Delta$ satisfies the Nash inequality

$$\|u\|_{L^2}^{2+4/n} \lesssim \langle -\Delta u, u\rangle \|u\|_{L^1}^{4/n} \qquad \forall u \in H_0^1,$$

a result by Bendikov and Maheux, see [9, Theorem 1.3], implies that the fractional power $(-\Delta)^\alpha$ satisfies the Nash inequality

$$\|u\|_{L^2}^{2+4\alpha/n} \lesssim \langle (-\Delta)^\alpha u, u\rangle \|u\|_{L^1}^{4\alpha/n} \qquad \forall u \in \mathsf{D}(\mathscr{E}).$$

It follows by a general criterion of Varopoulos, Saloff-Coste and Coulhon (attributed to Ph. Bénilan), see [39, Theorem II.5.2], that the semigroup S_α on H generated by $(-\Delta)^\alpha$ satisfies the ultracontractivity estimate

$$\|S_\alpha(t)\|_{\mathscr{L}(L^1,L^\infty)} \lesssim t^{-n/2\alpha},$$

from which corresponding estimates for suitable powers of the resolvent can be deduced, as in the previous example.

Related results on ultracontractivity and smoothing properties of semigroups generated by non-local operators, arising as generators of Markov processes, can be found, e.g., in [20, 26].

We proceed with a brief discussion about the relation between our hypotheses on A and those needed in the deterministic setting, where it is enough to prove that $A+\beta$ is maximal monotone in H to get well-posedness of the nonlinear equation, for any right-hand side belonging to $L^1(0,T;H)$. Probably the most widely used criterion for the maximal monotonicity of the sum of two maximal monotone operators on H, at least with applications to PDE in mind, is the following: let F be a maximal monotone operator on H and φ a lower semi-continuous proper convex function on H. If

$$\varphi\big((I+\lambda F)^{-1}u\big) \leq \varphi(u) + C\lambda \qquad \forall \lambda > 0,\ \forall u \in \mathsf{D}(\varphi), \tag{3.1}$$

then $F+\partial\varphi$ is maximal monotone (see [12, Theorem 9, p. 108]). In the case of semilinear perturbations of the Laplacian of the type $-\Delta+\beta$, this result is used as follows: let φ be such that $-\Delta = \partial\varphi$, and

$$\psi : u \mapsto \begin{cases} \displaystyle\int_D j(u)\,dx, & \text{if } j(u) \in L^1(D), \\ +\infty, & \text{if } j(u) \notin L^1(D). \end{cases}$$

Then $\psi : H \to \mathbb{R} \cup \{+\infty\}$ is proper convex lower semicontinuous, and $F := \partial\psi$ is maximal monotone, with $F(u) = \beta(u)$ a.e. for all $u \in H$ such that $j(u) \in L^1(D)$. Then one has, recalling that $(I+\lambda\beta)^{-1}$ is a contraction on $\mathbb{R}$,

$$\begin{aligned} \varphi\big((I+\lambda F)^{-1}u\big) &= \int_D \big|\nabla(I+\lambda\beta)^{-1}u\big|^2\,dx \\ &\leq \int_D |\nabla u|^2\,dx = \varphi(u), \end{aligned}$$

so that (3.1) is satisfied, and $-\Delta + \beta$ is maximal monotone. If one replaces $-\Delta$ with a general positive self-adjoint operator A on H, it is not clear how to adapt such reasoning. However, if we assume that A is the generator of a symmetric Dirichlet form $\mathscr{E}$ on H, then (3.1) is satisfied, with $C = 0$ and $\varphi = \mathscr{E}$. This follows from the fact that $(I+\lambda\beta)^{-1}$ is a normal contraction on $\mathbb{R}$ and that, for any normal contraction T on $\mathbb{R}$, $u \in \mathsf{D}(\mathscr{E})$ implies $Tu \in \mathsf{D}(\mathscr{E})$ and $\mathscr{E}(Tu,Tu) \leq \mathscr{E}(u,u)$, a proof of which can be found, e.g., in [29, Theorem 4.12, p. 36].

On the other hand, if A is maximal monotone but not self-adjoint, we cannot express it as the subdifferential of a convex function on H. Hence we are led to "dualize" the previous argument, i.e. we can try to show that

$$\psi\big((I+\lambda A)^{-1}u\big) \leq \psi(u) + C\lambda \qquad \forall \lambda > 0,\ \forall u \in \mathsf{D}(\varphi).$$

Knowing only that the resolvent is a contraction does not seem enough to proceed. However, if we assume that the resolvent is sub-Markovian, we can apply Jensen's inequality (see Lemma 4.2 below), so that

$$j\big((I+\lambda A)^{-1}u\big) \leq (I+\lambda A)^{-1}j(u),$$

hence, integrating,

$$\psi\big((I+\lambda A)^{-1}u\big) = \int_D j\big((I+\lambda A)^{-1}u\big)\,dx \leq \int_D (I+\lambda A)^{-1}j(u)\,dx.$$

Assuming also that the resolvent is contracting in L^1, we obtain $\psi\big((I+\lambda A)^{-1}u\big) \leq \psi(u)$, hence that $A+\beta$ is maximal monotone in H. Recall that A is contracting in L^1 if it is the generator of a (nonsymmetric) Dirichlet form. It results from this discussion that our conditions (ii) and (iii) on A are not restrictive and are probably close to optimal, while the ultracontractivity condition (iv) is completely superfluous in the deterministic setting. Moreover, while condition (i') is always satisfied if A is self-adjoint, it is equally superfluous in the deterministic case if A is non-symmetric.

Let us now comment on the Lipschitz continuity assumption on B. It is natural to ask whether a well-posedness result analogous to Theorem 2.2 holds under the weaker assumption that B is progressively measurable, linearly growing, and just locally Lipschitz continuous, i.e. assuming that there exists a sequence $(L_B^n)_n$ of positive real numbers such that

$$\big\|B(\omega,t,x) - B(\omega,t,y)\big\|_{\mathscr{L}^2(U,H)} \leq L_B^n \|x-y\|_H$$

for every $(\omega,t) \in \Omega \times [0,T]$ and $x,y \in H$ with $\|x\|_H, \|y\|_H \leq n$, for every $n \in \mathbb{N}$. In this case, introducing the globally Lipschitz continuous truncated operators

$$B_n : \Omega \times [0,T] \times H \to \mathscr{L}^2(U,H), \qquad B_n(\omega,t,x) := B(\omega,t,nPx),$$

for all $n \in \mathbb{N}$, where $P : H \to H$ is the projection on the closed unit ball in H, the stochastic evolution equation

$$dX_n + AX_n\,dt + \beta(X_n)\,dt \ni B_n(t,X_n)\,dW, \qquad X_n(0) = X_0,$$

is well-posed in $\mathscr{J}$ for all $n \in \mathbb{N}$. One would now expect to be able to construct a global solution by suitably "gluing" the solutions (X_n, ξ_n). In fact, this technique has been successfully applied in several situations (cf., e.g., [14, 25, 38]): the key argument is to introduce the sequence of stopping times $(\tau_n)_n$ defined as

$$\tau_n := \inf\bigl\{t \in [0,T] : \|X_n(t)\| \geq n\bigr\} \wedge T,$$

and to show that, for any $m > n$, one has $X_m = X_n$ on

$$[\![0, \tau_n]\!] := \bigl\{(\omega, t) \in \Omega \times [0,T] : \, 0 \leq t \leq \tau_n(\omega)\bigr\}.$$

For this construction to work, it seems essential to assume that X_n has continuous trajectories for all $n \in \mathbb{N}$ (as is the case in *op. cit.*). However, in our case, we only know that the trajectories of X_n are weakly continuous in H, hence the above construction does not seem to work. On the other hand, we conjecture that strong solutions in $\mathscr{J}$ to (1.1) are indeed pathwise continuous under suitable polynomial boundedness assumption on β, and that, in this case, equations with locally Lipschitz diffusion coefficient can be shown to be well-posed. This will be treated in forthcoming work. We conclude remarking that such a well-posedness result for semilinear equations with polynomially growing drift does not follow from the classical variational approach (see, e.g., [28, Example 5.1.8]).

4 Preliminaries

We collect, for the reader's convenience, several notions and results that we are going to use in the following sections.

4.1 Convex analysis and monotone operators

We recall basic concepts of convex analysis and their connections with the theory of maximal monotone operators. We limit ourselves to the case of functions (and operators) defined on the real line, as we will not need the general setting of Banach spaces. For a comprehensive treatment we refer, e.g., to [5, 13, 22].

A graph γ in $\mathbb{R} \times \mathbb{R}$ is called monotone if

$$(x_1 - x_2)(y_1 - y_2) \geq 0$$

for all (x_1, y_1), $(x_2, y_2) \in \gamma$. If γ is maximal in the family of monotone subsets of $\mathbb{R} \times \mathbb{R}$, endowed with the partial order relation of set inclusion, then it is said to be maximal monotone. In other words, γ is maximal monotone if it does not admit any proper monotone extension. This maximality property is equivalent to the range condition

$$\mathsf{R}(I + \lambda\gamma) = \mathbb{R} \qquad \forall \lambda > 0,$$

where I stands for the identity function. Monotonicity implies that the inverse $(I + \lambda\gamma)^{-1}$, called the resolvent of γ, is single-valued (hence a function, not just a graph) and contracting. Moreover, $(I + \lambda\gamma)^{-1}$ converges pointwise to the projection on the closed convex set $\overline{\mathsf{D}(\gamma)}$ as $\lambda \to 0$. An essential tool is the Yosida regularization $\gamma_\lambda : \mathbb{R} \to \mathbb{R}$, defined as

$$\gamma_\lambda := \frac{1}{\lambda}\bigl(I - (I + \lambda\gamma)^{-1}\bigr), \qquad \lambda > 0.$$

The following properties will be used extensively:

(a) γ_λ is monotone and Lipschitz continuous, with Lipschitz constant bounded by $1/\lambda$;

(b) $\gamma_\lambda \in \gamma \circ (I + \lambda\gamma)^{-1}$.

Let $\varphi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be a function not identically equal to $+\infty$ (i.e., proper), convex and lower-semicontinuous. Denoting the set of subsets of $\mathbb{R}$ by $\mathfrak{P}(\mathbb{R})$, the map

$$\begin{aligned} \partial\varphi : \mathbb{R} &\longrightarrow \mathfrak{P}(\mathbb{R}) \\ x &\longmapsto \bigl\{ z \in \mathbb{R} : \varphi(y) - \varphi(x) \geq z(y-x) \quad \forall y \in \mathbb{R} \bigr\} \end{aligned}$$

is called the subdifferential of φ. The multivalued map $\gamma := \partial\varphi$, that can equivalently be considered as a graph in $\mathbb{R} \times \mathbb{R}$, is maximal monotone. Conversely, every maximal monotone graph of $\mathbb{R} \times \mathbb{R}$ is the subdifferential of a convex proper function, which is, roughly speaking, its indefinite integral.

The Moreau-Yosida regularization of φ is the convex differentiable function $\varphi_\lambda : \mathbb{R} \to \mathbb{R}$ defined by

$$\varphi_\lambda(x) := \inf_{y \in \mathbb{R}} \Bigl(\varphi(y) + \frac{|x-y|^2}{2\lambda} \Bigr), \qquad \lambda > 0.$$

It enjoys the following fundamental properties:

(c) $\varphi'_\lambda = \gamma_\lambda$, where γ_λ denotes the Yosida regularization of $\gamma = \partial\varphi$;

(d) φ_λ converges pointwise to φ from below as $\lambda \to 0$;

The (Fenchel-Legendre) conjugate of φ is the proper convex lower-semicontinuous function $\varphi^* : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ defined as

$$\varphi^* : x \mapsto \sup_{y \in \mathbb{R}} \bigl(xy - \varphi(y) \bigr).$$

The Young inequality

$$xy \leq \varphi(y) + \varphi^*(x) \qquad \forall x, y \in \mathbb{R}$$

follows immediately from the definition. The following properties will be particularly useful:

(e) equality holds in the Young inequality if and only if $x \in \partial\varphi(y)$;

(f) if $\mathsf{D}(\gamma) = \mathbb{R}$, then φ^* is superlinear at infinity, i.e.

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

We shall also need a result about passing to the limit "within" maximal monotone graphs due to Brézis, see [12, Theorem 18, p. 126].

Lemma 4.1. *Let γ be a maximal monotone graph in $\mathbb{R} \times \mathbb{R}$ with $\mathsf{D}(\gamma) = \mathbb{R}$ and $0 \in \gamma(0)$. Assume that the sequences $(y_n)_{n\in\mathbb{N}}$, $(g_n)_{n\in\mathbb{N}}$ of real-valued measurable functions on a finite measure space $(Y, \mathscr{A}, \mu)$ are such that $y_n \to y$ μ-a.e. as $n \to \infty$, $g_n \in \gamma(y_n)$ μ-a.e. for all $n \in \mathbb{N}$, and $(g_n y_n)$ is a bounded subset of $L^1(Y, \mathscr{A}, \mu)$. Then there exists $g \in L^1(Y, \mathscr{A}, \mu)$ and a subsequence n' such that $g_{n'} \to g$ weakly in $L^1(Y, \mathscr{A}, \mu)$ as $n' \to \infty$ and $g \in \gamma(y)$ μ-almost everywhere.*

Finally, we recall a simplified version of an "abstract" Jensen's inequality, due to Haase (see [21, Theorem 3.4]), that will be used to prove a priori estimates for convex functionals of stochastic processes.

Lemma 4.2. *Let $(Y, \mathscr{A}, \mu)$, $(Z, \mathscr{B}, \nu)$ be measure spaces, $E \subset L^0(Y, \mathscr{A}, \mu)$ a Banach function space, and*
$$T : E \longrightarrow L^0(Z, \mathscr{B}, \nu)$$
a linear continuous sub-Markovian operator. Moreover, let $\varphi : \mathbb{R} \to [0, \infty[$ be a convex lower semicontinuous function with $\varphi(0) = 0$. Then
$$\varphi(Tf) \leq T\varphi(f)$$
for all $f \in E$ such that $\varphi(f) \in E$.

4.2 Hilbert-Schmidt operators

Let us recall now some standard facts about linear maps. We recall that the space of continuous linear operators from a Banach space E to another one F, equipped with the strong operator topology, is denoted by $\mathscr{L}_s(E, F)$. If E and F are Hilbert spaces, the space of Hilbert-Schmidt operators $\mathscr{L}^2(E, F)$ is an operator ideal, in particular it is stable with respect to pre-composition as well as post-composition with continuous linear operators: if E' and F' are also Hilbert spaces, and
$$E' \xrightarrow{R} E \xrightarrow{T} F \xrightarrow{L} F',$$
with R and L continuous linear operators, then $LTR \in \mathscr{L}^2(E', F')$,[6] with
$$\|LTR\|_{\mathscr{L}^2(E',F')} \leq \|L\|_{\mathscr{L}(F,F')} \|T\|_{\mathscr{L}^2(E,F)} \|R\|_{\mathscr{L}(E',E)}$$
(see, e.g., [11, p. V.52]). It follows from these properties that, for any $T \in \mathscr{L}^2(E, F)$, the mapping
$$\begin{aligned} \Phi_T : \mathscr{L}_s(F, F') &\longrightarrow \mathscr{L}^2(E, F') \\ L &\longmapsto LT \end{aligned}$$
is continuous: $L_n \to L$ in $\mathscr{L}_s(F, F')$ implies that $L_n T \to LT$ in $\mathscr{L}^2(E, F')$. If E and F are separable, then $\mathscr{L}^2(E, F)$ is itself a separable Hilbert space.

Lemma 4.3. *If G is a progressively measurable $\mathscr{L}^2(U, H)$-valued process such that*
$$\mathbb{E} \int_0^T \|G(s)\|^2_{\mathscr{L}^2(U,H)} \, ds < \infty$$
and F is a progressively measurable H-valued process such that $\mathbb{E}(F^*_T)^2 < \infty$, then, for any $\varepsilon > 0$,
$$\mathbb{E}\big((FG) \cdot W\big)^*_T \leq \varepsilon \, \mathbb{E}\big(F^*_T\big)^2 + N(\varepsilon) \, \mathbb{E} \int_0^T \|G(s)\|^2_{\mathscr{L}^2(U,H)} \, ds.$$

[6] One may say, in a shorter but perhaps cryptic way, that $\mathscr{L}^2$ is functorial, more precisely that $\mathscr{L}^2(E, \cdot)$ and $\mathscr{L}^2(\cdot, F)$ are a covariant and a contravariant functor, respectively.

Proof. By the ideal property of Hilbert-Schmidt operators, one has

$$\begin{aligned}\bigl\|F(s)G(s)\bigr\|_{\mathscr{L}^2(U,\mathbb{R})} &\leq \bigl\|F(s)\bigr\|_H \bigl\|G(s)\bigr\|_{\mathscr{L}^2(U,H)}\\ &\leq (F_T^*)\bigl\|G(s)\bigr\|_{\mathscr{L}^2(U,H)}\end{aligned}$$

for all $s \in [0,T]$, hence

$$\int_0^T \bigl\|F(s)G(s)\bigr\|^2_{\mathscr{L}^2(U,\mathbb{R})}\,ds \leq (F_T^*)^2 \int_0^T \bigl\|G(s)\bigr\|^2_{\mathscr{L}^2(U,H)}\,ds,$$

where the right-hand side is finite $\mathbb{P}$-a.s. thanks to the assumptions on F and G. Then $(FG) \cdot W$ is a local martingale, for which Davis' inequality yields

$$\begin{aligned}\mathbb{E}\bigl((FG)\cdot W\bigr)^*_T &\lesssim \mathbb{E}\bigl[(FG)\cdot W, (FG)\cdot W\bigr]^{1/2}_T\\ &= \mathbb{E}\biggl(\int_0^T \bigl\|F(s)G(s)\bigr\|^2_{\mathscr{L}^2(U,\mathbb{R})}\,ds\biggr)^{1/2}\\ &\leq \mathbb{E}(F_T^*)\biggl(\int_0^T \bigl\|G(s)\bigr\|^2_{\mathscr{L}^2(U,H)}\,ds\biggr)^{1/2}.\end{aligned}$$

The proof is finished invoking the elementary inequality

$$ab \leq \frac{1}{2}\bigl(\varepsilon a^2 + \frac{1}{\varepsilon}b^2\bigr) \qquad \forall a,b \in \mathbb{R}.$$ □

4.3 Continuity and compactness in spaces of vector-valued functions

The following result by Strauss, see [37, Theorem 2.1], provides sufficient conditions for a vector-valued function to be weakly continuous. It will be used to establish the pathwise weak continuity of solutions to several stochastic equations. We recall that, given a Banach space E and an interval $I \subseteq \mathbb{R}$, the space of weakly continuous functions from I to E is denoted by $C_w(I;E)$.

Lemma 4.4. *Let E and F be Banach spaces such that E is dense in F, $E \hookrightarrow F$, and E is reflexive. Then*

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

The next result is a classical integration-by-parts formula, whose proof can be found, for instance, in [5, §1.3]. Let $\mathcal{V}$ and $\mathcal{H}$ be Hilbert spaces such that $\mathcal{V} \hookrightarrow \mathcal{H} \hookrightarrow \mathcal{V}^*$, and denote by $W(a,b;\mathcal{V})$ the set of functions $u \in L^2(a,b;\mathcal{V})$ such that $u' \in L^2(a,b;\mathcal{V}^*)$, where the derivative u' is meant in the sense of $\mathcal{V}^*$-valued distributions. The duality of $\mathcal{V}$ and $\mathcal{V}^*$ as well as the scalar product of $\mathcal{H}$ will be denoted by $\langle\cdot,\cdot\rangle$.

Lemma 4.5. *Let $u \in W(a,b;\mathcal{V})$. Then there exists $\tilde{u} \in C([a,b];\mathcal{H})$ such that $u(t) = \tilde{u}(t)$ for almost all $t \in [a,b]$. Moreover, for any $v \in W(a,b;\mathcal{V})$, $\langle u,v\rangle$ is absolutely continuous on $[a,b]$ and*

$$\frac{d}{dt}\bigl\langle u(t),v(t)\bigr\rangle = \bigl\langle u'(t),v(t)\bigr\rangle + \bigl\langle u(t),v'(t)\bigr\rangle.$$

The following compactness criterion is due to Simon, see [36, Corollary 4, p. 85].

Lemma 4.6. *Let E_1, E_2, E_3 be Banach spaces such that $E_1 \hookrightarrow E_2$ and $E_2 \hookrightarrow E_3$ compactly. Assume that F is a bounded subset of $L^p(0,T;E_1) \cap W^{1,1}(0,T;E_3)$ for some $p \geq 1$. Then F is relatively compact in $L^p(0,T;E_2)$.*

5 Well-posedness for a regularized equation

Let V_0 be a separable Hilbert space such that V_0 is a dense subset of V, $V_0 \hookrightarrow V$, and $V_0 \hookrightarrow L^\infty(D)$. The goal of this section is to establish existence and uniqueness of solutions to the stochastic evolution equation

$$dX(t) + AX(t)\,dt + \beta(X(t))\,dt \ni B(t)\,dW(t), \qquad X(0) = X_0, \tag{5.1}$$

where B is an $\mathscr{L}^2(U, V_0)$-valued process. In particular, this stochastic equation can be interpreted as a version of (1.1) with additive and more regular noise.

Proposition 5.1. *Assume that $X_0 \in L^2(\Omega, \mathscr{F}_0, \mathbb{P}; H)$ and that*

$$B \in L^2(\Omega; L^2(0,T; \mathscr{L}^2(U, V_0)))$$

is measurable and adapted. Then equation (5.1) admits a unique strong solution (X, ξ) such that

$$\begin{gathered} X \in L^2(\Omega; L^\infty(0,T;H)) \cap L^2(\Omega; L^2(0,T;V)), \\ j(X) + j^*(\xi) \in L^1((0,T) \times D) \qquad \mathbb{P}\text{-almost surely.} \end{gathered}$$

Moreover, $X(\omega, \cdot) \in C_w([0,T]; H)$ for $\mathbb{P}$-almost all $\omega \in \Omega$.

The rest of this section is devoted to the proof of Proposition 5.1, which is structured as a follows: we consider a regularized version of (5.1), where the nonlinear term β is replaced by its Yosida approximation, and obtain suitable a priori estimates, both pathwise and in expectation. Taking limits in appropriate topologies of the solutions to these regularized equations, we construct solutions to (5.1), that are finally shown to be unique.

Let

$$\beta_\lambda := \frac{1}{\lambda}\bigl(I - (I + \lambda\beta)^{-1}\bigr), \qquad \lambda > 0,$$

be the Yosida approximation of β, and consider the regularized equation

$$dX_\lambda(t) + AX_\lambda(t)\,dt + \beta_\lambda(X_\lambda(t))\,dt = B(t)\,dW(t), \qquad X_\lambda(0) = X_0.$$

Since β_λ is monotone and Lipschitz continuous, it is easy to check that the operator $A + \beta_\lambda$ satisfies, for any $\lambda > 0$, the classical conditions of Pardoux, Krylov and Rozovskiĭ [24, 33]. For completeness, a proof is given next.

Lemma 5.2. *Let $\lambda > 0$. The operator $A_\lambda := A + \beta_\lambda : V \to V^*$ satisfies the following conditions:*

(i) *A_λ is hemicontinuous, i.e. the map $\mathbb{R} \ni \eta \mapsto \langle A_\lambda(u + \eta v), x\rangle$ is continuous for all u, v, $x \in V$;*

(ii) *A_λ is monotone, i.e. $\langle A_\lambda u - A_\lambda v, u - v\rangle \geq 0$ for all u, $v \in V$;*

(iii) *A_λ is coercive, i.e. there exists a constant $C_1 > 0$ such that $\langle A_\lambda v, v\rangle \geq C_1 \|v\|_V^2$ for all $v \in V$;*

(iv) A_λ *is bounded, i.e. there exists a constant* $C_2 > 0$ *such that* $\|A_\lambda v\|_{V^*} \leq C_2\|v\|_V$ *for all* $v \in V$.

Proof. (i) For any $u, v, x \in V$, one has

$$\langle A_\lambda(u+\eta v), x\rangle = \langle Au, x\rangle + \eta\langle Av, x\rangle + \int_D \beta_\lambda(u+\eta v)x.$$

It clearly suffices to check that the last term depends continuously on η, which follows immediately by the Lipschitz continuity of β_λ. (ii) Since both A and β_λ are monotone, one has

$$\langle A_\lambda u - A_\lambda v, u-v\rangle = \langle Au - Av, u-v\rangle + \int_D (\beta_\lambda(u) - \beta_\lambda(v)(u-v) \geq 0.$$

(iii) Similarly, since $0 \in \beta(0)$ implies $\beta_\lambda(0) = 0$, coercivity of A and monotonicity of β_λ imply

$$\langle A_\lambda v, v\rangle = \langle Av, v\rangle + \int_D \beta_\lambda(v)v \geq \langle Av, v\rangle \geq C\|v\|_V^2$$

(in particular, C_1 can be chosen equal to C, the coercivity constant of A itself). (iv) Using again the fact that $\beta_\lambda(0) = 0$, and recalling that β_λ is Lipschitz continuous with Lipschitz constant bounded by $1/\lambda$, one has

$$\begin{aligned}\langle A_\lambda v, u\rangle = \langle Av, u\rangle + \int_D \beta_\lambda(v)u &\leq \|Av\|_{V^*}\|u\|_V + \frac{1}{\lambda}\|v\|_H\|u\|_H \\ &\leq \big(\|A\|_{\mathscr{L}(V,V^*)} + k/\lambda\big)\|v\|_V\|u\|_V,\end{aligned}$$

where k is the norm of the continuous embedding $\iota : V \to H$. □

Hence (5.2) admits a unique variational solution, that is, there exists a unique adapted process

$$X_\lambda \in L^2(\Omega; C([0,T]; H)) \cap L^2(\Omega; L^2(0,T;V))$$

such that, in V^*,

$$X_\lambda(t) + \int_0^t AX_\lambda(s)\,ds + \int_0^t \beta_\lambda(X_\lambda(s))\,ds = X_0 + \int_0^t B(s)\,dW(s) \tag{5.2}$$

for all $t \in [0,T]$.

In the next lemmata we establish a priori estimates for X_λ and $\beta_\lambda(X_\lambda)$. We begin with a pathwise estimate.

Lemma 5.3. *There exists* $\Omega' \subseteq \Omega$ *with* $\mathbb{P}(\Omega') = 1$ *and* $M : \Omega' \to \mathbb{R}$ *such that*

$$\big\|X_\lambda(\omega)\big\|^2_{C([0,T];H)\cap L^2(0,T;V)} + \big\|j_\lambda(X_\lambda(\omega))\big\|_{L^1(0,T;L^1(D))} < M(\omega)$$

for all $\omega \in \Omega'$.

Proof. Setting $Y_\lambda := X_\lambda - B\cdot W$, Itô's formula[7] yields

$$\big\|Y_\lambda(t)\big\|_H^2 + 2\int_0^t \big\langle AX_\lambda(s), Y_\lambda(s)\big\rangle\,ds + 2\int_0^t \big\langle \beta_\lambda(X_\lambda), Y_\lambda(s)\big\rangle\,ds = \big\|X_0\big\|_H^2,$$

where $\|X_\lambda\|_H \le \|Y_\lambda\|_H + \|B\cdot W\|_H$ by the triangle inequality, hence

$$\|Y_\lambda(t)\|_H^2 \ge \frac{1}{2}\|X_\lambda(t)\|_H^2 - \|B\cdot W(t)\|_H^2.$$

Moreover, writing $\langle AX_\lambda, Y_\lambda\rangle = \langle AX_\lambda, X_\lambda\rangle - \langle AX_\lambda, B\cdot W\rangle$, one has

$$\langle AX_\lambda, X_\lambda\rangle \ge C\|X_\lambda\|_V^2$$

by the coercivity of A, and

$$\begin{aligned}\langle AX_\lambda, B\cdot W\rangle &\le \|A\|_{\mathscr{L}(V,V^*)}\|X_\lambda\|_V\|B\cdot W\|_V\\ &\le \frac{1}{2}C\|X_\lambda\|_V^2 + \frac{1}{2\varepsilon}\|B\cdot W\|_V^2,\end{aligned}$$

where we have used the elementary inequality $ab \le \frac{1}{2}(\varepsilon a^2 + b^2/\varepsilon)$ for all $a,b\in\mathbb{R}$, with $\varepsilon := C\|A\|_{\mathscr{L}(V,V^*)}^{-2}$. Then

$$\langle AX_\lambda, Y_\lambda\rangle \ge \frac{1}{2}C\|X_\lambda\|_V^2 - \frac{1}{2\varepsilon}\|B\cdot W\|_V^2,$$

so that

$$2\int_0^t \big\langle AX_\lambda(s), Y_\lambda(s)\big\rangle\,ds \ge C\int_0^t \|X_\lambda(s)\|_V^2\,ds - \frac{1}{\varepsilon}\int_0^t \|B\cdot W(s)\|_V^2\,ds$$

and

$$\begin{aligned}&\frac{1}{2}\|X_\lambda(t)\|_H^2 + C\int_0^t \|X_\lambda(s)\|_V^2\,ds + 2\int_0^t \big\langle \beta_\lambda(X_\lambda(s)), Y_\lambda(s)\big\rangle\,ds\\ &\qquad\le \|X_0\|_H^2 + \|B\cdot W(t)\|_H^2 + \frac{1}{\varepsilon}\int_0^t \|B\cdot W(s)\|_V^2\,ds.\end{aligned} \tag{5.3}$$

Let j_λ be the Moreau-Yosida regularization of j, that is

$$j_\lambda(x) := \inf_{y\in\mathbb{R}}\Big(j(y) + \frac{|x-y|^2}{2\lambda}\Big), \qquad \lambda > 0.$$

We recall that j_λ is a convex, proper differentiable function, with $j'_\lambda = \beta_\lambda$, that converges pointwise to j from below. In particular,

$$\beta_\lambda(x)(x-y) \ge j_\lambda(x) - j_\lambda(y) \ge j_\lambda(x) - j(y) \qquad \forall x,y\in\mathbb{R}.$$

This implies

$$\begin{aligned}\int_0^t \big\langle \beta_\lambda(X_\lambda(s)), Y_\lambda(s)\big\rangle\,ds &= \int_0^t\!\!\int_D \beta_\lambda(X_\lambda(s,x))(X_\lambda(s,x) - B\cdot W(s,x))\,dx\,ds\\ &\ge \int_0^t\!\!\int_D j_\lambda(X_\lambda(s,x))\,dx\,ds - \int_0^t\!\!\int_D j(B\cdot W(s,x))\,dx\,ds,\end{aligned}$$

[7] Whenever we refer to Itô's formula, we shall always mean the version in [24].

hence also

$$\begin{aligned}\frac{1}{2}\|X_\lambda(t)\|_H^2 + C\int_0^t \|X_\lambda(s)\|_V^2\,ds + 2\int_0^t\int_D j_\lambda(X_\lambda(s,x))\,dx\,ds \\ \leq \|X_0\|_H^2 + \|B\cdot W(t)\|_H^2 + \frac{1}{\varepsilon}\int_0^t \|B\cdot W(s)\|_V^2\,ds \\ + 2\int_0^t\int_D j(B\cdot W(s,x))\,dx\,ds.\end{aligned}$$

Taking the supremum with respect to t yields

$$\begin{aligned}&\big\|X_\lambda\big\|^2_{C([0,T];H)} + \big\|X_\lambda\big\|^2_{L^2(0,T;V)} + \big\|j_\lambda(X_\lambda)\big\|_{L^1(0,T;L^1(D))} \\ &\qquad \lesssim \big\|X_0\big\|_H^2 + \big\|B\cdot W\big\|^2_{C([0,T];H)} + \big\|B\cdot W\big\|^2_{L^2(0,T;V)} + \big\|j(B\cdot W)\big\|_{L^1(0,T;L^1(D))},\end{aligned}$$

where the implicit constant depends only on the operator norm of A. It follows by Itô's isometry and Doob's inequality that

$$\big\|B\cdot W\big\|_{L^2(\Omega;C([0,T];V_0))} \lesssim \big\|B\big\|_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,V_0)))},$$

where the right-hand side is finite by assumption, hence, recalling that V_0 is continuously embedded in V,

$$\big\|B\cdot W\big\|_{C([0,T];H)} + \big\|B\cdot W\big\|_{L^2(0,T;V)} \lesssim_T \big\|B\cdot W\big\|_{C([0,T];V_0)}.$$

Analogously, denoting the norm of the continuous embedding $\iota : V_0 \to L^\infty(D)$ by k, one has, recalling that j is symmetric and increasing on $\mathbb{R}_+$,

$$\big\|j(B\cdot W(t)\big\|_{L^1(D)} \lesssim_{|D|} j\big(\|B\cdot W(t)\|_{L^\infty(D)}\big) \leq j\big(k\|B\cdot W(t)\|_{V_0}\big),$$

for all $t \in [0,T]$, hence

$$\big\|j(B\cdot W)\big\|_{L^1(0,T;L^1(D))} \lesssim_{|D|,T} j\big(k\|B\cdot W\|_{C([0,T];V_0)}\big).$$

The proof is complete choosing $\Omega' \subset \Omega$ such that $\|X_0(\omega)\|_H$ and $\|B\cdot W(\omega)\|_{C([0,T];V_0)}$ are finite for all $\omega \in \Omega'$, and defining $M : \Omega' \to \mathbb{R}$ as

$$M := \big\|X_0\big\|_H^2 + \big\|B\cdot W\big\|^2_{C([0,T];H)} + \big\|B\cdot W\big\|^2_{L^2(0,T;V)} + \big\|j(B\cdot W)\big\|_{L^1(0,T;L^1(D))}. \qquad \square$$

Remark 5.4. The above estimates can be obtained by purely deterministic arguments, without invoking Itô's formula. In fact, note that equation (5.2) can equivalently be written as

$$Y_\lambda(t) + \int_0^t \big(AX_\lambda(s) + \beta_\lambda(X_\lambda(s))\big)\,ds = 0.$$

One has $Y_\lambda \in L^2(0,T;V)$, which follows at once by the properties of X_λ and by $B\cdot W \in L^2(\Omega;C([0,T];V_0))$. Similarly, since AX_λ and $\beta_\lambda(X_\lambda)$ belong to $L^2(\Omega;L^2(0,T;V^*))$, one also has, by the previous identity, $Y'_\lambda \in L^2(0,T;V^*)$. In particular, there exists $\Omega' \subset \Omega$, with $\mathbb{P}(\Omega') = 1$, such that

$$Y_\lambda(\omega) \in L^2(0,T;V), \qquad Y'_\lambda(\omega) \in L^2(0,T;V^*) \qquad \forall \omega \in \Omega'.$$

Lemma 4.5 then yields

$$\frac{1}{2}\big\|Y_\lambda(t)\big\|_H^2 + \int_0^t \big\langle AX_\lambda(s), Y_\lambda(s)\big\rangle\,ds + \int_0^t \big\langle \beta_\lambda(X_\lambda), Y_\lambda(s)\big\rangle\,ds = \frac{1}{2}\big\|X_0\big\|_H^2.$$

Lemma 5.5. *There exists a constant $N > 0$ such that*

$$\begin{aligned}\|X_\lambda\|^2_{L^2(\Omega;C([0,T];H))} + \|X_\lambda\|^2_{L^2(\Omega;L^2(0,T;V)} + \|\beta_\lambda(X_\lambda)X_\lambda\|_{L^1(\Omega;L^1(0,T;L^1(D)))}\\ < N\Big(\|X_0\|^2_{L^2(\Omega;H)} + \|B\|^2_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))}\Big).\end{aligned}$$

Proof. Itô's formula yields

$$\begin{aligned}\|X_\lambda(t)\|^2_H + 2\int_0^t \big\langle AX_\lambda(s), X_\lambda(s)\big\rangle\,ds + 2\int_0^t \big\langle \beta_\lambda(X_\lambda(s)), X_\lambda(s)\big\rangle\,ds\\ = \|X_0\|^2_H + 2\int_0^t X_\lambda(s)B(s)\,dW(s) + \frac{1}{2}\int_0^t \|B(s)\|^2_{\mathscr{L}^2(U,H)}\,ds,\end{aligned}$$

where X_λ in the stochastic integral on the right-hand side has to be interpreted as taking values in $H^* \simeq H$. The coercivity of A and the monotonicity of β_λ readily imply, after taking supremum in time and expectation,

$$\begin{aligned}\mathbb{E}\|X_\lambda\|^2_{C([0,T];H)} + 2C\,\mathbb{E}\|X_\lambda\|^2_{L^2(0,T;V)} + \mathbb{E}\int_0^T \big\langle \beta_\lambda(X_\lambda(s)), X_\lambda(s)\big\rangle\,ds\\ \lesssim \mathbb{E}\|X_0\|^2_H + \mathbb{E}\|B\|^2_{L^2(0,T;\mathscr{L}^2(U,H))} + \mathbb{E}\sup_{t\in[0,T]}\left|\int_0^t X_\lambda(s)B(s)\,dW(s)\right|,\end{aligned}$$

where, by Lemma 4.3,

$$\mathbb{E}\sup_{t\in[0,T]}\left|\int_0^t X_\lambda(s)B(s)\,dW(s)\right| \le \varepsilon\,\mathbb{E}\|X_\lambda\|^2_{C([0,T];H)} + N(\varepsilon)\,\mathbb{E}\int_0^T \|B(s)\|^2_{\mathscr{L}^2(U,H)}\,ds$$

for any $\varepsilon > 0$, whence the result follows choosing ε small enough. □

We now establish weak compactness properties for the sequence $(\beta_\lambda(X_\lambda))$.

Lemma 5.6. *The sequence $(\beta_\lambda(X_\lambda))$ is relatively weakly compact in $L^1(\Omega\times(0,T)\times D)$. Moreover, there exists a set $\Omega'' \subset \Omega$, with $\mathbb{P}(\Omega'') = 1$, such that $(\beta_\lambda(X_\lambda(\omega,\cdot))$ is weakly relatively compact in $L^1((0,T)\times D)$ for all $\omega \in \Omega''$.*

Proof. Recalling that, for any $y, r \in \mathbb{R}$, $j(y)+j^*(r) = ry$ if and only if $r \in \partial j(y) = \beta(y)$, one has

$$j\big((I+\lambda\beta)^{-1}x\big) + j^*\big(\beta_\lambda(x)\big) = \beta_\lambda(x)(I+\lambda\beta)^{-1}x \le \beta_\lambda(x)x \qquad \forall x \in \mathbb{R}. \tag{5.4}$$

In fact, since $\beta_\lambda \in \beta\circ(I+\lambda\beta)^{-1}$, it follows from $\beta = \partial j$ that $\beta_\lambda(x) \in \partial j\big((I+\lambda\beta)^{-1}x\big)$. Moreover, $\beta\big((I+\lambda\beta)^{-1}x\big)(I+\lambda\beta)^{-1}x \ge 0$ by monotonicity of β, hence the inequality in (5.4) follows since $(I+\lambda\beta)^{-1}$ is a contraction. The previous lemma thus implies, thanks to the symmetry of j^*, that there exists a constant N, independent of λ, such that, setting

$$\bar{N}(X_0,B) := N\Big(\|X_0\|^2_{L^2(\Omega;H)} + \|B\|^2_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))}\Big),$$

one has

$$\mathbb{E}\int_0^T\!\!\int_D j^*\big(|\beta_\lambda(X_\lambda)|\big) \le \mathbb{E}\int_0^T\!\!\int_D \beta_\lambda(X_\lambda)X_\lambda < \bar{N}(X_0,B).$$

Since j^* is superlinear at infinity, the sequence $(\beta_\lambda(X_\lambda))$ is uniformly integrable on $\Omega \times (0,T) \times D$ by the de la Vallée-Poussin criterion, hence weakly relatively compact in $L^1(\Omega \times (0,T) \times D)$ by a well-known theorem of Dunford and Pettis. The first assertion is thus proved.

By (5.3), since $Y_\lambda = X_\lambda - B \cdot W$, it follows that

$$\begin{aligned}\int_0^t \bigl\langle \beta_\lambda(X_\lambda(s)), X_\lambda(s) \bigr\rangle \, ds \lesssim \|X_0\|_H^2 + \|B \cdot W(t)\|_H^2 + \int_0^t \|B \cdot W(s)\|_V^2 \, ds \\ + \int_0^t \bigl\langle \beta_\lambda(X_\lambda(s)), B \cdot W(s) \bigr\rangle \, ds,\end{aligned}$$

where, by Young's inequality and convexity (recalling that $j^*(0) = 0$),

$$\int_0^t \bigl\langle \beta_\lambda(X_\lambda(s)), B \cdot W(s) \bigr\rangle \, ds \leq \frac{1}{2} \int_0^t \int_D j^*\bigl(\beta_\lambda(X_\lambda)\bigr) + \int_0^t \int_D j(2B \cdot W).$$

Rearranging terms and proceeding as in the (end of the) proof of Lemma 5.3, we infer that there exists a set $\Omega'' \subset \Omega$, with $\mathbb{P}(\Omega'') = 1$, and a function $M : \Omega'' \to \mathbb{R}$ such that

$$\int_0^T \bigl\langle \beta_\lambda(X_\lambda(\omega, s)), X_\lambda(\omega, s) \bigr\rangle \, ds < M(\omega) \qquad \forall \omega \in \Omega''. \tag{5.5}$$

The symmetry of j^* and (5.4) yield, as before, that, for any $\omega \in \Omega''$, $(\beta_\lambda(X_\lambda(\omega, \cdot)))$ is weakly relatively compact in $L^1((0,T) \times D)$. $\square$

In order to pass to the limit as $\lambda \to 0$, we are going to use Simon's compactness criterion, i.e. Lemma 4.6, and Brézis' Lemma 4.1.

Proposition 5.7. *There exists $\Omega' \subseteq \Omega$, with $\mathbb{P}(\Omega') = 1$, such that, for any $\omega \in \Omega'$, there exists a subsequence $\lambda' = \lambda'(\omega)$ of λ such that, as $\lambda' \to 0$,*

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

Proof. The first two convergence statements follow by Lemma 5.3, and the fourth one follows by Lemma 5.6. Let us show that the third convergence statement holds. In the following we omit the indication of ω, as no confusion can arise. Setting $Y_\lambda = X_\lambda - B \cdot W$, (5.2) can equivalently be written as the deterministic equation (with random coefficients) on V^*

$$Y_\lambda' + AX_\lambda + \beta_\lambda(X_\lambda) = 0,$$

where

$$\begin{aligned}
\bigl\|AX_\lambda\bigr\|_{L^1(0,T;V_0^*)} \lesssim \bigl\|AX_\lambda\bigr\|_{L^1(0,T;V^*)} \lesssim \bigl\|X_\lambda\bigr\|_{L^1(0,T;V)},\\
\bigl\|\beta_\lambda(X_\lambda)\bigr\|_{L^1(0,T;V_0^*)} \lesssim \bigl\|\beta_\lambda(X_\lambda)\bigr\|_{L^1(0,T;V^*)} \lesssim \bigl\|\beta_\lambda(X_\lambda)\bigr\|_{L^1(0,T;L^1(D))},
\end{aligned}$$

hence, again by Lemmata 5.3 and 5.6, $\|Y_\lambda'\|_{L^1(0,T;V_0^*)}$ is bounded uniformly over λ. Moreover, since $B\cdot W \in L^2(\Omega;C([0,T];V_0))$ and

$$\big\|Y_\lambda\big\|_{L^2(0,T;V)} \leq \big\|X_\lambda\big\|_{L^2(0,T;V)} + \big\|B\cdot W\big\|_{L^2(0,T;V)},$$

we conclude that (Y_λ) is bounded in $L^2(0,T;V)$. Simon's compactness criterion then implies that Y_λ, hence also X_λ, is relatively compact in $L^2(0,T;H)$. Since $X_{\lambda'} \rightharpoonup X$ in $L^2(0,T;V)$, it follows that

$$X_{\lambda'}(\omega,\cdot) \longrightarrow X(\omega,\cdot) \qquad \text{in } L^2(0,T;H),$$

thus completing the proof. □

We are now going to show that the couple (X,ξ) just constructed is indeed the unique solution to the equation with "smoothed" noise (5.1).

Proof of Proposition 5.1. In spite of the above preparations, the argument is quite long, so we subdivide it into several steps.

STEP 1. In the notation of Proposition 5.7, let $\omega \in \Omega'$ be arbitrary but fixed. Note that $X_{\lambda'} \to X$ in $L^2(0,T;H)$ implies that, passing to a further subsequence of λ', denoted with the same symbol for simplicity, $X_{\lambda'}(t) \to X(t)$ in H for almost all $t \in [0,T]$. Moreover, $X_{\lambda'} \rightharpoonup X$ in $L^2(0,T;V)$ implies that

$$\int_0^t AX_\lambda(s)\,ds \rightharpoonup \int_0^t AX(s)\,ds \qquad \text{in } V^*$$

for all $t\in[0,T]$. In fact, taking $\phi_0 \in V$ and $\phi := s \mapsto 1_{[0,t]}(s)\phi_0 \in L^2(0,t;V)$, one obviously has $A^*\phi \in L^2(0,t;V^*)$ and

$$\begin{aligned}\int_0^t \langle AX_\lambda(s),\phi_0\rangle\,ds &= \int_0^T \langle AX_\lambda(s),\phi(s)\rangle\,ds = \int_0^T \langle X_\lambda(s),A^*\phi(s)\rangle\,ds\\ &\longrightarrow \int_0^T \langle X(s),A^*\phi(s)\rangle\,ds = \int_0^t \langle AX(s),\phi_0\rangle\,ds.\end{aligned}$$

Similarly, $\beta_{\lambda'}(X_{\lambda'}) \rightharpoonup \xi$ in $L^1((0,T)\times D)$ implies

$$\int_0^t \beta_{\lambda'}(X_{\lambda'}(s))\,ds \rightharpoonup \int_0^t \xi(s)\,ds \qquad \text{in } L^1(D)$$

for all $t\in[0,T]$. In particular, passing to the limit as $\lambda'\to 0$ in the regularized equation (5.2) yields

$$X(t) + \int_0^t AX(s)\,ds + \int_0^t \xi(s)\,ds = X_0 + B\cdot W(t) \qquad \text{in } V_0^* \text{ for a.a. } t\in[0,T].$$

Since $AX \in L^2(0,T;V^*) \hookrightarrow L^1(0,T;V_0^*)$ and $\xi \in L^1(0,T;L^1(D)) \hookrightarrow L^1(0,T;V_0^*)$, recalling that $B\cdot W \in C([0,T];V_0)$, we infer that $X\in C([0,T];V_0^*)$, hence the previous identity is true for all $t\in[0,T]$. Moreover, it follows from $X\in L^\infty(0,T;H)$ that $X \in C_w([0,T];H)$, thanks Lemma 4.4. Note also that all terms expect the second one on the left-hand side take values in $L^1(D)$, and all terms except the third one on the left-hand side take values in V^*, hence the above identity holds true also in $L^1(D)\cap V^*$.

Let us now show that $\xi \in \beta(X)$ a.e. in $(0,T)\times D$: $X_{\lambda'} \to X$ in $L^2(0,T;H)$ implies that, passing to a subsequence of λ', still denoted by the same symbol, $X_{\lambda'} \to X$ a.e. in $(0,T)\times D$, hence also $(I+\lambda'\beta)^{-1}X_{\lambda'} \to X$ a.e. in $(0,T)\times D$. Since $\beta_{\lambda'}(X_{\lambda'}) \in \beta((I+\lambda'\beta)^{-1}X_{\lambda'})$ a.e. in $(0,T)\times D$ and $\beta_{\lambda'}(X_{\lambda'})(I+\lambda'\beta)^{-1}X_{\lambda'}$ is bounded in $L^1((0,T)\times D)$ by (5.5), Brézis' Lemma 4.1 implies the claim. These relations and the weak convergence $\beta_{\lambda'}(X_{\lambda'}) \rightharpoonup \xi$ in $L^1((0,T)\times D)$ also imply, by the weak lower semicontinuity of convex integrals, that

$$\begin{aligned}\int_0^T\!\!\int_D \big(j(X)+j^*(\xi)\big) &\le \liminf_{\lambda'\to 0}\int_0^T\!\!\int_D \big(j((I+\lambda' A)^{-1}X_{\lambda'})+j^*(\beta_{\lambda'}(X_{\lambda'}))\big)\\ &= \liminf_{\lambda'\to 0}\int_0^T\!\!\int_D \beta_{\lambda'}(X_{\lambda'})(I+\lambda' A)^{-1}X_{\lambda'} \le N,\end{aligned}$$

where N is a constant that depends on ω.

Step 2. Still keeping ω fixed as in the previous step, we are going to show that the limits X and ξ constructed above are unique. Suppose there exist (X_i,ξ_i), $\xi_i \in \beta(X_i)$ a.e. in $(0,T)\times D$, $i=1,2$, such that

$$X_i(t)+\int_0^t AX_i(s)\,ds+\int_0^t \xi_i(s)\,ds = X_0 + B\cdot W(t)$$

in $L^1(D)\cap V^*$ for all $t\in[0,T]$. Setting $X=X_1-X_2$ and $\xi=\xi_1-\xi_2$, it is enough to show that

$$X(t)+\int_0^t AX(s)\,ds+\int_0^t \xi(s)\,ds = 0 \tag{5.6}$$

in $L^1(D)\cap V^*$ for all $t\in[0,T]$ implies $X=0$ and $\xi=0$. By the hypotheses on A, there exists $m\in\mathbb{N}$ such that $(I+\delta A)^{-m}$ maps $L^1(D)$ in $L^\infty(D)$. Therefore, setting

$$X^\delta := (I+\delta A)^{-m}X, \qquad \xi^\delta := (I+\delta A)^{-m}\xi,$$

one has

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

for all $t\in[0,T]$, for which Itô's formula and monotonicity of A yield

$$\frac{1}{2}\big\|X^\delta(t)\big\|_H^2+\int_0^t\!\!\int_D \xi^\delta(s,x)X^\delta(s,x)\,dx\,ds \le 0.$$

We can now take the limit as $\delta\to 0$. Since $(I+\delta A)^{-m}$ converges, in the strong operator topology, to the identity in $\mathscr{L}(H)$, one has $\|X^\delta(t)\|_H \to \|X(t)\|_H$ for all $t\in[0,T]$. Passing to a subsequence of δ, still denoted by the same symbol, we also have $X^\delta \to X$ and $\xi^\delta \to \xi$ a.e. in $(0,T)\times D$, hence $X^\delta\xi^\delta \to X\xi$ a.e. in $(0,T)\times D$. Let us show that $(X^\delta\xi^\delta)$ is uniformly integrable: by the symmetry of j and j^*, and the abstract Jensen inequality of Lemma 4.2, we have

$$|X_\delta\xi_\delta| \le j(X_\delta)+j^*(\xi_\delta) \le (I+\delta A)^{-m}\big(j(X)+j^*(\xi)\big),$$

where the term on the right-hand side converges to $j(X)+j^*(\xi)$ in $L^1((0,T)\times D)$ as $\delta\to 0$, hence $(X^\delta\xi^\delta)$ is indeed uniformly integrable on $(0,T)\times D$. It follows by Vitali's convergence theorem that, for any $t\in[0,T]$,

$$\int_0^t\int_D X^\delta\xi^\delta \to \int_0^t\int_D X\xi,$$

hence also

$$\frac{1}{2}\big\|X(t)\big\|_H^2+\int_0^t\int_D X(s,x)\xi(s,x)\,dx\,ds\le 0.$$

The monotonicity of β immediately implies that $X(t)=0$ for all $t\in[0,T]$. Substituing in (5.6), we are left with $\int_0^t \xi(s)\,ds=0$ in $L^1(D)$ for all $t\in[0,T]$, so that also $\xi=0$, and uniqueness is proved.

Step 3. The solution (X,ξ) does *not* have, a priori, any measurability in ω, because of the way it has been constructed. We are going to show that in fact X and ξ are predictable processes. The reasoning for X is simple: with ω fixed, we have proved that from any subsequence of λ one can extract a further subsequence λ', depending on ω, such that the convergences of Proposition 5.7 take place, and the limit (X,ξ) is unique. This implies, by a well-known criterion of classical analysis, that the same convergences hold along the original sequence λ, which does *not* depend on ω. The convergence of $X_\lambda(\omega,\cdot)$ to $X(\omega,\cdot)$ in $L^2(0,T;H)$ implies that $X:\Omega\to L^2(0,T;H)$ is measurable and $X_\lambda(\omega,t)$ converges to $X(\omega,t)$ in H in $\mathbb{P}\otimes dt$-measure, hence $X_{\bar\lambda}(\omega,t)\to X(\omega,t)$ in H $\mathbb{P}\otimes dt$-a.e. along a subsequence $\bar\lambda$ of λ. Since X_λ is predictable, being adapted with continuous trajectories in H, we infer that X is predictable. Unfortunately a similar reasoning does *not* work for ξ, because $\xi_\lambda(\omega):=\beta_\lambda(X_\lambda(\omega))$ converges only weakly in $L^1((0,T)\times D)$ for $\mathbb{P}$-a.a. $\omega\in\Omega$.[8] We shall prove instead that a subsequence of $\xi_\lambda:=\beta_\lambda(X_\lambda)$ converges weakly to ξ in $L^1(\Omega\times(0,T)\times D)$. In fact, let $g\in L^\infty((0,T)\times D)$ be arbitrary but fixed. Then, setting

$$F_\lambda(\omega):=\int_0^T\int_D \xi_\lambda(\omega,s,x)g(s,x)\,dx\,ds,\qquad F(\omega):=\int_0^T\int_D \xi(\omega,s,x)g(s,x)\,dx\,ds,$$

we have $F_\lambda\to F$ in probability, and we claim that $F_\lambda\to F$ weakly in $L^1(\Omega)$. Let $h\in L^\infty(\Omega)$ be arbitrary but fixed, and introduce the even convex function

$$j_0:=j^*(\cdot/M),\qquad M:=\frac{1}{\big(\|g\|_{L^\infty((0,T)\times D)}\vee 1\big)\big(\|h\|_{L^\infty(\Omega)}\vee 1\big)}.$$

Then, by Jensen's inequality,

$$\begin{aligned}\mathbb{E}\,j_0(F_\lambda h)&=\mathbb{E}\,j_0\biggl(\int_0^T\int_D \xi_\lambda(\omega,s,x)g(s,x)h(\omega)\,dx\,ds\biggr)\\&\lesssim_{T,|D|}\mathbb{E}\int_0^T\int_D j_0\big(\xi_\lambda(\omega,s,x)g(s,x)h(\omega)\big)\,dx\,ds\\&\le \mathbb{E}\int_0^T\int_D j^*\big(\xi_\lambda(\omega,s,x)\big)\,dx\,ds,\end{aligned}$$

[8] One may indeed deduce, using Mazur's lemma, that there exists, for each ω in a set of probability one, a sequence $(\tilde\xi_{\mu(\omega)}(\omega))_{\mu(\omega)}$ in the convex envelope of $(\xi_\lambda(\omega))_\lambda$ that converges to $\xi(\omega)$. However, the map $\omega\mapsto\tilde\xi_{\mu(\omega)}(\omega)$ needs not be measurable, hence we cannot infer measurability of its limit ξ.

where the last term is bounded by a constant independent of λ, as proved in Lemma 5.6. Since j_0 inherits the superlinearity at infinity of j^*, the criterion of de la Vallée Poussin implies that $F_\lambda h$ is uniformly integrable, hence, since $F_\lambda h \to Fh$ in probability, that $F_\lambda h \to Fh$ strongly in $L^1(\Omega)$ by Vitali's theorem. As h was arbitrary, this implies that $F_\lambda \to F$ weakly in $L^1(\Omega)$, thus also that $\xi_\lambda \to \xi$ weakly in $L^1(\Omega \times (0,T) \times D)$ by arbitrariness of g. By the canonical identification of $L^1(\Omega \times (0,T) \times D)$ with $L^1(\Omega \times (0,T); L^1(D))$ and Mazur's lemma (see, e.g., [11, 7), p. 360]), there exists a sequence $(\zeta_n)_{n\in\mathbb{N}}$ of convex combinations of (ξ_λ) that converges strongly to ξ in $L^1(D)$ in $\mathbb{P} \otimes dt$-measure, hence $\mathbb{P} \otimes dt$-a.e. passing to a subsequence of n. Since ξ_λ, hence ζ_n, are predictable for all λ and n, respectively, it follows that ξ is a predictable $L^1(D)$-valued process and $\xi : \Omega \to L^1((0,T) \times D))$ is measurable. Moreover, since $X_\lambda(\omega, \cdot) \to X(\omega, \cdot)$ in $L^2(0,T;H)$ for $\mathbb{P}$-a.a. ω and $(X_\lambda)_\lambda$ is bounded in $L^2(\Omega; L^2(0,T;V))$, it follows that $X_\lambda \rightharpoonup X$ in $L^2(\Omega; L^2(0,T;V))$. Therefore, an entirely analogous argument based on Mazur's lemma yields that $X : \Omega \to L^2(0,T;V)$ is measurable.

Step 4. As last step, we are going to show that X and ξ satisfy also estimates in expectation. In particular, the weak and weak* lower semicontinuity of the norm ensures that, for $\mathbb{P}$-almost all $\omega \in \Omega$,

$$\begin{aligned}
\big\|X(\omega,\cdot)\big\|_{L^2(0,T;V)} &\le \liminf_{\lambda\to 0}\big\|X_\lambda(\omega,\cdot)\big\|_{L^2(0,T;V)},\\
\big\|X(\omega,\cdot)\big\|_{L^\infty(0,T;H)} &\le \liminf_{\lambda\to 0}\big\|X_\lambda(\omega,\cdot)\big\|_{L^\infty(0,T;H)},\\
\big\|\xi(\omega,\cdot)\big\|_{L^1(Q)} &\le \liminf_{\lambda\to 0}\big\|\beta_\lambda(X_\lambda(\omega,\cdot))\big\|_{L^1(Q)}.
\end{aligned}$$

Taking expectations and recalling Lemmata 5.5 and 5.6, it follows by Fatou's lemma that, for a constant N,

$$\begin{aligned}
\mathbb{E}\big\|X\big\|^2_{L^2(0,T;V)} &\le \mathbb{E}\Big(\liminf_{\lambda\to 0}\big\|X_\lambda\big\|^2_{L^2(0,T;V)}\Big) \le \liminf_{\lambda\to 0}\mathbb{E}\big\|X_\lambda\big\|^2_{L^2(0,T;V)} < N,\\
\mathbb{E}\big\|X\big\|^2_{L^\infty(0,T;H)} &\le \mathbb{E}\Big(\liminf_{\lambda\to 0}\big\|X_\lambda\big\|^2_{L^\infty(0,T;H)}\Big) \le \liminf_{\lambda\to 0}\mathbb{E}\big\|X_\lambda\big\|^2_{L^\infty(0,T;H)} < N,\\
\mathbb{E}\big\|\xi\big\|_{L^1(0,T;L^1(D))} &\le \mathbb{E}\Big(\liminf_{\lambda\to 0}\big\|\xi_\lambda\big\|_{L^1(0,T;L^1(D))}\Big) \le \liminf_{\lambda\to 0}\mathbb{E}\big\|\xi_\lambda\big\|_{L^1(0,T;L^1(D))} < N,
\end{aligned}$$

i.e.

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

The proof is thus complete. □

We conclude this section with a corollary that will be used in the following.

Corollary 5.8. *There exists a constant N such that*

$$\mathbb{E}\int_0^T\!\!\int_D \big(j(X) + j^*(\xi)\big) < N\Big(\big\|X_0\big\|^2_{L^2(\Omega;H)} + \big\|B\big\|^2_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))}\Big).$$

Proof. Thanks to Step 3 in the previous proof, there exists a sequence λ, independent of ω, such that $X_\lambda \to X$ a.e. in $(0,T) \times D$ and $\beta_\lambda(X_\lambda) \to \xi$ weakly in $L^1((0,T) \times D)$.

Proceeding as in the first part of the proof of Lemma 5.6, Lemma 5.5 implies that there exists a constant N such that

$$\mathbb{E}\int_0^T\int_D \big(j(I+\lambda\beta)^{-1}X_\lambda) + j^*(\beta_\lambda(X_\lambda))\big)\,dx\,ds < \bar N(X_0,B),$$

where $\bar N(X_0,B) := N\big(\|X_0\|^2_{L^2(\Omega;H)} + \|B\|^2_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))}\big)$. Therefore, in analogy to Step 4 of the previous proof, two applications of Fatou's lemma yield

$$\mathbb{E}\int_0^T\int_D j(X) \leq \liminf_{\lambda\to 0}\mathbb{E}\int_0^T\int_D j((I+\lambda\beta)^{-1}X_\lambda) < \bar N(X_0,B),$$

as well as, by the weak lower semicontinuity of convex integrals and Fatou's lemma again,

$$\mathbb{E}\int_0^T\int_D j^*(\xi) \leq \liminf_{\lambda\to 0}\mathbb{E}\int_0^T\int_D j^*(\beta_\lambda(X_\lambda)) < \bar N(X_0,B). \qquad \square$$

6 Well-posedness with additive noise

In this section we prove well-posedness for the equation

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

where B is an $\mathscr{L}^2(U,H)$-valued process. Note that this is just equation (1.1) with additive noise.

Proposition 6.1. *Assume that $X_0 \in L^2(\Omega,\mathscr{F}_0,\mathbb{P};H)$ and that*

$$B \in L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))$$

is measurable and adapted. Then equation (5.1) *is well posed in $\mathscr{J}$. Moreover, $X(\omega,\cdot) \in C_w([0,T];H)$ for $\mathbb{P}$-almost all $\omega\in\Omega$.*

Proof. We shall proceed in several steps: first we approximate the coefficient B in such a way that the corresponding equation can be uniquely solved by the methods of the previous section. Then we pass to the limit in an appropriate way, obtaining a solution to (6.1), which is then shown to be unique.

STEP 1. By Assumption A(iv), there exists $m\in\mathbb{N}$ such that $(I+A)^{-m}$ maps continuously L^1 to L^∞. The space $V_0 := \mathsf{D}(A^m)$, endowed with inner product

$$\langle u,v\rangle_{V_0} := \langle u,v\rangle_H + \langle A^m u, A^m v\rangle_H, \qquad u,v\in\mathsf{D}(A^m),$$

is a Hilbert space densely and continuously embedded in V. Moreover, the diagram

$$\mathsf{D}(A^m) \xrightarrow{(I+A)^m} L^1(D) \xrightarrow{(I+A)^{-m}} L^\infty(D)$$

immediately shows that V_0 is also continuously embedded in $L^\infty(D)$. In particular, all hypotheses on V_0 of the previous section are met. Moreover, by the ideal property of Hilbert-Schmidt operators, setting, for any $\varepsilon>0$,

$$B^\varepsilon := (I+\varepsilon A)^{-m}B,$$

we have $B^\varepsilon \in L^2(\Omega; L^2(0,T; \mathscr{L}^2(U,V_0)))$. Then it follows by Proposition 5.1 that, for any $\varepsilon > 0$, there exist predictable processes

$$X^\varepsilon \in L^2(\Omega; L^\infty(0,T;H)) \cap L^2(\Omega; L^2(0,T;V)),$$
$$\xi^\varepsilon \in L^1(\Omega \times (0,T) \times D),$$

with $X^\varepsilon(\omega,\cdot) \in C_w([0,T];H)$ for $\mathbb{P}$-almost all $\omega \in \Omega$, such that

$$X^\varepsilon(t) + \int_0^t AX^\varepsilon(s)\,ds + \int_0^t \xi^\varepsilon(s)\,ds = X_0 + \int_0^t B^\varepsilon(s)\,dW(s) \tag{6.2}$$

in $V^* \cap L^1(D)$ for all $t \in [0,T]$. Moreover, $\xi^\varepsilon \in \beta(X^\varepsilon)$ a.e. in $(0,T) \times D$ and $j(X^\varepsilon) + j^*(\xi^\varepsilon) \in L^1((0,T) \times D)$ $\mathbb{P}$-almost surely.

Step 2. For any $\varepsilon > 0$, the equation in V^*

$$X^\varepsilon_\lambda(t) + \int_0^t AX^\varepsilon_\lambda(s)\,ds + \int_0 \beta_\lambda(X^\varepsilon_\lambda(s))\,ds = X_0 + \int_0^t B^\varepsilon(s)\,dW(s)$$

admits a unique (variational) strong solution X^ε_λ. Taking into account the coercivity of A and the monotonicity of β_λ, Itô's formula yields, for any $\delta > 0$,

$$\begin{aligned}&\big\|X^\varepsilon_\lambda(t) - X^\delta_\lambda(t)\big\|_H^2 + \int_0^t \big\|X^\varepsilon_\lambda(s) - X^\delta_\lambda(s)\big\|_V^2\,ds \\ &\qquad \lesssim \int_0^t \big(X^\varepsilon_\lambda(s) - X^\delta_\lambda(s)\big)\big(B^\varepsilon(s) - B^\delta(s)\big)\,dW(s) + \int_0^t \big\|B^\varepsilon(s) - B^\delta(s)\big\|^2_{\mathscr{L}^2(U,H)}\,ds.\end{aligned}$$

Taking supremum in time and expectation, it easily follows from Lemma 4.3 that

$$\begin{aligned}\big\|X^\varepsilon_\lambda - X^\delta_\lambda\big\|_{L^2(\Omega;L^\infty(0,T;H))} &+ \big\|X^\varepsilon_\lambda - X^\delta_\lambda\big\|_{L^2(\Omega;L^2(0,T;V))} \\ &\lesssim \big\|B^\varepsilon - B^\delta\big\|_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))}.\end{aligned}$$

On the other hand, the proof of Proposition 5.1 shows that there exists a sequence λ, independent of ε, such that, for $\mathbb{P}$-almost all $\omega \in \Omega$,

$$\begin{aligned} X^\varepsilon_\lambda(\omega,\cdot) &\xrightarrow{\;*\;} X^\varepsilon(\omega,\cdot) && \text{in } L^\infty(0,T;H),\\ X^\varepsilon_\lambda(\omega,\cdot) &\rightharpoonup X^\varepsilon(\omega,\cdot) && \text{in } L^2(0,T;V),\\ \beta_\lambda(X^\varepsilon_\lambda(\omega,\cdot)) &\rightharpoonup \xi^\varepsilon(\omega,\cdot) && \text{in } L^1((0,T)\times D)\end{aligned}$$

as $\lambda \to 0$. Since the weak* limit in $L^\infty(0,T;H)$ as $\lambda \to 0$ of $X^\varepsilon_\lambda - X^\delta_\lambda$ is $X^\varepsilon - X^\delta$, the weak* lower semicontinuity of the norm implies

$$\big\|X^\varepsilon - X^\delta\big\|_{L^\infty(0,T;H)} \leq \liminf_{\lambda \to 0} \big\|X^\varepsilon_\lambda - X^\delta_\lambda\big\|_{L^\infty(0,T;H)},$$

thus also, by Fatou's lemma,

$$\mathbb{E}\big\|X^\varepsilon - X^\delta\big\|^2_{L^\infty(0,T;H)} \leq \mathbb{E}\liminf_{\lambda\to 0}\big\|X^\varepsilon_\lambda - X^\delta_\lambda\big\|^2_{L^\infty(0,T;H)} \lesssim \mathbb{E}\big\|B^\varepsilon - B^\delta\big\|^2_{L^2(0,T;\mathscr{L}^2(U,H))}.$$

An entirely similar argument yields

$$\mathbb{E}\big\|X^\varepsilon - X^\delta\big\|^2_{L^2(0,T;V)} \lesssim \mathbb{E}\big\|B^\varepsilon - B^\delta\big\|^2_{L^2(0,T;\mathscr{L}^2(U,H))},$$

so that

$$\begin{aligned}\big\|X^\varepsilon - X^\delta\big\|_{L^2(\Omega;L^\infty(0,T;H))} + \big\|X^\varepsilon - X^\delta\big\|_{L^2(\Omega;L^2(0,T;V))}\\ \lesssim \big\|B^\varepsilon - B^\delta\big\|_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))}.\end{aligned}$$

Taking into account that $\|B^\varepsilon - B\|_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))} \to 0$ as $\varepsilon \to 0$, it follows that (X^ε) is a Cauchy sequence in $E := L^2(\Omega; L^\infty(0,T;H)) \cap L^2(\Omega; L^2(0,T;V))$, hence there exists $X \in E$ such that X^ε converges (strongly) to X in E as $\varepsilon \to 0$. In particular, the limit process X is predictable. Moreover, by Corollary 5.8, there exists a constant N such that

$$\begin{aligned}\mathbb{E}\int_0^T\int_D \big(j(X^\varepsilon) + j^*(\xi^\varepsilon)\big)\,dx\,ds &< N\Big(\big\|X_0\big\|^2_{L^2(\Omega;H)} + \big\|B^\varepsilon\big\|^2_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))}\Big)\\ &\le N\Big(\big\|X_0\big\|^2_{L^2(\Omega;H)} + \big\|B\big\|^2_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))}\Big),\end{aligned} \tag{6.3}$$

as it follows by the ideal property of Hilbert-Schmidt operators and the contractivity of $(I+\varepsilon A)^{-1}$. The criterion by de la Vallée Poussin then implies that (ξ^ε) is uniformly integrable on $\Omega \times (0,T) \times D$, hence, by the Dunford-Pettis theorem, (ξ^ε) is weakly relatively compact in $L^1(\Omega \times (0,T) \times D)$. Therefore, passing to a subsequence of ε, denoted by the same symbol, there exists ξ belonging to the latter space such that $\xi^\varepsilon \to \xi$ therein in the weak topology. In particular, by an argument based on Mazur's lemma, entirely analogous to that used in Step 3 of the proof of Proposition 5.1, one infers that ξ is a predictable process.

Step 3. We can now pass to the limit as $\varepsilon \to 0$ in Equation (6.2), by a reasoning analogous to the one use in Step 1 of the proof of Proposition 5.1. As proved in the previous step, X^ε converges strongly to X in $L^2(\Omega; L^\infty(0,T;H))$, hence

$$\operatorname*{ess\,sup}_{t\in[0,T]}\big\|X^\varepsilon(t) - X(t)\big\|_H \to 0$$

in probability as $\varepsilon \to 0$. Let $\phi_0 \in V_0$ be arbitrary. Since $V_0 \hookrightarrow L^\infty(D)$, one has

$$\big\langle X^\varepsilon(t), \phi_0\big\rangle \to \big\langle X(t), \phi_0\big\rangle$$

in probability for almost all $t \in [0,T]$. Let us set, for an arbitrary but fixed $t \in [0,T]$, $\phi : s \mapsto 1_{[0,t]}(s)\phi_0 \in L^2(0,T;V)$, so that $A\phi \in L^2(0,T;V^*)$. Recalling that $X^\varepsilon \to X$ (strongly, hence also weakly) in $L^2(\Omega; L^2(0,T;V))$, it follows immediately that $X^\varepsilon \rightharpoonup X$ in $L^2(0,T;V)$ in measure, hence

$$\begin{aligned}\int_0^t \langle AX^\varepsilon, \phi_0\rangle\,ds = \int_0^T \langle AX^\varepsilon(s), \phi(s)\rangle\,ds = \int_0^T \langle X^\varepsilon(s), A\phi(s)\rangle\,ds\\ \to \int_0^T \langle X(s), A\phi(s)\rangle\,ds = \int_0^t \langle AX(s), \phi_0\rangle\,ds\end{aligned}$$

in probability as $\varepsilon \to 0$. A completely analogous reasoning shows that

$$\int_0^t \langle \xi^\varepsilon(s), \phi_0\rangle\,ds \to \int_0^t \langle \xi(s), \phi_0\rangle\,ds$$

in probability as $\varepsilon \to 0$. Doob's maximal inequality and the convergence

$$\left\| B^\varepsilon - B \right\|_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))} \xrightarrow{\varepsilon\to 0} 0$$

readily yield also that $B^\varepsilon \cdot W(t) \to B \cdot W(t)$ in H in probability for all $t \in [0,T]$. In particular, since $\phi_0 \in V_0$ and $t \in [0,T]$ are arbitrary, we infer that

$$X(t) + \int_0^t AX(s)\,ds + \int_0^t \xi(s)\,ds = X_0 + \int_0^t B(s)\,dW(s)$$

holds in V_0^* for almost all t. Recalling that $\xi \in L^1(0,T;L^1(D)) \hookrightarrow L^1(0,T;V_0^*)$, so that all terms except the first on the left-hand side have trajectories in $C([0,T];V_0^*)$, we conclude that the identity holds for all $t \in [0,T]$. Moreover, thanks to Lemma 4.4, $X \in C([0,T];V_0^*)$ and $X \in L^\infty(0,T;H)$ imply $X \in C_w([0,T];H)$. Note also that all terms bar the second [third] one on the left-hand side are $L^1(D)$-valued [V^*-valued], hence the identity holds in $L^1(D) \cap V^*$ for all $t \in [0,T]$.

Step 4. Convergence of $X^\varepsilon \to X$ in $L^2(\Omega;L^\infty(0,T;H))$ implies convergence in measure in $\Omega \times (0,T) \times D$, hence, by Fatou's lemma, (6.3) yields

$$\mathbb{E}\int_0^T\int_D j(X) < \bar{N}(X_0,B),$$

where $\bar{N}(X_0,B)$ is the constant appearing in the last term of (6.3). Similarly, since $\xi^\varepsilon \to \xi$ weakly in $L^1(\Omega \times (0,T) \times D)$, (6.3) and the weak lower semicontinuity of convex integrals yield

$$\mathbb{E}\int_0^T\int_D j^*(\xi) < \bar{N}(X_0,B).$$

To complete the proof of existence, we only need to show that $\xi \in \beta(X)$ a.e. in $\Omega \times (0,T) \times D$. Note that, passing to a subsequence of ε, still denoted by the same symbol, we have $X^\varepsilon \to X$ a.e. in $\Omega \times (0,T) \times D$. Recalling that $\xi^\varepsilon \in \beta(X^\varepsilon)$ a.e. in $\Omega \times (0,T) \times D$, (6.3) again implies

$$\mathbb{E}\int_0^T\int_D X^\varepsilon\xi^\varepsilon = \mathbb{E}\int_0^T\int_D \bigl(j(X^\varepsilon) + j^*(\xi^\varepsilon)\bigr) < \bar{N}(X_0,B).$$

It follows by monotonicity that $X^\varepsilon\xi^\varepsilon \geq 0$, hence $X^\varepsilon\xi^\varepsilon \in L^1(\Omega \times (0,T) \times D)$. Brézis' Lemma 4.1 then yields $\xi \in \beta(X)$ a.e. in $\Omega \times (0,T) \times D$.

Uniqueness and continuous dependence of the solution on the initial datum is an immediate consequence of the next result. □

We first need to introduce weighted (in time) versions of some spaces of processes. For any $p \in [1,\infty]$ and $\alpha \geq 0$, we shall denote by $L^p_\alpha(0,T)$ the space $L^p(0,T)$ endowed with the norm $\|f\|_{L^p_\alpha(0,T)} := \|t \mapsto e^{-\alpha t}f(t)\|_{L^p(0,T)}$. It is clear that $L^p(0,T)$ and $L^p_\alpha(0,T)$, for different values of α, are all isomorphic (their norms are equivalent). Completely similar notation will be used for vector-valued L^p and L^p_α spaces. For typographical economy, restricted only to the formulation of the following proposition, let us define the Banach space

$$F_\alpha := L^2(\Omega;L^\infty_\alpha(0,T;H)) \cap L^2(\Omega;L^2_\alpha(0,T;V)),$$

endowed with the norm

$$\|\cdot\|_{F_\alpha} := \|\cdot\|_{L^2(\Omega;L^\infty_\alpha(0,T;H))\cap L^2(\Omega;L^2_\alpha(0,T;V))} + \sqrt{\alpha}\|\cdot\|_{L^2(\Omega;L^2_\alpha(0,T;H))}$$

Proposition 6.2. *Let (X_1,ξ_1), $(X_2,\xi_2) \in \mathscr{J}$ be solutions to* (6.1) *with initial values X_{01}, $X_{02} \in L^2(\Omega,\mathscr{F}_0;H)$ and progressively measurable diffusion coefficients B_1, $B_2 \in L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))$, respectively. Then, for any $\alpha \geq 0$,*

$$\|X_1 - X_2\|_{F_\alpha} \lesssim \|X_{01} - X_{02}\|_{L^2(\Omega;H)} + \|B_1 - B_2\|_{L^2(\Omega;L^2_\alpha(0,T;\mathscr{L}^2(U,H)))}.$$

In particular, there is a unique solution $(X,\xi) \in \mathscr{J}$ to (6.1)*.*

Proof. Setting

$$Y := X_1 - X_2, \qquad Y_0 := X_{01} - X_{02}, \qquad G := B_1 - B_2,$$

one has

$$Y(t) + \int_0^t AY(s)\,ds + \int_0^t \zeta(s)\,ds = Y_0 + \int_0^t G(s)\,dW(s)$$

in $V^* \cap L^1(D)$, where $\zeta := \xi_1 - \xi_2$, and ξ_1, ξ_2 are defined in the obvious way. By the hypotheses on A, there exists $m \in \mathbb{N}$ such that, using the notation $h^\delta := (I + \delta A)^{-m} h$ for any h for which it makes sense,

$$AY^\delta,\ \zeta^\delta \in L^1(\Omega;L^1(0,T;H)),$$

while Y_0^δ and G^δ have the same integrability properties of Y, Y_0 and G, respectively. In particular, we have

$$Y^\delta(t) + \int_0^t AY^\delta(s)\,ds + \int_0^t \zeta^\delta(s)\,ds = Y_0^\delta + \int_0^t G^\delta(s)\,dW(s)$$

in V^*. Let $\alpha > 0$ be arbitrary but fixed, and add a superscript α to any process that is multiplied pointwise by the function $t \mapsto e^{-\alpha t}$. The integration by parts formula yields

$$Y^{\delta,\alpha}(t) + \int_0^t (A + \alpha I) Y^{\delta,\alpha}(s)\,ds + \int_0^t \zeta^{\delta,\alpha}(s)\,ds = Y_0^\delta + \int_0^t G^{\delta,\alpha}(s)\,dW(s),$$

to which we can apply Itô's formula for the square of the norm in H, obtaining, using the coercivity of A,

$$\begin{aligned}
&\big\|Y^{\delta,\alpha}(t)\big\|_H^2 + 2\alpha \int_0^t \big\|Y^{\delta,\alpha}(s)\big\|_H^2\,ds + 2C \int_0^t \big\|Y^{\delta,\alpha}(s)\big\|_V^2\,ds \\
&\qquad + 2\int_0^t \big\langle Y^{\delta,\alpha}(s), \zeta^{\delta,\alpha}(s)\big\rangle\,ds \\
&\qquad\qquad \leq \big\|Y_0^\delta\big\|_H^2 + \int_0^t Y^{\delta,\alpha}(s) G^{\delta,\alpha}(s)\,dW(s) + \int_0^t \big\|G^{\delta,\alpha}(s)\big\|_{\mathscr{L}^2(U,H)}^2\,ds.
\end{aligned}$$

We are now going to pass to the limit as $\delta \to 0$: the first term on the left-hand side and on the right-hand side clearly converge to $\|Y^\alpha(t)\|_H^2$ and $\|Y_0\|_H^2$, respectively. Since $(I+\delta A)^{-1}$ converges to the identity in H as well as in V in the strong operator topology, the dominated convergence theorem yields

$$\int_0^t \big\|Y^{\delta,\alpha}(s)\big\|_V^2\,ds \longrightarrow \int_0^t \big\|Y^\alpha(s)\big\|_V^2\,ds,$$

$$\int_0^t \big\|G^{\delta,\alpha}(s)\big\|_{\mathscr{L}^2(U,H)}^2\,ds \longrightarrow \int_0^t \big\|G^\alpha(s)\big\|_{\mathscr{L}^2(U,H)}^2\,ds$$

as $\delta \to 0$ for all $t \in [0,T]$. Defining the real local martingales

$$M^{\delta,\alpha} := (Y^{\delta,\alpha}G^{\delta,\alpha}) \cdot W, \qquad M^\alpha := (Y^\alpha G^\alpha) \cdot W,$$

in order to establish convergence in probability (uniformly on compact sets) of the sequence $M^{\delta,\alpha}$ to M^α as $\delta \to 0$, it is sufficient to show that $[M^{\delta,\alpha} - M^\alpha, M^{\delta,\alpha} - M^\alpha]_T$ converges to zero in probability. To this purpose, note that

$$\begin{aligned}
[M^{\delta,\alpha} - M^\alpha, M^{\delta,\alpha} - M^\alpha]_T^{1/2} &= \big\|Y^{\delta,\alpha}G^{\delta,\alpha} - Y^\alpha G^\alpha\big\|_{L^2(0,T;\mathscr{L}^2(U,\mathbb{R}))} \\
&\leq \big\|Y^{\delta,\alpha}G^{\delta,\alpha} - Y^{\delta,\alpha}G^\alpha\big\|_{L^2(0,T;\mathscr{L}^2(U,\mathbb{R}))} \\
&\quad + \big\|Y^{\delta,\alpha}G^\alpha - Y^\alpha G^\alpha\big\|_{L^2(0,T;\mathscr{L}^2(U,\mathbb{R}))},
\end{aligned}$$

where

$$\big\|Y^{\delta,\alpha}(t)G^{\delta,\alpha}(t) - Y^{\delta,\alpha}(t)G^\alpha(t)\big\|_{\mathscr{L}^2(U,\mathbb{R})} \leq \big\|Y^\alpha(t)\big\|_H \big\|G^{\delta,\alpha}(t) - G^\alpha(t)\big\|_{\mathscr{L}^2(U,H)}$$

for all $t \in [0,T]$. Since the right-hand side converges to 0 as $\delta \to 0$ and it is bounded by $2\|Y^\alpha\|_{L^\infty(0,T;H)}\|G^\alpha(t)\|_{\mathscr{L}^2(U,H)}$, and $G^\alpha \in L^2(0,T;\mathscr{L}^2(U,H))$, the dominated convergence theorem yields

$$\big\|Y^{\delta,\alpha}G^{\delta,\alpha} - Y^{\delta,\alpha}G^\alpha\big\|_{L^2(0,T;\mathscr{L}^2(U,\mathbb{R}))} \to 0$$

as $\delta \to 0$. A completely analogous argument shows that $\big\|Y^{\delta,\alpha}G^\alpha - Y^\alpha G^\alpha\big\|_{L^2(0,T;\mathscr{L}^2(U,\mathbb{R}))}$ tends to 0 as $\delta \to 0$ as well.

We are now going to show that $Y^{\delta,\alpha}\zeta^{\delta,\alpha} \to Y^\alpha\zeta^\alpha$ in $L^1(\Omega \times (0,T) \times D)$, which clearly implies that

$$\int_0^t\int_D Y^{\delta,\alpha}\zeta^{\delta,\alpha} \to \int_0^t\int_D Y^\alpha\zeta^\alpha$$

in probability for all $t \in [0,T]$. Since $Y^{\delta,\alpha} \to Y^\alpha$ and $\zeta^{\delta,\alpha} \to \zeta^\alpha$ in measure in $\Omega \times (0,T) \times D$, Vitali's theorem implies strong convergence in L^1 if the sequence $(Y^{\delta,\alpha}\zeta^{\delta,\alpha})$ is uniformly integrable in $\Omega \times (0,T) \times D$. In turn, the latter is certainly true if $\big(|Y^{\delta,\alpha}\zeta^{\delta,\alpha}|\big)$ is dominated by a sequence that converges strongly in L^1. In order to prove this property, note that j and j^* are increasing on $\mathbb{R}_+$, hence

$$\begin{aligned}
\frac{1}{4}\big|Y^{\delta,\alpha}(\omega,t,x)\zeta^{\delta,\alpha}(\omega,t,x)\big| &\leq j\big(e^{-\alpha t}|Y^\delta(\omega,t,x)|/2\big) + j^*\big(e^{-\alpha t}|\zeta^\delta(\omega,t,x)|/2\big) \\
&\leq j\big(|Y^\delta(\omega,t,x)|/2\big) + j^*\big(|\zeta^\delta(\omega,t,x)|/2\big),
\end{aligned}$$

so that, by the symmetry of j and j^*, and by the Jensen inequality of Lemma 4.2,

$$\frac{1}{4}\big|Y^{\delta,\alpha}\zeta^{\delta,\alpha}\big| \leq j(Y^\delta/2) + j^*(\zeta^\delta/2) \leq (I + \delta A)^{-m}\big(j(Y/2) + j^*(\zeta/2)\big),$$

where, by convexity and symmetry,

$$j(Y/2) = j\Big(\frac{1}{2}X_1 + \frac{1}{2}(-X_2)\Big) \leq \frac{1}{2}\big(j(X_1) + j(X_2)\big) \in L^1(\Omega \times (0,T) \times D),$$

and, completely analogously,

$$j^*(\zeta/2) \leq \frac{1}{2}\big(j^*(\xi_1) + j^*(\xi_2)\big) \in L^1(\Omega \times (0,T) \times D),$$

hence

$$\big|Y^{\delta,\alpha}\zeta^{\delta,\alpha}\big| \lesssim (I+\delta A)^{-m}\big(j(X_1)+j(X_2)+j^*(\xi_1)+j^*(\xi_2)\big).$$

Since the right-hand side of this expression converges strongly in $L^1(\Omega\times(0,T)\times D)$ as $\delta\to 0$, it is, a fortiori, uniformly integrable, and so is the left-hand side.

We have thus obtained

$$\begin{aligned}
\big\|Y^\alpha(t)\big\|_H^2 &+ 2\alpha\int_0^t \big\|Y^\alpha(s)\big\|_H^2\,ds + 2\int_0^t \mathscr{E}\big(Y^\alpha(s),Y^\alpha(s)\big)\,ds\\
&+ 2\int_0^t\!\!\int_D Y^\alpha(s,x)\zeta^\alpha(s,x)\,dx\,ds\\
&\qquad \le \big\|Y_0\big\|_H^2 + \int_0^t Y^\alpha(s)G^\alpha(s)\,dW(s) + \int_0^t \big\|G^\alpha(s)\big\|^2_{\mathscr{L}^2(U,H)}\,ds,
\end{aligned}$$

where, by monotonicity, $Y^\alpha\zeta^\alpha = e^{-2\alpha\cdot}(X_1-X_2)(\xi_2-\xi_2)\ge 0$, hence, taking the $L^\infty(0,T)$ norm and expectation on both sides,

$$\begin{aligned}
&\big\|Y^\alpha\big\|_{L^2(\Omega;L^\infty(0,T;H))} + \sqrt{\alpha}\big\|Y^\alpha\big\|_{L^2(\Omega;L^2(0,T;H))} + \big\|Y^\alpha\big\|_{L^2(\Omega;L^2(0,T;V))}\\
&\quad\lesssim \big\|Y_0\big\|_{L^2(\Omega;H)} + \Big(\mathbb{E}\sup_{t\le T}\Big|\int_0^t Y^\alpha(s)G^\alpha(s)\,dW(s)\Big|\Big)^{1/2} + \big\|G^\alpha\big\|_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))}.
\end{aligned}$$

By Lemma 4.3, one has

$$\begin{aligned}
\Big(\mathbb{E}\sup_{t\le T}\Big|\int_0^t Y^\alpha(s)G^\alpha(s)\,dW(s)\Big|\Big)^{1/2} &\le \varepsilon\big\|Y^\alpha\big\|_{L^2(\Omega;L^\infty(0,T;H))}\\
&\quad + N(\varepsilon)\big\|G^\alpha\big\|_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))},
\end{aligned}$$

with $\varepsilon>0$ arbitrary. Choosing ε sufficiently small and rearranging terms, one obtains

$$\|X_1-X_2\|_{F_\alpha} \lesssim \|X_{01}-X_{02}\|_{L^2(\Omega;H)} + \|B_1-B_2\|_{L^2(\Omega;L^2_\alpha(0,T;\mathscr{L}^2(U,H)))}$$

as claimed.

Choosing $\alpha=0$, $X_{01}=X_{02}$, and $B_1=B_2$, one gets immediately $X_1=X_2$, hence also, by substitution,

$$\int_0^t (\xi_1(s)-\xi_2(s))\,ds = 0 \qquad \forall t\in[0,T],$$

which implies uniqueness of ξ. □

7 Proof of the main result

Let $Y\in L^2(\Omega;L^2(0,T;H))$ be a progressively measurable process, $X_0\in L^2(\Omega,\mathscr{F}_0,\mathbb{P};H)$, and consider the equation

$$dX(t) + AX(t)\,dt + \beta(X(t))\,dt \ni B(t,Y(t))\,dW(t), \qquad X(0)=X_0. \tag{7.1}$$

Since $B(\cdot, Y)$ is U-measurable, adapted, and belongs to $L^2(\Omega; L^2(0, T; \mathscr{L}^2(U, H)))$, the above equation is well-posed in $\mathscr{J}$ by Proposition 6.1, hence one can define a map

$$\begin{aligned}\Gamma : L^2(\Omega; H) \times L^2(\Omega; L^2(0, T; H)) &\longrightarrow L^2(\Omega; L^2(0, T; H)) \times L^1(\Omega \times (0, T) \times D)\\ (X_0, Y) &\longmapsto (X, \xi),\end{aligned}$$

where (X, ξ) is the unique process in $\mathscr{J}$ solving (7.1). Denoting the $L^2(\Omega; L^2(0, T; H))$-valued component of Γ by Γ_1 and the $L^1(\Omega \times (0, T) \times D)$-valued component by Γ_2, we are going to show that $Y \mapsto \Gamma_1(X_0, Y)$ is a (strict) contraction of $L^2(\Omega; L^2(0, T; H))$, if endowed with a suitably chosen equivalent norm. Let $X_i = \Gamma_1(X_{0i}, Y_i)$, $i = 1, 2$, with obvious meaning of the symbols. For any $\alpha \geq 0$, Proposition 6.2 yields

$$\begin{aligned}&\|X_1 - X_2\|_{L^2(\Omega; L^\infty_\alpha(0,T;H)) \cap L^2(\Omega; L^2_\alpha(0,T;V))} + \sqrt{\alpha}\|X_1 - X_2\|_{L^2(\Omega; L^2_\alpha(0,T;H))}\\ &\qquad \lesssim \|X_{01} - X_{02}\|_{L^2(\Omega;H)} + \|B(\cdot, Y_1) - B(\cdot, Y_2)\|_{L^2(\Omega; L^2_\alpha(0,T;\mathscr{L}^2(U,H)))},\end{aligned} \tag{7.2}$$

in particular, by the Lipschitz continuity of B,

$$\begin{aligned}\|X_1 - X_2\|_{L^2(\Omega; L^2_\alpha(0,T;H))} &\lesssim \frac{1}{\sqrt{\alpha}}\|X_{01} - X_{02}\|_{L^2(\Omega;H)}\\ &\qquad + \frac{1}{\sqrt{\alpha}}\|B(\cdot, Y_1) - B(\cdot, Y_2)\|_{L^2(\Omega; L^2_\alpha(0,T;\mathscr{L}^2(U,H)))}\\ &\lesssim \frac{1}{\sqrt{\alpha}}\Big(\|X_{01} - X_{02}\|_{L^2(\Omega;H)} + \|Y_1 - Y_2\|_{L^2(\Omega; L^2_\alpha(0,T;H))}\Big),\end{aligned} \tag{7.3}$$

where the implicit constant does not depend on α. In particular, if $X_{01} = X_{02}$, choosing α large enough, one has that, for any $X_0 \in L^2(\Omega, H)$, $Y \mapsto \Gamma_1(X_0, Y)$ is a contraction of $L^2(\Omega; L^2_\alpha(0, T; H))$. It follows by the Banach fixed-point theorem that $\Gamma_1(X_0, \cdot)$ has a unique fixed point X therein, hence also in $L^2(\Omega; L^2(0, T; H))$ by equivalence of norms. Setting $\xi := \Gamma_2(X_0, X)$, by definition of the map Γ, (X, ξ) is a solution to (1.1) and it belongs to $\mathscr{J}$.

Let X_{01}, $X_{02} \in L^2(\Omega, \mathscr{F}_0; H)$ and X_1, X_2 be the unique fixed points of the maps $\Gamma_1(X_{0i}, \cdot)$, $i = 1, 2$, respectively, and $\xi_i := \Gamma_2(X_{0i}, X_i)$, $i = 1, 2$. Replacing Y_i with $X_i = \Gamma_1(X_{0i}, X_i)$, $i = 1, 2$, in (7.3) yields

$$\|X_1 - X_2\|_{L^2(\Omega; L^2_\alpha(0,T;H))} \leq C_1\|X_{01} - X_{02}\|_{L^2(\Omega;H)} + C_2\|X_1 - X_2\|_{L^2(\Omega; L^2_\alpha(0,T;H))},$$

with $C_1 > 0$, $C_2 \in]0, 1[$, hence, by equivalence of norms,

$$\|X_1 - X_2\|_{L^2(\Omega; L^2(0,T;H))} \lesssim \|X_{01} - X_{02}\|_{L^2(\Omega;H)}.$$

This implies, substituting Y_i with $X_i = \Gamma(X_{0i}, X_i)$, $i = 1, 2$, in (7.2), with $\alpha = 0$,

$$\begin{aligned}&\|X_1 - X_2\|_{L^2(\Omega; L^\infty(0,T;H)) \cap L^2(\Omega; L^2(0,T;V))}\\ &\qquad \lesssim \|X_{01} - X_{02}\|_{L^2(\Omega;H)} + \|B(\cdot, X_1) - B(\cdot, X_2)\|_{L^2(\Omega; L^2(0,T;\mathscr{L}^2(U,H)))}\\ &\qquad \lesssim \|X_{01} - X_{02}\|_{L^2(\Omega;H)} + \|X_1 - X_2\|_{L^2(\Omega; L^2(0,T;H))}\\ &\qquad \lesssim \|X_{01} - X_{02}\|_{L^2(\Omega;H)}.\end{aligned}$$

Choosing $\alpha = 0$ and $X_{01} = X_{02}$, one gets immediately $X_1 = X_2$, hence also, by substitution,

$$\int_0^t (\xi_1(s) - \xi_2(s))\, ds = 0 \qquad \forall t \in [0, T],$$

which implies uniqueness of ξ.

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] W. Arendt, *Heat kernels*, Lecture Notes of the 9th Internet Seminar on Evolution Equations, 2006, available at `https://www.uni-ulm.de/mawi/iaa/members/arendt/`.

[3] W. Arendt and A. V. Bukhvalov, *Integral representations of resolvents and semigroups*, Forum Math. **6** (1994), no. 1, 111–135. MR 1253180

[4] W. Arendt, R. Chill, C. Seifert, D. Vogt, and J. Voigt, *Form methods for evolution equations and applications*, Lecture Notes of the 18th Internet Seminar on Evolution Equations, 2015, available at `https://www.mat.tuhh.de/isem18/Phase_1:_The_lectures`.

[5] V. Barbu, *Analysis and control of nonlinear infinite-dimensional systems*, Academic Press Inc., Boston, MA, 1993. MR MR1195128 (93j:49002)

[6] ———, *Existence for semilinear parabolic stochastic equations*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. **21** (2010), no. 4, 397–403. MR 2746091 (2012d:35424)

[7] V. Barbu, G. Da Prato, and M. Röckner, *Existence of strong solutions for stochastic porous media equation under general monotonicity conditions*, Ann. Probab. **37** (2009), no. 2, 428–452. MR MR2510012

[8] V. Barbu and C. Marinelli, *Strong solutions for stochastic porous media equations with jumps*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **12** (2009), no. 3, 413–426. MR 2572464

[9] A. Bendikov and P. Maheux, *Nash type inequalities for fractional powers of non-negative self-adjoint operators*, Trans. Amer. Math. Soc. **359** (2007), no. 7, 3085–3097 (electronic). MR 2299447

[10] L. Boccardo and G. Croce, *Elliptic partial differential equations*, De Gruyter, Berlin, 2014. MR 3154599

[11] N. Bourbaki, *Espaces vectoriels topologiques. Chapitres 1 à 5*, new ed., Masson, Paris, 1981. MR 633754

[12] H. Brézis, *Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations*, Contributions to nonlinear functional analysis (Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1971), Academic Press, New York, 1971, pp. 101–156. MR MR0394323 (52 #15126)

[13] ______, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Publishing Co., Amsterdam, 1973. MR MR0348562 (50 #1060)

[14] Z. Brzeźniak, B. Maslowski, and J. Seidler, *Stochastic nonlinear beam equations*, Probab. Theory Related Fields **132** (2005), no. 1, 119–149. MR 2136869

[15] S. Cerrai, *Stochastic reaction-diffusion systems with multiplicative noise and non-Lipschitz reaction term*, Probab. Theory Related Fields **125** (2003), no. 2, 271–304. MR 1961346 (2004a:60117)

[16] G. Da Prato, *Kolmogorov equations for stochastic PDEs*, Birkhäuser Verlag, Basel, 2004. MR MR2111320 (2005m:60002)

[17] R. Dautray and J.-L. Lions, *Mathematical analysis and numerical methods for science and technology. Vol. 2*, Springer-Verlag, Berlin, 1988. MR 969367

[18] E. B. Davies, *Heat kernels and spectral theory*, Cambridge University Press, Cambridge, 1990. MR MR1103113 (92a:35035)

[19] K.-J. Engel and R. Nagel, *One-parameter semigroups for linear evolution equations*, Springer-Verlag, New York, 2000. MR MR1721989 (2000i:47075)

[20] I. Gentil and C. Imbert, *The Lévy-Fokker-Planck equation: Φ-entropies and convergence to equilibrium*, Asymptot. Anal. **59** (2008), no. 3-4, 125–138. MR 2450356

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

[22] J.-B. Hiriart-Urruty and C. Lemaréchal, *Fundamentals of convex analysis*, Springer-Verlag, Berlin, 2001. MR 1865628 (2002i:90002)

[23] T. Kato, *Perturbation theory for linear operators*, Springer-Verlag, Berlin, 1995, Reprint of the 1980 edition. MR 1335452

[24] 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)

[25] M. Kunze and J. van Neerven, *Continuous dependence on the coefficients and global existence for stochastic reaction diffusion equations*, J. Differential Equations **253** (2012), no. 3, 1036–1068. MR 2922662

[26] Se. Kusuoka and C. Marinelli, *On smoothing properties of transition semigroups associated to a class of SDEs with jumps*, Ann. Inst. Henri Poincaré Probab. Stat. **50** (2014), no. 4, 1347–1370. MR 3269997

[27] J.-L. Lions and E. Magenes, *Problèmes aux limites non homogènes et applications. Vol. 1*, Dunod, Paris, 1968. MR 0247243

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

[29] Zhi Ming Ma and M. Röckner, *Introduction to the theory of (nonsymmetric) Dirichlet forms*, Springer-Verlag, Berlin, 1992. MR 1214375

[30] C. Marinelli, *On well-posedness of semilinear stochastic evolution equations on L_p spaces*, arXiv:1512.04323.

[31] C. Marinelli and Ll. Quer-Sardanyons, *Existence of weak solutions for a class of semilinear stochastic wave equations*, SIAM J. Math. Anal. **44** (2012), no. 2, 906–925. MR 2914254

[32] E. M. Ouhabaz, *Analysis of heat equations on domains*, Princeton University Press, Princeton, NJ, 2005. MR 2124040

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

[34] É. Pardoux and A. Răşcanu, *Stochastic differential equations, backward SDEs, partial differential equations*, Springer, Cham, 2014. MR 3308895

[35] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Springer-Verlag, New York, 1983. MR 85g:47061

[36] 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)

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

[38] J. van Neerven, M. C. Veraar, and L. Weis, *Stochastic evolution equations in UMD Banach spaces*, J. Funct. Anal. **255** (2008), no. 4, 940–993. MR 2433958 (2009h:35465)

[39] N. Th. Varopoulos, L. Saloff-Coste, and T. Coulhon, *Analysis and geometry on groups*, Cambridge University Press, Cambridge, 1992. MR 1218884