



**HAL**  
open science

## Existence of weak solutions to stochastic evolution inclusions

Adam Jakubowski, Mikhail Kamenskii, Paul Raynaud de Fitte

► **To cite this version:**

Adam Jakubowski, Mikhail Kamenskii, Paul Raynaud de Fitte. Existence of weak solutions to stochastic evolution inclusions. *Stochastic Analysis and Applications*, 2005, 23 (4), pp.723-749. 10.1081/SAP-200064469 . hal-00002365v2

**HAL Id: hal-00002365**

**<https://hal.science/hal-00002365v2>**

Submitted on 8 Aug 2004

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Existence of weak solutions to stochastic evolution inclusions

Adam JAKUBOWSKI\*  
 Mikhail KAMENSKIĬ†  
 Paul RAYNAUD DE FITTE‡

August 6, 2004

## Abstract

We prove the existence of a weak mild solution to the Cauchy problem for the semilinear stochastic differential inclusion in a Hilbert space

$$dX_t \in AX_t dt + F(t, X_t) dt + G(t, X_t) dW_t$$

where  $W$  is a Wiener process,  $A$  is a linear operator which generates a  $C_0$ -semigroup,  $F$  and  $G$  are multifunctions with convex compact values satisfying some growth condition and, with respect to the second variable, a condition weaker than the Lipschitz condition. The weak solution is constructed in the sense of Young measures.

## 1 Introduction

Ordinary and stochastic differential and inclusions in infinite dimensional spaces have many important and interesting applications on which we do not stop here, addressing the reader to the books and to the reviews (see for example [1, 2, 3]).

For ordinary differential equations, after the famous example of Dieudonné [4] and the fundamental results of A. Godunov [5, 6], it became evident that, in the case of infinite dimensional Banach spaces, it is necessary to suppose an auxiliary condition on the right hand side of the equation

$$(1) \quad x' = f(t, x)$$

in order to have an existence theorem for the initial value problem with the initial condition

$$(2) \quad x(0) = x_0.$$

It is well known that the estimation

$$(3) \quad \|f(t, x) - f(t, y)\| \leq L(t, \|x - y\|),$$

\*Nicholas Copernicus University, Faculty of Mathematics and Informatics, ul. Chopina 12/18, 87–100 Toruń, Poland. E-Mail: adjakubo@mat.uni.torun.pl

†Departement of Mathematics, State University of Voronezh, Voronezh, Universitetskaja pl. 1, 394693, Russia. E-Mail: mikhaikamenski@mail.ru

‡Laboratoire Raphaël Salem, UMR CNRS 6085, UFR Sciences, Université de Rouen, 76821 Mont Saint Aignan Cedex, France. E-Mail: prf@univ-rouen.fr

where  $L$  is a real function such that the integral inequality

$$u(t) \leq \int_0^t L(s, u(s)) ds,$$

has a unique solution  $u(t) = 0$ , gives the existence condition mentioned below. In [7] it was shown that it is possible to add to  $f$  satisfying (3) a continuous compact operator. As generalisations of this fact many papers in the seventies were devoted to a condition of the form

$$(4) \quad \varphi(f(t, \Lambda)) \leq L(s, \varphi(\Lambda)),$$

where  $\varphi$  is a measure of noncompactness (see e.g. [8]). The abstract fixed point theorem for condensing operators (see [8]) was successfully applied in this way.

In the end of the XXth century, in many papers (see e.g. [9, 10, 11]), it was shown that Condition (4) implies the existence of solutions to the semilinear equation

$$x' = Ax + f(t, x)$$

and the semilinear inclusion (e.g. [12], see also the references in [13])

$$(5) \quad x' \in Ax + f(t, x),$$

with the initial condition (2) and a linear operator  $A$  generating a  $C_0$ -semigroup. In the inclusion (5),  $f$  is a multivalued map with convex compact values.

In the same period from the seventies, it was remarked that the existence of a weak solution and the condition of unicity of trajectories implies the existence of a strong solution to the stochastic differential equation

$$(6) \quad dX_t = a(t, X_t) dt + b(t, X_t) dW_t,$$

where  $W$  is a standard Wiener process (here, “weak” and “strong” are taken in the probabilistic sense; the existence of a weak solution with only continuous  $a$  and  $b$  was known since Skorokhod [14]). The direct proof of the existence of a strong solution for (6) using a convenient measure of noncompactness when  $a$  and  $b$  satisfy a condition like (3) was presented in [15], see also [8].

The passage from the finite dimensional case to the infinite dimensional case with  $a$  and  $b$  satisfying a Lipschitz condition is presented in [2, 3] see there the bibliography. The generalisation to the case of a semilinear stochastic differential equation

$$dX_t = AX_t dt + f(t, X_t)dt + \sigma(t, X_t) dW_t,$$

deals with the estimations for the stochastic convolution operator

$$\int_0^t e^{(t-s)A} v(s) dW_s$$

(see [16, 17, 2, 18, 19]) and for Lipschitz  $f$  and  $\sigma$  is presented in [2].

In the work of Da Prato and Frankowska [20], the existence result is proved for the semilinear stochastic inclusion

$$(7) \quad dX_t \in AX_t dt + f(t, X_t) dt + \sigma(t, X_t) dW_t,$$

with multivalued  $f$  and  $\sigma$  which are Lipschitz with respect to the Hausdorff metric.

In the present paper we aim to prove the existence of a “weak” solution (or a “solution measure”) for (7) in the case where  $f$  and  $\sigma$  are compact valued multifunctions satisfying a condition like (3) where the norm in the left hand side is replaced by the Hausdorff metric. More precisely, we prove the existence of a “weak” mild solution  $X$  to the Cauchy problem

$$(8) \quad \begin{cases} dX_t \in AX_t + F(t, X_t) dt + G(t, X_t) dW(t) \\ X(0) = \xi \end{cases}$$

where  $X$  takes its values in a Hilbert space  $\mathbb{H}$ ,  $W$  is a Brownian motion on a Hilbert space  $\mathbb{U}$ ,  $A$  is a linear operator on  $\mathbb{H}$  and  $F$  and  $G$  are multivalued mappings with compact convex values, continuous in the second variable, and satisfy an assumption which is much more general than the usual Lipschitz one.

Instead of constructing the weak solution with the help of Skorokhod’s representation theorem, we define our weak solution as a Young measure (ie as a *solution measure* in the sense of Jacod and Mémmin [21]).

As in the case of the ordinary differential equation (1) with the right hand side satisfying (4) (see [22]) we apply the Tonelli scheme, and prove that its solutions are uniformly tight. We then pass to the limit and we obtain the existence of a weak mild solution to the initial problem.

To get rid of the Lipschitz assumption on  $F$  and  $G$  has a cost: not only do we obtain “weak” solutions, but our techniques based on compactness lead us to consider mappings  $F$  and  $G$  with convex compact values, whereas the more geometric methods of [20] deal with unbounded closed valued mappings. Furthermore, in our work, the multifunctions  $F$  and  $G$  are deterministic, whereas they are random in [20].

## 2 Formulation of the problem, statement of the result

### Notations

Throughout,  $0 < T < +\infty$  is a fixed time and  $\mathfrak{F} = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$  is a stochastic basis satisfying the usual conditions. The  $\sigma$ -algebra of predictable subsets of  $[0, T] \times \Omega$  is denoted by  $\mathcal{P}$ .

If  $\mathbb{E}$  is a separable metric space, we denote  $\mathcal{B}_{\mathbb{E}}$  its Borel  $\sigma$ -algebra, and by  $\mathcal{M}^{1,+}(\mathbb{E})$  the space of probability laws on  $(\mathbb{E}, \mathcal{B}_{\mathbb{E}})$ , endowed with the usual narrow (or weak) topology. The law of a random element  $X$  of  $\mathbb{E}$  is denoted by  $\mathcal{L}(X)$ . The space of random elements of  $\mathbb{E}$  defined on  $(\Omega, \mathcal{F})$  is denoted by  $L^0(\Omega; \mathbb{E})$  or simply  $L^0(\mathbb{E})$ , and it is endowed with the topology of convergence in  $\mathbb{P}$ -probability. We identify random elements which are equal  $\mathbb{P}$ -a.e. Recall that a subset  $\Lambda$  of  $L^0(\mathbb{E})$  is said to be tight if, for any  $\delta > 0$ , there exists a compact subset  $\mathcal{K}$  of  $\mathbb{E}$  such that, for every  $X \in \Lambda$ ,  $\mathbb{P}(X \in \mathcal{K}) \geq 1 - \delta$ .

### Spaces of processes

In all this paper,  $p > 2$  is a fixed number. For any  $t \in [0, T]$ , we denote by  $\mathcal{N}_c^p(\mathfrak{F}, [0, t]; \mathbb{H})$  the space of continuous  $(\mathcal{F}_t)$ -adapted  $\mathbb{H}$ -valued processes  $X$  such that

$$\|X\|_{\mathcal{N}_c^p(\mathfrak{F}, [0, t]; \mathbb{H})}^p := \mathbf{E} \left( \sup_{0 \leq s \leq t} \|X(s)\|_{\mathbb{H}}^p \right) < +\infty.$$

For any  $t \in [0, T]$  we denote by  $C([0, t]; \mathbb{H})$  the space of continuous mappings from  $[0, t]$  to  $\mathbb{H}$ . The space  $C([0, t]; \mathbb{H})$  is endowed with the topology of uniform convergence defined by the distance  $d_\infty(u, v) = \sup_{s \in [0, t]} \|u(s) - v(s)\|$ . The space  $\mathcal{N}_c^p(\mathfrak{F}, [0, t]; \mathbb{H})$  is thus a closed separable subspace of the space  $L^0(C([0, t]; \mathbb{H}))$  of  $C([0, t]; \mathbb{H})$ -valued random variables.

## Measures of noncompactness

If  $(\mathbb{E}, d)$  is a metric space and  $\Lambda \subset \mathbb{E}$ , we say that a subset  $\Lambda'$  of  $\mathbb{E}$  is an  $\epsilon$ -net of  $\Lambda$  if

$$\inf_{X \in \Lambda} d(X, \Lambda') \leq \epsilon$$

(note that  $\Lambda'$  is not necessarily a subset of  $\Lambda$ ). Let  $\Lambda$  be a subset of  $\mathcal{N}_c^p(\mathfrak{F}, [0, T]; \mathbb{H})$ . For any  $s \in [0, T]$ , we denote by  $\Lambda|_s$  the subset of  $\mathcal{N}_c^p(\mathfrak{F}, [0, s]; \mathbb{H})$  of restrictions to  $[0, s]$  of elements of  $\Lambda$ . We denote

$$\Psi(\Lambda)(s) := \inf \left\{ \epsilon > 0; \Lambda|_s \text{ has a tight } \epsilon\text{-net in } \mathcal{N}_c^p(\mathfrak{F}, [0, s]; \mathbb{H}) \right\}.$$

We shall see in Lemma 3.1 that  $\Lambda$  is tight if and only if  $\Psi(\Lambda)(T) = 0$ . Note also that the mapping  $s \mapsto \Psi(\Lambda)(s)$  is nondecreasing. We denote

$$\Psi(\Lambda) := (\Psi(X)(s))_{0 \leq s \leq T}.$$

The family  $(\Psi(X)(s))_{0 \leq s \leq t}$  is called the *measure of noncompactness* of  $\Lambda$  (see [8, 23] about measures of noncompactness).

## Spaces of closed compact subsets

For any metric space  $\mathbb{E}$ , the set of nonempty compact (resp. nonempty compact convex) subsets of  $\mathbb{E}$  is denoted by  $\mathfrak{K}(\mathbb{E})$  (resp.  $\mathfrak{K}_c(\mathbb{E})$ ). We endow  $\mathfrak{K}(\mathbb{E})$  and its subspace  $\mathfrak{K}_c(\mathbb{E})$  with the Hausdorff distance

$$\text{Hausd}_{\mathbb{E}}(B, C) := \max \left( \inf_{b \in B} d(b, C), \inf_{c \in C} d(c, B) \right).$$

Recall that, if  $\mathbb{E}$  is Polish, then  $\mathfrak{K}_c(\mathbb{E})$  is Polish too (see *e.g.* [24, Corollary II-9]).

## Stable convergence and Young measures

Let  $\mathbb{E}$  be a complete metric space. We denote by  $\mathcal{Y}(\Omega, \mathcal{F}, P; \mathbb{E})$  (or simply  $\mathcal{Y}(\mathbb{E})$ ) the set of measurable mappings

$$\mu : \begin{cases} \Omega & \mapsto \mathcal{M}^{1,+}(\mathbb{E}) \\ \omega & \mapsto \mu_\omega \end{cases}$$

Each element  $\mu$  of  $\mathcal{Y}(\Omega, \mathcal{F}, P; \mathbb{E})$  can be identified with the measure  $\tilde{\mu}$  on  $(\Omega \times \mathbb{E}, \mathcal{F} \otimes \mathcal{B}_{\mathbb{E}})$  defined by  $\tilde{\mu}(A \times B) = \int_A \mu_\omega(B) dP(\omega)$  (and the mapping  $\mu \mapsto \tilde{\mu}$  is onto, see *e.g.* [25]). In the sequel, we shall use freely this identification. For instance, if  $f : \Omega \times \mathbb{E} \rightarrow \mathbb{R}$  is a bounded measurable mapping, the notation  $\mu(f)$  denotes  $\int_\Omega \mu_\omega(f(\omega, \cdot)) dP(\omega)$ . The elements of  $\mathcal{Y}(\Omega, \mathcal{F}, P; \mathbb{E})$  are called *Young measures* on  $\Omega \times \mathbb{E}$ .

Let  $\mathcal{C}_b(\mathbb{E})$  denote the set of continuous bounded real valued functions defined on  $\mathbb{E}$ . The set  $\mathcal{Y}(\Omega, \mathcal{F}, P; \mathbb{E})$  is endowed with a metrizable topology, such that a sequence  $(\mu^n)$  of

Young measures converges to a Young measure  $\mu$  if, for each  $A \in \mathcal{F}$  and each  $f \in \mathcal{C}_b(\mathbb{E})$ , the sequence  $(\mu^n(\mathbf{1}_A \otimes f))$  converges to  $\mu(\mathbf{1}_A \otimes f)$  (where  $\mathbf{1}_A$  is the indicator function of  $A$ ). We then say that  $(\mu^n)$  converges *stably* to  $\mu$ .

Each element  $X$  of  $L^0(\Omega; \mathbb{E})$  can (and will sometimes) be identified with the Young measure

$$\underline{\delta}_X : \omega \mapsto \delta_{X(\omega)},$$

where, for any  $x \in \mathbb{E}$ ,  $\delta_x$  denotes the probability concentrated on  $x$ . Note that the restriction to  $L^0(\Omega; \mathbb{E})$  of the topology of stable convergence is the topology of convergence in probability. If  $(X_n)$  is a tight sequence in  $L^0(\Omega; \mathbb{E})$ , then, by Prohorov's compactness criterion for Young measures [26, 27], each subsequence of  $(X_n)$  has a further subsequence, say  $(X'_n)$ , which converges stably to some  $\mu \in \mathcal{Y}(\mathbb{E})$ , that is, for every  $A \in \mathcal{F}$  and every  $f \in \mathcal{C}_b(\mathbb{E})$ ,

$$\lim_n \int_A f(X'_n) dP = \mu(\mathbf{1}_A \otimes f).$$

This entails in particular that  $(X'_n)$  converges in distribution to the measure  $\mu(\Omega \times \cdot) \in \mathcal{M}^{1,+}(\mathbb{E})$ .

See [28] or [29] for an introduction to Young measures and their applications.

## Hypothesis

In the sequel, we are given two separable Hilbert spaces  $\mathbb{H}$  and  $\mathbb{U}$  and an  $(\mathcal{F}_t)_{t \in [0, T]}$ -Brownian motion  $W$  (possibly cylindrical) on  $\mathbb{U}$ . We denote by  $\mathbb{L}$  the space of Hilbert-Schmidt operators from  $\mathbb{U}$  to  $\mathbb{H}$ .

We shall consider the following hypothesis:

(HS)  $A$  is the generator of a  $\mathcal{C}_0$  semigroup  $(S(t))_{t \geq 0}$ . In particular (see *e.g.* [1, Theorem 1.3.1]), there exist  $M > 0$  and  $\beta \in ]-\infty, +\infty[$  such that, for every  $t \geq 0$ ,

$$\|S(t)\| \leq M e^{\beta t}.$$

For  $t \in [0, T]$ , we denote  $M_t = \sup_{0 \leq s \leq t} M e^{\beta s}$ .

(HFG)  $F : \Omega \times [0, T] \times \mathbb{H} \rightarrow \mathfrak{K}_c(\mathbb{H})$  and  $G : \Omega \times [0, T] \times \mathbb{H} \rightarrow \mathfrak{K}_c(\mathbb{L})$  are measurable mappings which satisfy the following conditions:

(i) There exists a constant  $C_{\text{growth}} > 0$  such that, for all  $(t, x) \in [0, T] \times \mathbb{H}$ ,

$$\begin{aligned} \text{Hausd}_{\mathbb{H}}(0, F(t, x)) &\leq C_{\text{growth}}(1 + \|x\|) \\ \text{Hausd}_{\mathbb{L}}(0, G(t, x)) &\leq C_{\text{growth}}(1 + \|x\|). \end{aligned}$$

(ii) For all  $(t, x, y) \in [0, T] \times \mathbb{H} \times \mathbb{H}$ ,

$$\begin{aligned} (\text{Hausd}_{\mathbb{H}}(F(t, x), F(t, y)))^p &\leq L(t, \|x - y\|^p) \\ (\text{Hausd}_{\mathbb{L}}(G(t, x), G(t, y)))^p &\leq L(t, \|x - y\|^p), \end{aligned}$$

where  $L : [0, T] \times [0, +\infty] \rightarrow [0, +\infty]$  is a given continuous mapping such that

(a) for every  $t \in [0, T]$ , the mapping  $L(t, \cdot)$  is nondecreasing and convex,

(b) for every measurable mapping  $z : [0, T] \rightarrow [0, +\infty]$  and for every constant  $K > 0$ , the following implication holds true:

$$(9) \quad \left[ (\forall t \in [0, T]) \ z(t) \leq K \int_0^t L(s, z(s)) ds \right] \Rightarrow z = 0.$$

In particular, we have  $L(0) = 0$ , thus Hypothesis (HFG)-(ii) entails that, for each  $t \in [0, T]$ , the mappings  $F(t, \cdot)$  and  $G(t, \cdot)$  are continuous for the Hausdorff distances  $\text{Hausd}_{\mathbb{H}}$  and  $\text{Hausd}_{\mathbb{L}}$  respectively. Such a function  $L$  is considered in e.g. [30, 15, 31, 8, 32, 33, 34]. Concrete examples can be found in [35, Section 6 of Chapter 3].

$$(HI) \quad \xi \in L^p(\Omega, \mathcal{F}_0, \mathbb{P}|_{\mathcal{F}_0}; \mathbb{H}).$$

Recall that, under Hypothesis (HS), there exists a constant  $C_{\text{Conv}}$  such that, for any predictable process  $Z \in L^p(\Omega \times [0, T]; \mathbb{L})$ , we have

$$(10) \quad \mathbf{E} \left[ \sup_{s \leq t} \left\| \int_0^s S(s-r)Z(r) dW(r) \right\|^p \right] \leq C_{\text{Conv}} t^{(p/2)-1} \mathbf{E} \int_0^t \|Z(s)\|_{\mathbb{L}}^p ds$$

(see [18, 2], and also [19] for a strikingly short proof in the case when  $S$  is contractive).

### Weak and strong mild solutions

We say that  $X \in \mathcal{N}_c^p(\mathfrak{F}, [0, T]; \mathbb{H})$  is a (*strong*) *mild solution* to Equation (8) if there exist two predictable processes  $f$  and  $g$  defined on  $\mathfrak{F}$  satisfying

$$(11) \quad \begin{cases} X(t) = S(t)\xi + \int_0^t S(t-s)f(s) ds + \int_0^t S(t-s)g(s) dW(s) \\ f(s) \in F(s, X(s)) \text{ P-a.e.} \\ g(s) \in G(s, X(s)) \text{ P-a.e.} \end{cases}$$

So, “mild solution” refers to the variation of constant formula, whereas “strong” refers to the fact that the solution is defined on the given stochastic basis.

We say that a process  $X$  is a *weak mild solution* or a *mild solution-measure* to (8) if there exists a stochastic basis  $\underline{\mathfrak{F}} = (\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_t, \mu)$  satisfying the following conditions:

1.  $\underline{\Omega}$  has the form  $\underline{\Omega} = \Omega \times \Omega'$ ,  $\underline{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}'$  for some  $\sigma$ -algebra  $\mathcal{F}'$  on  $\Omega'$ ,  $\underline{\mathcal{F}}_t = \mathcal{F}_t \otimes \mathcal{F}'_t$  for some right continuous filtration  $(\mathcal{F}'_t)$  on  $(\Omega', \mathcal{F}')$ , and the probability  $\mu$  satisfies  $\mu(A \times \Omega') = \mathbb{P}(A)$  for every  $A \in \mathcal{F}$ .
2. The process  $W$  is a Brownian motion on  $\underline{\mathfrak{F}}$  (we identify here every random variable  $X$  on  $\Omega$  with the random variable  $(\omega, \omega') \mapsto X(\omega)$  defined on  $\underline{\Omega}$ ).
3.  $X \in \mathcal{N}_c^p(\underline{\mathfrak{F}}, [0, T]; \mathbb{H})$  and there exist two predictable processes  $f$  and  $g$  defined on  $\underline{\mathfrak{F}}$  satisfying (11).

The terminology *solution-measure* is that of [21]. If  $\mathcal{F}'$  is the Borel  $\sigma$ -algebra of some topology on  $\Omega'$ , a solution-measure can also be seen as a Young measure. This is the point of view adopted by Pellaumail [36, 37], who calls Young measures *rules*.

## Main result and corollaries

We can now state the main result of this paper. The proofs will be given in Section 4.

**Theorem 2.1 (Main result)** *Under Hypothesis (HS), (HFG) and (HI), Equation (8) has a weak mild solution.*

An easy adaptation of our reasoning also yields, as a by-product, a well-known strong existence result:

**Proposition 2.2 (Strong existence in the single valued case [33, 34])** *Under Hypothesis (HS), (HFG) and (HI), if furthermore  $F$  and  $G$  are single-valued, then (8) has a strong mild solution.*

Using the Steiner point for the choice of selections of  $F$  and  $G$ , we deduce the following

**Proposition 2.3 (Strong existence in the finite dimensional case)** *Under Hypothesis (HS), (HFG) and (HI), if furthermore  $\mathbb{H}$  and  $\mathbb{U}$  are finite dimensional, then (8) has a strong mild solution.*

## 3 Preliminary results

### Tightness results and boundedness results

We start with a very simple and useful lemma.

**Lemma 3.1 (A tightness criterion)** *Let  $(\mathbb{E}, d)$  be a separable complete metric space. Let  $\Lambda$  be a set of random elements of  $\mathbb{E}$  defined on  $(\Omega, \mathcal{F}, \mathbb{P})$ . Let  $r \geq 1$ . Assume that, for every  $\epsilon > 0$ , there exists a tight subset  $\Lambda_\epsilon$  such that*

$$\Lambda \subset \Lambda_\epsilon^{[\epsilon]} := \{X \in L^0(\mathbb{E}); \exists Y \in \Lambda_\epsilon, \mathbf{E} d^r(X, Y) < \epsilon\}.$$

*Then  $\Lambda$  is tight.*

**Proof.** From Jensen inequality, we only need to prove Lemma 3.1 for  $p = 1$ . Indeed, we have  $\mathbf{E} d(X, Y) \leq (\mathbf{E} d^r(X, Y))^{1/r}$  thus

$$\Lambda \subset \{X \in L^0(\mathbb{E}); \exists Y \in \Lambda_\epsilon, \mathbf{E} d(X, Y) < \epsilon^{1/r}\}.$$

Let  $\text{BL}_1(d)$  be the set of all mappings  $f : \mathbb{E} \rightarrow [0, 1]$  which are 1-Lipschitz with respect to  $d$ . Let  $\beta$  denote the Dudley distance on  $\mathcal{M}^{1,+}(\mathbb{E})$ , that is, for all  $\mu, \nu \in \mathcal{M}^{1,+}(\mathbb{E})$ ,

$$\beta(\mu, \nu) = \sup_{f \in \text{BL}_1(d)} \mu(f) - \nu(f).$$

It is well known that the narrow topology on  $\mathcal{M}^{1,+}(\mathbb{E})$  is induced by  $\beta$  and that  $\beta$  is complete, see e.g. [38]. For every  $X \in \Lambda$ , there exists  $Y \in \Lambda_\epsilon$  such that  $\mathbf{E} d(X, Y) < \epsilon$ , which implies  $\beta(\mathcal{L}(X), \mathcal{L}(Y)) < \epsilon$ . We thus have

$$\{\mathcal{L}(X); X \in \Lambda\} \subset \bigcap_{\epsilon > 0} \{\mathcal{L}(X); X \in \Lambda_\epsilon\}^{[\epsilon]},$$

where, for any  $\Xi \subset \mathcal{M}^{1,+}(\mathbb{E})$  and any  $\epsilon > 0$ ,  $\Xi^{[\epsilon]} = \{\mu \in \mathcal{M}^{1,+}(\mathbb{E}); \exists \nu \in \Xi, \beta(\mu, \nu) < \epsilon\}$ . This proves that  $\{\mathcal{M}^{1,+}(X); X \in \Lambda\}$  is totally bounded for  $\beta$ , thus relatively compact in the narrow topology.  $\square$



**Lemma 3.2** *Let  $\Lambda$  be a set of continuous adapted processes on  $\mathbb{H}$ . Assume that each element of  $\Lambda$  is in  $L^p(\Omega \times [0, T]; \mathbb{H})$  and that  $\Lambda$ , considered as a set of  $C([0, T]; \mathbb{H})$ -valued random variables, is tight. Assume furthermore that  $G$  satisfies Hypothesis (HFG)-(i). Let  $\Xi$  be the set of processes  $Y$  of the form*

$$Y(t) = \int_0^t S(t-s)g(s) dW(s)$$

where  $g$  is  $\mathcal{P}$ -measurable and  $g(s) \in G(s, X(s))$  a.e. for some  $X \in \Lambda$ . The set  $\Xi$  is a tight set of  $C([0, T]; \mathbb{H})$ -valued random variables.

**Proof.** We will prove Lemma 3.2 through a series of reductions.

*First step:* We can assume without loss of generality that there exists a compact subset  $\mathcal{Q}$  of  $C([0, T]; \mathbb{H})$  such that, for each  $X \in \Lambda$  and for each  $\omega \in \Omega$ ,  $X(\omega, \cdot) \in \mathcal{Q}$ .

Indeed, assume that Lemma 3.2 is true under this additional hypothesis. Let  $\epsilon > 0$ . There exists a compact subset  $\mathcal{Q}_\epsilon$  of  $C([0, T]; \mathbb{H})$  such that, for each  $X \in \Lambda$ ,  $P(X \in \mathcal{Q}_\epsilon) \geq 1 - \epsilon$ . For every  $X \in \Lambda$ , there exists a measurable subset  $\Omega_X$  of  $\Omega$  such that  $P(\Omega_X) \geq 1 - \epsilon$  and, for every  $\omega \in \Omega_X$ ,  $X(\omega) \in \mathcal{Q}_\epsilon$ . For each  $X \in \Lambda$  and each  $t \in [0, T]$ , let us denote

$$X^\epsilon(t) = \begin{cases} X(t \wedge \tau) & \text{if } \tau < +\infty \\ X(0) & \text{if } \tau = +\infty, \end{cases}$$

where  $\tau(\omega)$  is the infimum of all  $s \in [0, T]$  such that there exists  $u \in \mathcal{Q}$  which coincides with  $X(\omega, \cdot)$  on  $[0, s]$  (we take  $\inf \emptyset = +\infty$ ). The process  $X^\epsilon$  is continuous and adapted, thus predictable. Furthermore, for every  $\omega \in \Omega_X$ , we have  $X(\omega, \cdot) = X^\epsilon(\omega, \cdot)$ .

Let  $\Lambda^\epsilon = \{X^\epsilon; X \in \Lambda\}$ . The set  $\Lambda^\epsilon$  is tight, thus, from our hypothesis, the set  $\Xi^\epsilon$ , obtained by replacing  $\Lambda$  by  $\Lambda^\epsilon$  in the definition of  $\Xi$ , is tight. There exists a compact subset  $\mathcal{Z}_\epsilon$  of  $C([0, T]; \mathbb{H})$  such that, for each  $Y^\epsilon \in \Xi^\epsilon$ ,  $P(Y^\epsilon \in \mathcal{Z}_\epsilon) \geq 1 - \epsilon$ . Let  $Y \in \Xi$ . The process  $Y$  has the form

$$Y(t) = \int_0^t S(t-s)g(s) dW(s)$$

where  $g$  is  $\mathcal{P}$ -measurable and  $g(s) \in G(s, X(s))$  a.e. for some  $X \in \Lambda$ . The set  $\{X \neq X^\epsilon\}$  is predictable, thus there exists a predictable process  $g^\epsilon$  such that  $g^\epsilon(s) \in G(s, X^\epsilon(s))$  a.e. and  $g^\epsilon(t) = g(t)$  for  $t \leq \tau$  (we can construct  $g^\epsilon$  as a selection of the predictable multifunction  $H$  defined by  $H(t) = \{g(t)\}$  if  $t \leq \tau$  and  $H(t) = G(t, X^\epsilon(t))$  otherwise). Let  $Y^\epsilon \in \Xi^\epsilon$  be defined by

$$Y^\epsilon(t) = \int_0^t S(t-s)g^\epsilon(s) dW(s).$$

We have

$$Y(t) - Y^\epsilon(t) = \int_{t \wedge \tau}^t S(t-s)(g(s) - g^\epsilon(s)) dW(s)$$

We thus have  $Y = Y^\epsilon$  on  $\Omega_X$ . This shows that, for every  $Y \in \Xi$ , we have

$$\begin{aligned} P(Y \notin \mathcal{Z}_\epsilon) &= P(Y = Y^\epsilon \text{ and } Y^\epsilon \notin \mathcal{Z}_\epsilon) + P(Y \neq Y^\epsilon \text{ and } Y \notin \mathcal{Z}_\epsilon) \\ &\leq P(Y^\epsilon \notin \mathcal{Z}_\epsilon) + P(Y^\epsilon \neq Y) \leq 2\epsilon. \end{aligned}$$

Thus  $\Xi$  is tight.

*Second step:* We can furthermore assume without loss of generality that there exists a compact subset  $\mathcal{K}$  of  $\mathbb{L}$  such that  $G(t, X(\omega, t)) \subset \mathcal{K}$  a.e. for all  $X \in \Lambda$ .

Assume that Lemma 3.2 holds under this hypothesis.

For any adapted process  $g \in L^p(\Omega \times [0, T]; \mathbb{L})$ , we denote by  $Y^{(g)}$  the process defined by

$$Y^{(g)}(t) = \int_0^t S(t-s)g(\omega, s) dW(s).$$

We denote by  $\Gamma$  the set of  $\mathcal{P}$ -measurable  $\mathbb{L}$ -valued processes  $g$  such that  $g(\omega, t) \in G(s, X(t))$  for every  $t \in [0, T]$  a.e. for some  $X \in \Lambda$ .

Let  $\mathcal{R}_0$  be a compact subset of  $\mathbb{H}$  such that  $\mathcal{Q}(t) \subset \mathcal{R}_0$  for all  $t \in [0, T]$  (where  $\mathcal{Q}$  is as in Step 1 and  $\mathcal{Q}(t) = \{u(t); u \in \mathcal{Q}\}$ ). The multifunction  $t \mapsto G(t, \mathcal{R}_0)$  is measurable and has compact values in  $\mathbb{L}$ . Let  $\epsilon > 0$ . There exists a compact subset  $\mathfrak{K}_\epsilon$  of  $\mathfrak{K}_c(\mathbb{L})$  and a measurable subset  $I_\epsilon$  of  $[0, T]$  such that, denoting by  $\text{meas}$  the Lebesgue measure on  $[0, T]$ ,

$$(12) \quad \text{meas}([0, T] \setminus I_\epsilon) \leq \epsilon, \quad \text{and} \quad (\forall t \in I_\epsilon) \quad G(t, \mathcal{R}_0) \in \mathfrak{K}_\epsilon.$$

Now, from a well known characterization of the compact subsets of  $\mathfrak{K}(\mathbb{L})$  ([39, Theorem 2.5.2], see also [40, Theorem 3.1] for the converse implication), the set

$$\mathcal{K}_\epsilon = \bigcup_{K \in \mathfrak{K}_\epsilon} K$$

is a compact subset of  $\mathbb{L}$ . Let us define a multifunction

$$G_\epsilon : \begin{cases} [0, T] \times \mathbb{H} & \rightarrow \mathfrak{K}_c(\mathbb{L}) \\ (t, x) & \mapsto \mathbf{1}_{I_\epsilon}(t)G(t, x). \end{cases}$$

For every  $g \in \Gamma$  and every  $t \in [0, T]$ , let us set

$$g_\epsilon(t) = \mathbf{1}_{I_\epsilon}(t)g(t).$$

The process  $g_\epsilon$  is  $\mathcal{P}$ -measurable and satisfies, for some  $X \in \Lambda$ ,

$$g_\epsilon(t) \in G_\epsilon(t, X(t)) \subset \mathcal{K}_\epsilon \quad \text{a.e. for every } t \in [0, T]$$

(we assume without loss of generality that  $0 \in \mathcal{R}_0$ ). Thus, from our hypothesis, the set

$$\Xi_\epsilon = \{Y^{g_\epsilon}; g \in \Gamma\}$$

is tight. Let  $R_0 = \sup_{x \in \mathcal{R}_0} \|x\|_{\mathbb{H}}$ . By (HFG)-(i), we have, for every  $g \in \Gamma$  and every  $t \in [0, T]$ ,

$$\|g(t) - g_\epsilon(t)\| \begin{cases} = 0 & \text{if } t \in I_\epsilon \\ = \|g(t)\| \leq C_{\text{growth}}(1 + R_0) & \text{if } t \notin I_\epsilon. \end{cases}$$

Using (10), we thus have

$$\begin{aligned} \text{Hausd}_{\mathcal{N}_c^p(\mathfrak{F}, [0, T]; \mathbb{H})}(\Xi, \Xi_\epsilon) &\leq \sup_{g \in \Gamma} C_{\text{Conv}} T^{(p/2)-1} \mathbf{E} \int_0^t \|g(s) - g_\epsilon(s)\|_{\mathbb{L}}^p ds \\ &\leq C_{\text{Conv}} T^{(p/2)-1} \int_0^t (C_{\text{growth}}(1 + R_0) \mathbf{1}_{[0, T] \setminus I_\epsilon}(s))^p ds \\ &\leq C_{\text{Conv}} T^{(p/2)-1} C_{\text{growth}}^p (1 + R_0)^p \epsilon^p. \end{aligned}$$

From Lemma 3.1, we deduce that  $\Xi$  is tight.

*Third step:* We can also assume without loss of generality that  $A$  is a bounded operator on  $\mathbb{H}$ .

Let  $A_n$  ( $n > \beta$ ) be the Yosida approximations of  $A$ . We are going to prove that

$$(13) \quad \sup_{g \in \Gamma} \mathbf{E} \sup_{0 \leq t \leq T} \left\| \int_0^t \left( S(t-s) - e^{(t-s)A_n} \right) g(s) dW(s) \right\|^p \rightarrow 0 \quad \text{as } n \rightarrow +\infty.$$

Let  $D(A)$  be the domain of  $A$ . We have  $\overline{D(A)} = \mathbb{H}$  thus, for each  $\epsilon > 0$ , we can find a finite  $\epsilon$ -net  $\mathcal{K}_\epsilon$  of  $\mathcal{K}$  which lies in  $D(A)$ . Then, for each  $g \in \Gamma$ , we can define a predictable  $\mathcal{K}_\epsilon$ -valued process  $g_\epsilon$  such that  $\|g - g_\epsilon\|_\infty \leq \epsilon$ . For each  $n$ , the semigroup  $e^{tA_n}$  satisfies an inequality similar to (10) with same constant  $C_{\text{Conv}}$ , because  $C_{\text{Conv}}$  depends only on the parameters  $M$  and  $\beta$  in Hypothesis (HS). We thus have

$$\begin{aligned} \sup_{g \in \Gamma} \mathbf{E} \sup_{0 \leq t \leq T} \left\| \int_0^t S(t-s) (g(s) - g_\epsilon(s)) dW(s) \right\|^p &\leq C_{\text{Conv}} T^{p/2} \epsilon^p \\ \sup_{g \in \Gamma} \mathbf{E} \sup_{0 \leq t \leq T} \left\| \int_0^t e^{(t-s)A_n} (g(s) - g_\epsilon(s)) dW(s) \right\|^p &\leq C_{\text{Conv}} T^{p/2} \epsilon^p. \end{aligned}$$

Therefore, we only need to prove (13) in the case when  $\mathcal{K} \subset D(A)$ .

For every  $x \in D(A)$  and every integer  $n > \beta$ , we have

$$\left( S(t-s) - e^{(t-s)A_n} \right) x = \left[ S(t-s-r) e^{rA_n} \right]_0^{t-s} x = \int_0^{t-s} S(t-s-r) e^{rA_n} (A_n - A) x dr.$$

Thus, assuming that  $\mathcal{K} \subset D(A)$ , and denoting by  $C$  a constant which may differ from line to line, we have, for every  $g \in \Gamma$ , using the stochastic Fubini theorem (see [2]) and the convolution inequality (10),

$$\begin{aligned} \mathbf{E} \sup_{0 \leq t \leq T} \left\| \int_0^t \left( S(t-s) - e^{(t-s)A_n} \right) g(s) dW(s) \right\|^p &= \mathbf{E} \sup_{0 \leq t \leq T} \left\| \int_0^t \int_0^{t-s} S(t-s-r) e^{rA_n} (A_n - A) g(s) dr dW(s) \right\|^p \\ &= \mathbf{E} \sup_{0 \leq t \leq T} \left\| \int_0^t e^{rA_n} \int_0^{t-r} S(t-s-r) e^{rA_n} (A_n - A) g(s) dW(s) dr \right\|^p \\ &\leq C \mathbf{E} \sup_{0 \leq t \leq T} \int_0^t \left\| \int_0^{t-r} S(t-s-r) (A_n - A) g(s) dW(s) \right\|^p dr \\ &\leq C \int_0^T \mathbf{E} \sup_{0 \leq t \leq T} \left\| \int_0^{t-r} S(t-s-r) (A_n - A) g(s) dW(s) \right\|^p dr \\ &\leq C \int_0^T \int_0^T \mathbf{E} \|(A_n - A) g(s)\|^p ds dr. \end{aligned}$$

But, for every  $x \in D(A)$ , we have  $(A_n - A)x \rightarrow 0$ . So, using the compactness of  $\mathcal{K}$  and Lebesgue's dominated convergence theorem, we obtain (13).

From Lemma 3.1, we conclude that we only need to check Lemma 3.2 for the semigroups  $e^{tA_n}$  ( $n > \beta$ ), which amounts to check Lemma 3.2 in the case when  $D(A) = \mathbb{H}$ . In this case,  $S(t)$  is the exponential  $e^{tA}$  in the usual sense.

*Fourth step: We can assume without loss of generality that  $\mathbb{H}$  is finite dimensional.*

Let  $(e_n)$  be an orthonormal basis of  $\mathbb{H}$ . For each  $n$ , let  $\mathbb{H}_n = \text{Span}(e_1, \dots, e_n)$  and let  $P_n$  be the orthogonal projection from  $\mathbb{H}$  onto  $\mathbb{H}_n$ . Let  $\Gamma$  be any contour around the spectrum of  $A$ , say  $\Gamma$  is a circle  $C(0, \rho)$ . Denoting by  $R$  the resolvent operator, we have, for any  $g \in \Gamma$ ,

$$\begin{aligned} & \int_0^t \left( e^{(t-s)A} - e^{(t-s)P_n A} \right) P_n g(s) dW(s) \\ &= \int_0^t \frac{1}{2\pi i} \int_{\Gamma} e^{\lambda(t-s)} (R(\lambda, A) - R(\lambda, P_n A)) d\lambda P_n g(s) dW(s) \\ &= \frac{1}{2\pi i} \int_{\Gamma} \int_0^t e^{\lambda(t-s)} (R(\lambda, A) - R(\lambda, P_n A)) P_n g(s) dW(s) d\lambda \\ &= \frac{\rho}{2\pi} \int_0^{2\pi} \int_0^t e^{\rho e^{i\theta}(t-s)} \left( R(\rho e^{i\theta}, A) - R(\rho e^{i\theta}, P_n A) \right) P_n g(s) dW(s) d\theta \end{aligned}$$

by the stochastic Fubini theorem. Denoting again by  $C$  a constant which may change from line to line, we thus have

$$\begin{aligned} & \mathbf{E} \sup_{0 \leq t \leq T} \left\| \int_0^t \left( e^{(t-s)A} - e^{(t-s)P_n A} \right) P_n g(s) dW(s) \right\|^p \\ & \leq C \mathbf{E} \int_0^{2\pi} \sup_{0 \leq t \leq T} \int_0^t \left\| e^{\rho e^{i\theta}(t-s)} \left( R(\rho e^{i\theta}, A) - R(\rho e^{i\theta}, P_n A) \right) P_n g(s) dW(s) \right\|^p d\theta \\ & \leq C \mathbf{E} \int_0^{2\pi} \sup_{0 \leq t \leq T} \int_0^t \left\| e^{\rho e^{i\theta}(t-s)} \left( R(\rho e^{i\theta}, A) - R(\rho e^{i\theta}, P_n A) \right) P_n g(s) dW(s) \right\|^p d\theta \\ & \leq C \int_0^{2\pi} \int_0^T \mathbf{E} \left\| \left( R(\rho e^{i\theta}, A) - R(\rho e^{i\theta}, P_n A) \right) P_n g(s) dW(s) \right\|^p d\theta \end{aligned}$$

using the convolution inequality for the semigroup  $t \mapsto e^{\rho e^{i\theta} t}$ . From the compactness of  $\mathcal{K}$  and Lebesgue's dominated convergence theorem, we get

$$\sup_{g \in \Gamma} \mathbf{E} \sup_{0 \leq t \leq T} \left\| \int_0^t \left( e^{(t-s)A} - e^{(t-s)P_n A} \right) P_n g(s) dW(s) \right\|^p$$

and we conclude as in Step 3.

*Fifth step: Assuming all preceding reductions, we now prove Lemma 3.2.*

Recall that  $R_1 = \sup_{x \in \mathcal{K}} \|x\|_{\mathbb{L}}$ . We have, for any  $\epsilon > 0$  and  $R > 0$ ,

$$\begin{aligned} \mathbf{P} \left\{ \sup_{0 \leq t \leq T} \|Y(t)\| \leq R \right\} & \leq \frac{4}{R^2} \mathbf{E} \|Y(t)\|^2 \\ & = \frac{4}{R^2} \mathbf{E} \int_0^T \left\| e^{(T-s)A} g(s) \right\|^2 ds \\ & \leq \frac{4}{R^2} T e^{T\|A\|} R_1^2 \end{aligned}$$

Taking  $R$  large enough, we get

$$(14) \quad (\forall \epsilon > 0) \quad (\exists R > 0) \quad \mathbf{P} \left\{ \sup_{0 \leq t \leq T} \|Y(t)\| \leq R \right\} \leq \epsilon.$$

Now, let  $\mathfrak{T}$  be the set of stopping times  $\tau$  such that  $0 \leq \tau \leq T$ . If  $\sigma, \tau \in \mathfrak{T}$  with  $0 < \tau - \sigma \leq \delta$  for some  $\delta > 0$ , we have, for any  $Y \in \Xi$  of the form  $Y = Y^{(g)}$ , with  $g \in G$ , and for any  $\eta > 0$ ,

$$\begin{aligned} \mathbf{P} \{ \|Y(\tau) - Y(\sigma)\| > \eta \} &\leq \frac{1}{\eta^2} \mathbf{E} \|Y(\tau) - Y(\sigma)\|^2 \\ &\leq \frac{2}{\eta^2} \left( \mathbf{E} \int_0^\sigma \left\| (e^{(\tau-s)A} - e^{(\sigma-s)A})g(s) \right\|^2 ds \right. \\ &\quad \left. + \mathbf{E} \int_\sigma^\tau \left\| e^{(\tau-s)A}g(s) \right\|^2 ds \right) \\ &\leq \frac{2}{\eta^2} \left( T\delta e^{2T\|A\|} \|A\|^2 R_1^2 + \delta e^{\delta\|A\|} R_1^2 \right) \\ &\leq \delta C(\eta) \end{aligned}$$

with  $C(\eta) = \frac{2e^{2T\|A\|} R_1^2 (1+T)}{\eta^2}$ . Taking  $\delta$  small enough, we get

$$(15) \quad (\forall \epsilon > 0) (\forall \eta > 0) (\exists \delta > 0) (\forall Y \in \Xi) \sup_{\substack{\sigma, \tau \in \mathfrak{T} \\ 0 < \tau - \sigma \leq \delta}} \mathbf{P} \{ \|Y(\tau) - Y(\sigma)\| > \eta \} \leq \epsilon.$$

From (14) and (15), we conclude, by a criterion of Aldous [41, 42], that  $\Xi$  is tight.  $\square$

### The multivalued operator $\Phi$

Let us denote by  $\Phi$  the mapping which, with every continuous adapted  $\mathbb{H}$ -valued process  $X$  such that  $\mathbf{E} \int_0^T \|X(s)\|^p ds < +\infty$ , associates the set of all processes of the form

$$S(t)\xi + \int_0^t S(t-s)f(s) ds + \int_0^t S(t-s)g(s) dW(s),$$

where  $f$  and  $g$  are predictable selections of  $(\omega, t) \mapsto F(t, X(\omega, t))$  and  $(\omega, t) \mapsto G(t, X(\omega, t))$  respectively. Lemma 3.2 will be used through the following corollary:

**Corollary 3.3 (The operator  $\Phi$  maps tight sets into tight sets)** *Let  $\Lambda$  be a set of continuous adapted processes on  $\mathbb{H}$ . Assume that each element of  $\Lambda$  is in  $L^p(\Omega \times [0, T]; \mathbb{H})$  and that  $\Lambda$ , considered as a set of  $C([0, T]; \mathbb{H})$ -valued random variables, is tight. Assume furthermore that Hypothesis (HS) and (HFG) are satisfied. Then  $\Phi \circ \Lambda := \cup_{X \in \Lambda} \Phi(X)$  is a tight set of  $C([0, T]; \mathbb{H})$ -valued random variables.*

**Proof.** Let us denote by  $\Phi_F$  the mapping which, with every continuous adapted  $\mathbb{H}$ -valued process  $X$  such that  $\mathbf{E} \int_0^T \|X(s)\|^p ds < +\infty$ , associates the set of all processes of the form

$$\int_0^t S(t-s)f(s) ds,$$

where  $f$  is a predictable selection of  $(\omega, t) \mapsto F(t, X(t))$ .

By Ascoli's theorem (see the details in the proof of [23, Lemma 4.2.1]), the mapping

$$f \mapsto \int_0^\cdot S(\cdot - s)f(s) ds$$

maps all measurable functions  $f : [0, T] \rightarrow \mathbb{H}$  with values in a given compact subset of  $\mathbb{H}$  into a compact subset of  $C([0, T]; \mathbb{H})$ . Thus, following the same lines as in Steps 1 and 2 of the proof of Lemma 3.2,  $\Phi_F$  maps tight bounded subsets of  $L^p(\Omega \times [0, T]; \mathbb{H})$  into tight sets of  $C([0, T]; \mathbb{H})$ -valued random variables. The conclusion immediately follows from Lemma 3.2.  $\square$

**Lemma 3.4 (The operator  $\Phi$  and the measure of noncompactness  $\Psi$ )** *Let  $\Lambda$  be a bounded subset of  $\mathcal{N}_c^p(\mathfrak{F}, [0, T]; \mathbb{H})$ . Assume Hypothesis (HFG). We then have*

$$\Psi^p(\Phi \circ \Lambda)(t) \leq k \int_0^t L(s, \Psi^p(\Lambda))(s) ds$$

for some constant  $k$  which depends only on  $T$ ,  $p$ ,  $M_T$ , and  $C_{\text{Conv}}$ .

**Proof.** The main arguments are inspired from [8, Lemma 4.2.6]. For simplicity, the space  $\mathcal{N}_c^p(\mathfrak{F}, [0, t]; \mathbb{H})$  will be denoted by  $\mathcal{N}_c^p(t)$ .

Let  $\epsilon > 0$ . The function  $t \mapsto \Psi(\Lambda)(t)$  is increasing, thus there exist at most a finite number of points  $0 \leq t_1 \leq \dots \leq t_n \leq T$  for which  $\Psi(\Lambda)$  makes a jump greater than  $\epsilon$ . Let us choose  $\delta_1 > 0$  such that  $i \neq j \Rightarrow ]t_i - \delta_1, t_i + \delta_1[ \cap ]t_j - \delta_1, t_j + \delta_1[ = \emptyset$ . Using points  $\beta_j$ ,  $j = 1, \dots, m$ , we divide the remaining part  $[0, T] \setminus (]t_1 - \delta_1, t_1 + \delta_1[ \cup \dots \cup ]t_n - \delta_1, t_n + \delta_1[)$  into disjoint intervals in such a way that, for each  $j$ ,

$$(16) \quad \sup_{s, t \in ]\beta_{j-1}, \beta_j]} |\Psi(\Lambda)(s) - \Psi(\Lambda)(t)| < \epsilon.$$

Let us then choose  $\delta_2 > 0$  such that  $j \neq k \Rightarrow ]\beta_j - \delta_2, \beta_j + \delta_2[ \cap ]\beta_k - \delta_2, \beta_k + \delta_2[ = \emptyset$ .

Now we start the construction of a tight net.

For each  $j = 1, \dots, m$ , we choose a tight  $(\Psi(\Lambda)(\beta_j) + \epsilon)$ -net  $N_j$  of  $\Lambda$  in  $\mathcal{N}_c^p(\beta_j)$ . (As  $\mathcal{N}_c^p(\beta_j)$  is separable, we can take a countable net  $N_j$ , but this fact will not be used here.) We obtain a (countable) family  $\mathcal{Z}$  in  $\mathcal{N}_c^p(T)$  by taking all continuous processes which coincide on each  $] \beta_{j-1} + \delta_2, \beta_j - \delta_2[$  ( $1 \leq j \leq m$ ) with some element of  $N_j$  and which have affine trajectories on the complementary segments. The set  $\mathcal{Z}$  is tight and bounded in  $L^p(\Omega \times [0, T]; \mathbb{H})$ , thus, by Corollary 3.3,  $\Phi(\mathcal{Z})$  is tight.

Consider a fixed  $X \in \Lambda$ .

We can find an element  $Z$  of  $\mathcal{Z}$  such that, for each  $j = 1, \dots, m$ ,

$$\|X - Z\|_{\mathcal{N}_c^p(\beta_j)} \leq \Psi(\Lambda)(\beta_j) + \epsilon.$$

For  $t \in ]\beta_{j-1} + \delta_2, \beta_j - \delta_2[$ , we have, using (16),

$$(17) \quad \begin{aligned} \mathbf{E} \|X(t) - Z(t)\|_{\mathbb{H}}^p &\leq \mathbf{E} \sup_{\beta_{j-1} + \delta_2 \leq s \leq \beta_j - \delta_2} \|X(s) - Z(s)\|_{\mathbb{H}}^p \\ &\leq \|X - Z\|_{\mathcal{N}_c^p(\beta_j)}^p \\ &\leq (\Psi(\Lambda)(\beta_j) + \epsilon)^p \\ &\leq (\Psi(\Lambda)(t) + 2\epsilon)^p \end{aligned}$$

Let  $Y \in \Phi(X)$ , say

$$Y(t) = S(t)X(0) + \int_0^t S(t-s)f(s) ds + \int_0^t S(t-s)g(s) dW(s),$$

where  $f$  and  $g$  are predictable selections of  $(\omega, t) \mapsto F(t, X(\omega, t))$  and  $(\omega, t) \mapsto G(t, X(\omega, t))$  respectively. We can find predictable selections  $\tilde{f}$  and  $\tilde{g}$  of  $(\omega, t) \mapsto F(t, Z(\omega, t))$  and  $(\omega, t) \mapsto G(t, Z(\omega, t))$  respectively, such that, for every  $t \in [0, T]$ ,

$$(18) \quad \begin{aligned} \|\tilde{f}(t) - f(t)\|^p &\leq 2 \text{Hausd}^p(F(t, X(t)), F(t, Z(t))) \\ &\leq 2L(t, \|X(t) - Z(t)\|^p) \end{aligned}$$

and

$$(19) \quad \begin{aligned} \|\tilde{g}(t) - g(t)\|^p &\leq 2 \text{Hausd}^p(G(t, X(t)), G(t, Z(t))) \\ &\leq 2L(t, \|X(t) - Z(t)\|^p). \end{aligned}$$

Let  $d_{\mathcal{N}_\epsilon^p(t)}$  be the distance in  $\mathcal{N}_\epsilon^p(t)$  associated with  $\|\cdot\|_{\mathcal{N}_\epsilon^p(t)}$ . We have, using (HFG) and the convexity of  $x \mapsto x^p$ ,

$$\begin{aligned} d_{\mathcal{N}_\epsilon^p(t)}^p(Y, \Phi(Z)) &\leq \mathbf{E} \sup_{0 \leq \tau \leq t} \left\| \int_0^\tau S(t-s)(f(s) - \tilde{f}(s)) ds \right. \\ &\quad \left. + \int_0^\tau S(t-s)(g(s) - \tilde{g}(s)) dW(s) \right\|^p \\ &\leq 2^{p-1} \mathbf{E} \sup_{0 \leq \tau \leq t} \left\| \int_0^\tau S(t-s)(f(s) - \tilde{f}(s)) ds \right\|^p \\ &\quad + \mathbf{E} \sup_{0 \leq \tau \leq t} \left\| \int_0^\tau S(t-s)(g(s) - \tilde{g}(s)) dW(s) \right\|^p \\ &\leq 2^{p-1} (M_t \int_0^t \mathbf{E} \|f(s) - \tilde{f}(s)\|^p ds \\ &\quad + C_{\text{Conv}} T^{p/2-1} \int_0^t \mathbf{E} \|g(s) - \tilde{g}(s)\|^p ds) \end{aligned}$$

(recall that  $M_t$  is defined in (HS)). Let us denote

$$(20) \quad \begin{aligned} k &= 2^p \max \left( M_T, C_{\text{Conv}} T^{p/2-1} \right), \\ J(t) &= [0, t] \cap \left( (\cup_{1 \leq i \leq n} [t_i - \delta_1, t_i + \delta_1]) \cup (\cup_{1 \leq j \leq m} [\beta_j - \delta_2, \beta_j + \delta_2]) \right), \\ I(t) &= [0, t] \setminus J(t). \end{aligned}$$

As  $\Lambda$  and  $\mathcal{Z}$  are bounded in  $L^p(\Omega \times [0, T]; \mathbb{H})$ , using (HFG) (18) and (19), and taking  $\delta_1$  and  $\delta_2$  sufficiently small, we get

$$d_{\mathcal{N}_\epsilon^p(t)}^p(Y, \Phi(Z)) \leq k \left( \int_{I(t)} \mathbf{E} L(s, \|X(s) - Z(s)\|^p) ds + \epsilon \right).$$

Then, using the convexity of  $L(t, \cdot)$ , and (17),

$$d_{\mathcal{N}_\epsilon^p(t)}^p(Y, \Phi(Z)) \leq k \left( \int_0^t L(s, (\Psi(\Lambda)(s) + 2\epsilon)^p) ds + \epsilon \right).$$

As  $\epsilon$ ,  $X$  and  $Y$  are arbitrary, the result follows.  $\square$

Here is a easy variant for the case when  $F$  and  $G$  are single-valued:

**Lemma 3.5** *Assume that  $F$  and  $G$  are single-valued. Assume furthermore Hypothesis (HFG). Let  $\Psi'$  be the measure of noncompactness on  $\mathcal{N}_c^p(\mathfrak{F}, [0, T]; \mathbb{H})$  defined by*

$$\Psi'(\Lambda)(s) := \inf \left\{ \epsilon > 0; \Lambda|_s \text{ has a finite } \epsilon\text{-net in } \mathcal{N}_c^p(\mathfrak{F}, [0, s]; \mathbb{H}) \right\} \quad (0 \leq s \leq T).$$

Let  $\Lambda$  be a bounded subset of  $\mathcal{N}_c^p(\mathfrak{F}, [0, T]; \mathbb{H})$ . We have

$$\Psi'^p(\Phi \circ \Lambda)(t) \leq k' \int_0^t L(s, \Psi'^p(\Lambda))(s) ds$$

for some constant  $k'$  which depends only on  $T$ ,  $p$ ,  $M_T$ , and  $C_{\text{Conv}}$ .

**Proof.** We only need to repeat the proof of Lemma 3.4, but we take for each  $j$  a finite  $(\Psi'(\Lambda)(\beta_j) + \epsilon)$ -net  $N_j$ . Then  $\mathcal{Z}$  is finite, thus  $\Phi(\mathcal{Z})$  is compact in  $\mathcal{N}_c^p(\mathfrak{F}, [0, T]; \mathbb{H})$ . The rest of the proof goes as in the proof of Lemma 3.4.  $\square$

In the proof of Theorem 2.1, we shall consider a variant of the operator  $\Phi$ . For each  $n \geq 1$ , let  $\Phi_n$  be the mapping which, with every continuous adapted  $\mathbb{H}$ -valued process  $X$  such that  $\mathbf{E} \int_0^T \|X(s)\|^p ds < +\infty$ , associates the set of all processes of the form

$$S(t - 1/n)\xi + \int_0^{t-1/n} S(t-s)f(s) ds + \int_0^{t-1/n} S(t-s)g(s) dW(s),$$

where  $f$  and  $g$  are predictable selections of  $(\omega, t) \mapsto F(t, X(\omega, t))$  and  $(\omega, t) \mapsto G(t, X(\omega, t))$ . The following lemma links the tightness properties of  $\Phi$  and  $\Phi_n$ .

**Lemma 3.6** *Let  $(X_n)$  be a sequence of continuous adapted  $\mathbb{H}$ -valued processes, which is bounded in  $L^p(\Omega \times [0, T]; \mathbb{H})$ . We then have*

$$\Psi(\cup_n \Phi_n(X_n)) \leq M_T \Psi(\cup_n \Phi(X_n)),$$

where  $M_T$  has been defined in (HS).

**Proof.** First, in the definition of  $\Phi_n$  and in that of  $\Phi$ , we can assume that  $\xi = 0$ , because  $\xi$  does not change the values of  $\Psi(\cup_n \Phi_n(X_n))$  and  $\Psi(\cup_n \Phi(X_n))$ . We thus have

$$\begin{aligned} \Phi_n(X_n)(t) &= \left\{ S(1/n) \int_0^{t-1/n} S((t-1/n)-s)f(s) ds \right. \\ &\quad \left. + S(1/n) \int_0^{t-1/n} S((t-1/n)-s)g(s) dW(s); \right. \\ &\quad \left. f \text{ and } g \text{ are predictable selections of } F \circ X \text{ and } G \circ X \right\} \\ &= S(1/n)\Phi(X_n)(t - 1/n). \end{aligned}$$

Let  $\tau_n : \mathcal{C}([0, T]; \mathbb{H}) \rightarrow \mathcal{C}([0, T]; \mathbb{H})$  be defined by

$$\tau_n(u)(t) = \begin{cases} u(t - 1/n) & \text{if } t \geq 1/n \\ u(0) & \text{if } t \leq 1/n. \end{cases}$$

For any set  $\Lambda$  of continuous processes, we have

$$\Psi(\tau_n(\Lambda)) \leq \Psi(\Lambda)$$



because, if  $\Xi$  is a tight  $\epsilon$ -net of  $\Lambda$ , then  $\tau_n(\Xi)$  is a tight  $\epsilon$ -net of  $\tau_n(\Lambda)$ . We thus have

$$\begin{aligned}\Psi(\cup_n \Phi_n(X_n)) &= \Psi(S(1/n) \cup_n \tau_n(\Phi(X_n))) \\ &\leq M_T \Psi(\cup_n \tau_n(\Phi(X_n))) \leq M_T \Psi(\cup_n \Phi(X_n)).\end{aligned}$$

□

## 4 Proof of the main results

### Proof of Theorem 2.1.

*First step: construction of a tight sequence of approximating solutions through Tonelli's scheme.*

For each integer  $n \geq 1$ , we can easily define a process  $\tilde{X}_n$  on  $[-1, T]$  by  $\tilde{X}_n(t) = 0$  if  $t \leq 0$  and, for  $t \geq 0$ ,

$$\tilde{X}_n(t) = S(t)\xi + \int_0^t S(t - (s - \frac{1}{n}))f_n(s) ds + \int_0^t S(t - (s - \frac{1}{n}))g_n(s) dW(s),$$

where  $f_n : \Omega \times [0, T] \times \rightarrow \mathbb{H}$  and  $g_n : \Omega \times [0, T] \times \rightarrow \mathbb{L}$  are predictable and

$$f_n(s) \in F(s, \tilde{X}_n(s - 1/n)) \text{ a.e.} \quad \text{and} \quad g_n(s) \in G(s, \tilde{X}_n(s - 1/n)) \text{ a.e.}$$

We then set, for  $t \in [1/n, T]$ ,

$$\begin{aligned}X_n(t) &= \tilde{X}_n(t - 1/n) \\ &= S(t - 1/n)\xi + \int_0^{t-1/n} S(t - s)f_n(s) ds + \int_0^{t-1/n} S(t - s)g_n(s) dW(s).\end{aligned}$$

For  $t \leq 1/n$ , we set  $X_n(t) = \xi$ . Let us show that  $(X_n)$  is bounded in  $L^p(\Omega \times [0, T]; \mathbb{H})$ . By convexity of  $t \mapsto |t|^p$ , we have

$$(21) \quad \|a_1 + \dots + a_m\|^p \leq m^{p-1} (\|a_1\|^p + \dots + \|a_m\|^p)$$

for any finite sequence  $a_1, \dots, a_m$  in any normed space. Recall also that  $M_T$  and  $C_{\text{growth}}$  have been defined in Hypothesis (HS) and (HFG)-(i) and  $C_{\text{Conv}}$  is the constant of stochastic convolution defined in (10). For every  $n \geq 1$ , we have the following chain of inequalities, where the supremum is taken over all predictable selections  $(f, g)$  of  $(\omega, t) \mapsto F(t, X_n(t)) \times G(t, X_n(t))$ :

$$\begin{aligned}\mathbf{E} \|X_n(t)\|^p &\leq \sup_{f, g} 3^{p-1} \left( \mathbf{E} \|S(t - 1/n)\xi\|^p + \mathbf{E} \int_0^{t-1/n} \|S(t - s)f(s)\|^p ds \right. \\ &\quad \left. + \mathbf{E} \int_0^{t-1/n} \|S(t - s)g(s)\|^p dW(s) \right) \\ &\leq \sup_g 3^{p-1} \left( M_T^p \mathbf{E} \|\xi\|^p + M_T^p C_{\text{growth}}^p \mathbf{E} \int_0^t (1 + \|X_n\|(s))^p ds \right. \\ &\quad \left. + M_T^p C_{\text{Conv}} \mathbf{E} \int_0^{t-1/n} \|g(s)\|^p ds \right) \\ &\leq 3^{p-1} M_T^p \left( \mathbf{E} \|\xi\|^p + C_{\text{growth}}^p (1 + C_{\text{Conv}}) \mathbf{E} \int_0^t (1 + \|X_n(s)\|)^p ds \right). \\ &\leq 3^{p-1} M_T^p \left( \mathbf{E} \|\xi\|^p + C_{\text{growth}}^p (1 + C_{\text{Conv}}) 2^{p-1} \mathbf{E} \int_0^t 1 + \|X_n(s)\|^p ds \right).\end{aligned}$$

Let  $K = 3^{p-1}M_T^p(\mathbf{E}\|\xi\|^p + 2^{p-1}C_{\text{growth}}^p(1 + C_{\text{Conv}})T)$ . We have

$$\mathbf{E}\|X_n(t)\|^p \leq K + K \int_0^t \mathbf{E}\|X_n(s)\|^p ds,$$

thus, by Gronwall Lemma,

$$\mathbf{E}\|X_n(t)\|^p \leq Ke^{Kt}$$

which provides the boundedness condition

$$(22) \quad \sup_n \mathbf{E} \int_0^T \|X_n(t)\|^p dt \leq TK e^{KT}.$$

Let us now show that  $(X_n)$  is a tight sequence of  $C([0, T]; \mathbb{H})$ -valued random variables. As  $(X_n)$  is bounded in  $L^p(\Omega \times [0, T]; \mathbb{H})$ , we can apply Lemma 3.6 and Lemma 3.4. Using the fact that  $X_n \in \Phi_n(X_n)$  for each  $n$ , we get, for every  $t \in [0, T]$ ,

$$\begin{aligned} (\Psi(\cup_n \{X_n\})(t))^p &\leq (\Psi(\cup_n \Phi_n(X_n)(t)))^p \leq M_T^p \Psi(\cup_n \Phi(X_n)(t))^p \\ &\leq M_T^p k \int_0^t L(s, \Psi^p(\cup_n \{X_n\})(s)) ds \end{aligned}$$

(where  $k$  is the constant we obtained in Lemma 3.4, see (20)). Thus, by (9) in Hypothesis (HFG)-(iib), we have  $\Psi(\cup_n \{X_n\})(t) = 0$  for each  $t$ , which, by Lemma 3.1, implies that  $(X_n)$  is tight.

*Second step: construction of a weak solution.*

By Prohorov's compactness criterion for Young measures, we can extract a subsequence of  $(X_n)$  which converges stably to a Young measure  $\mu \in \mathcal{Y}(\Omega, \mathcal{F}, \mathbb{P}; C([0, T]; \mathbb{H}))$ . For simplicity, we denote this extracted sequence by  $(X_n)$ .

It will be convenient to represent the limiting Young measure  $\mu$  as a random variable defined on an extended probability space. Let  $\mathcal{C}$  be the Borel  $\sigma$ -algebra of  $C([0, T]; \mathbb{H})$  and, for each  $t \in [0, T]$ , let  $\mathcal{C}_t$  be the sub- $\sigma$ -algebra of  $\mathcal{C}$  generated by  $C([0, t]; \mathbb{H})$ .

We define a stochastic basis  $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_t, \mu)$  by

$$\underline{\Omega} = \Omega \times C([0, T]; \mathbb{H}), \quad \underline{\mathcal{F}} = \mathcal{F} \otimes \mathcal{C} \quad \underline{\mathcal{F}}_t = \mathcal{F}_t \otimes \mathcal{C}_t$$

and we define  $X_\infty$  on  $\underline{\Omega}$  by

$$X_\infty(\omega, u) = u.$$

Clearly,  $\mathcal{L}(X_\infty) = \mu$  and  $X_\infty$  is  $(\underline{\mathcal{F}}_t)$ -adapted. Now, the random variables  $X_n$  can be seen as random elements defined on  $\underline{\Omega}$ , using the notation

$$X_n(\omega, u) := X_n(\omega) \quad (n \in \mathbb{N}).$$

Furthermore,  $X_n$  is  $(\underline{\mathcal{F}}_t)$ -adapted for each  $n$ . The  $\sigma$ -algebra  $\mathcal{F}$  can be identified with the sub- $\sigma$ -algebra  $\{A \times C([0, T]; \mathbb{H}); A \in \mathcal{F}\}$  of  $\underline{\mathcal{F}}$ . We thus have:

$$(23) \quad (\forall A \in \mathcal{F}) \quad (\forall f \in \mathcal{C}_b(\mathbb{H})) \quad \lim_n \int_A f(X_n) d\mu = \int_A f(X_\infty) d\mu.$$

To express (23), we say that  $(X_n)$  converges to  $X_\infty$   $\mathcal{F}$ -stably.

Let us show that  $W$  is an  $(\underline{\mathcal{F}}_t)$ -Wiener process under the probability  $\mu$ . Clearly,  $W$  is  $(\underline{\mathcal{F}}_t)$ -adapted, so we only need to prove that  $W$  has independent increments. By a result

of Balder [26, 43], each subsequence of  $(\underline{\delta}_{X_n})$  contains a further subsequence  $(\underline{\delta}_{X'_n})$  which K-converges to  $\mu$ , that is, for each subsequence  $(\underline{\delta}_{X''_n})$  of  $(\underline{\delta}_{X'_n})$ , we have

$$\lim_n \frac{1}{n} \sum_{i=1}^n \delta_{X''_n(\omega)} = \mu_\omega \text{ a.e.}$$

This entails that, for every  $A \in \mathcal{C}_t$ , the mapping  $\omega \mapsto \mu_\omega(A)$  is  $\mathcal{F}_t$ -measurable<sup>1</sup>. Let  $t \in [0, T]$  and let  $s > 0$  such that  $t + s \in [0, T]$ . Let us prove that, for any  $A \in \underline{\mathcal{F}}_t$  and any Borel subset  $C$  of  $\mathbb{U}$ , we have

$$(24) \quad \mu(A \cap \{\omega \in \Omega; W(t+s) - W(t) \in C\}) = \mu(A) \mu\{\omega \in \Omega; W(t+s) - W(t) \in C\}.$$

Let  $B = \{\omega \in \Omega; W(t+s) - W(t) \in C\}$ . We have

$$\begin{aligned} \mu(A \cap (B \times C([0, T]; \mathbb{H}))) &= \int_{\Omega \times C([0, T]; \mathbb{H})} \mathbf{1}_A(\omega, u) \mathbf{1}_B(\omega) d\mu(\omega, u) \\ &= \int_{\Omega} \mu_\omega(\mathbf{1}_A(\omega, \cdot)) \mathbf{1}_B(\omega) dP(\omega) \\ &= \int_{\Omega} \mu_\omega(\mathbf{1}_A(\omega, \cdot)) dP(\omega) P(B) \\ &= \mu(A) \mu(B \times C([0, T]; \mathbb{H})), \end{aligned}$$

which proves (24). Thus  $W(t+s) - W(t)$  is independent of  $\underline{\mathcal{F}}_t$ .

Now, there remains to prove that  $X = X_\infty$  satisfies (8). Note that the first step of the proof of Theorem 2.1 is valid for any choice of the selections  $f_n$  and  $g_n$ . In this second part, we will use a particular choice. As  $F$  and  $G$  are globally measurable, they have measurable graphs (see [24, Proposition III.13]). As furthermore they are continuous in the second variable, they admit Carathéodory selections [44, Corollary 1 of the Main Lemma], that is, there exist globally measurable mappings  $\mathfrak{f} : [0, T] \times \mathbb{H} \rightarrow \mathbb{H}$  and  $\mathfrak{g} : [0, T] \times \mathbb{H} \rightarrow \mathbb{L}$  such that  $\mathfrak{f}(t, \cdot)$  and  $\mathfrak{g}(t, \cdot)$  are continuous for every  $t \in [0, T]$  and  $\mathfrak{f}(t, x) \in F(t, x)$  and  $\mathfrak{g}(t, x) \in G(t, x)$  for all  $(t, x) \in [0, T] \times \mathbb{H}$ . We denote by  $\widehat{\mathbb{N}}$  the one-point compactification of  $\mathbb{N}$  and we set, for every  $n \in \widehat{\mathbb{N}}$  and every  $s \in [0, T]$ ,

$$\begin{aligned} f_n(s) &= \mathfrak{f}(s, X_n(s)) \\ g_n(s) &= \mathfrak{g}(s, X_n(s)). \end{aligned}$$

The sequence  $(f_n)$  converges in law to  $f_\infty$  in the space  $L^p([0, T]; \mathbb{H})$ . Indeed, by the growth condition (HFG)-(i) and Lebesgue's convergence theorem, the mapping

$$\mathfrak{h} : \begin{cases} C([0, T]; \mathbb{H}) & \rightarrow L^p([0, T]; \mathbb{H}) \\ u & \mapsto \mathfrak{f}(\cdot, (u(\cdot))) \end{cases}$$

is well defined and continuous, and we have  $f_n = \mathfrak{h} \circ X_n$ . Similarly,  $(g_n)$  converges in law to  $g_\infty$  in  $L^p([0, T]; \mathbb{L})$ . Actually, we have more: the sequence  $(X_n, W, f_n, g_n)$  converges in law to  $(X_\infty, W, f_\infty, g_\infty)$  in  $C([0, T]; \mathbb{H}) \times C([0, T]; \mathbb{U}) \times L^p([0, T]; \mathbb{H}) \times L^p([0, T]; \mathbb{L})$ . In particular, as  $p > 2$ , the convergence in law of  $(W, g_n)$  to  $(W, g_\infty)$  implies, for every  $t \in [0, T]$ , the convergence in law of the stochastic integrals  $\int_0^t S(t-s)g_n(s) dW(s)$  to

<sup>1</sup>From [21, Lemma 2.17], this means that  $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_t, \mu)$  is a *very good extension* of  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, P)$  in the sense of [21], that is, every martingale on  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, P)$  remains a martingale on  $(\underline{\Omega}, \underline{\mathcal{F}}, (\underline{\mathcal{F}}_t)_t, \mu)$ .

$\int_0^t S(t-s)g_\infty(s) dW(s)$ . To see this, one can e.g. use a Skorokhod representation  $(W'_n, g'_n)$  which converges a.e. to  $(W', g'_\infty)$ , then, by Vitali's convergence theorem,  $(W'_n, g'_n)$  converges in  $L^2(\Omega \times [0, T]; \mathbb{U} \times \mathbb{L})$  thus, for every  $t \in [0, T]$  the stochastic integrals  $\int_0^t S(t-s)g'_n(s) dW'(s)$  converge in  $L^2(\Omega \times [0, t]; \mathbb{H})$ . Let  $t \in [0, T]$  be fixed and let us set, for every  $n \in \widehat{N}$ ,

$$Z_n(t) = -X_n(t) + S(t)\xi + \int_0^{t-1/n} S(t-s)f_n(s) ds + \int_0^{t-1/n} S(t-s)g_n(s) dW(s)$$

(with  $1/\infty := 0$ ). For every  $t \in [0, T]$ , the sequence  $(Z_n(t))$  converges in law to  $Z_\infty(t)$ . Now, from the definition of  $X_n$  ( $n < +\infty$ ), we have

$$\begin{aligned} (\forall n < +\infty) \quad Z_n(t) &= (S(t) - S(t-1/n)) \xi \\ &\quad + \int_{t-1/n}^t S(t-s)f_n(s) ds + \int_{t-1/n}^t S(t-s)g_n(s) dW(s). \end{aligned}$$

But  $(S(t) - S(t-1/n)) \xi$  converges a.e. to 0 and we have, using the growth condition (HFG)-(i) and the boundedness property (22),

$$\begin{aligned} \mathbf{E} \left\| \int_{t-1/n}^t S(t-s)f_n(s) ds \right\| &\leq M e^{\beta T} \int_0^t \|f_n(s) \mathbf{1}_{[t-1/n, t]}(s)\| ds \\ &\leq M e^{\beta T} \mathbf{E} \left( \int_0^t \|f_n(s)\|^p ds \right)^{1/p} \left( \frac{1}{n} \right)^{p/(p-1)} \\ &\leq M e^{\beta T} \left( 2^{p-1} \mathbf{E} \int_0^t (1 + \|X_n(s)\|)^p ds \right)^{1/p} \left( \frac{1}{n} \right)^{p/(p-1)} \\ &\rightarrow 0 \text{ as } n \rightarrow \infty \end{aligned}$$

and, for any  $q$  such that  $2 < q < p$ ,

$$\begin{aligned} \mathbf{E} \left\| \int_{t-1/n}^t S(t-s)g_n(s) dW(s) \right\|^q &= \mathbf{E} \left\| \int_0^t S(t-s)g_n(s) \mathbf{1}_{[t-1/n, t]}(s) dW(s) \right\|^q \\ &\leq C'_{\text{Conv}} T^{q/2} \mathbf{E} \int_0^t \|g_n(s) \mathbf{1}_{[t-1/n, t]}(s)\|^q ds \\ &\leq C'_{\text{Conv}} T^{q/2} \left( \mathbf{E} \int_0^t \|g_n(s)\|^p ds \right)^{q/p} \left( \frac{1}{n} \right)^{(p-q)/p} \\ &\leq C'_{\text{Conv}} T^{q/2} \left( 2^{q-1} \mathbf{E} \int_0^t (1 + \|g_n(s)\|)^p ds \right)^{q/p} \left( \frac{1}{n} \right)^{(p-q)/p} \\ &\rightarrow 0 \text{ as } n \rightarrow \infty, \end{aligned}$$

where  $C'_{\text{Conv}}$  is the constant of stochastic convolution associated with  $q$ . Thus  $Z_n(t)$  converges to 0 in probability. Thus, for every  $t \in [0, T]$ , we have  $Z_\infty(t) = 0$  a.s. As  $Z_\infty$  is continuous, this means that  $Z_\infty = 0$  a.s. Thus  $X_\infty$  is a weak mild solution to (8).  $\square$

**Proof of Proposition 2.2.** Taking into account Lemma 3.5, and following the reasoning of the first part of the proof of Theorem 2.1, we see that the sequence  $(X_n)$  provided by the Tonelli scheme is relatively compact in  $\mathcal{N}_c^p(\mathfrak{F}, [0, T]; \mathbb{H})$ . Moreover, the limit of any convergent subsequence of  $(X_n)$  is a strong mild solution to (8).  $\square$

If  $\mathbb{E}$  is a Banach space, let us call *selection of  $\mathfrak{K}_c(\mathbb{E})$*  every mapping  $\mathfrak{S} : \mathfrak{K}_c(\mathbb{E}) \rightarrow \mathbb{E}$  such that  $\mathfrak{S}(K) \in K$  for every  $K \in \mathfrak{K}_c(\mathbb{E})$ .

**Proof of Proposition 2.3.** If  $\mathbb{L}$  is finite dimensional, it is well-known that there exists a selection  $\mathfrak{S}_{\mathbb{L}}$  of  $\mathfrak{K}_c(\mathbb{L})$  which is Lipschitz with respect to the Hausdorff distance. One such mapping is the Steiner point (see e.g. [45, 46, 47, 48]). In this case, we can define the Carathéodory mapping  $\mathfrak{g}$  of the proof of Theorem 2.1 by  $\mathfrak{g}(s, x) = \mathfrak{S}_{\mathbb{L}}(G(s, x))$ . Then the selection  $\mathfrak{g}$  satisfies all hypothesis satisfied by  $G$ , in particular (HFG)-(ii). Similarly, if  $\mathbb{H}$  is finite dimensional, we can define  $\mathfrak{f}$  as the Steiner point of  $F$ . So, the proof of Corollary 2.3 reduces to the case when  $F$  and  $G$  are single valued, and, in this case, from Proposition 2.2, there exists a strong mild solution to (8).  $\square$

**Remark 4.1** It is well known that, if  $\mathbb{L}$  is infinite dimensional, there exists no Lipschitz selection of  $\mathfrak{K}_c(\mathbb{L})$ : see [49, Theorem 4], where the reasoning given for the set of convex bounded sets also applies to the set of convex compact sets, and see also [50].

**Acknowledgements** We thank Professor Jan Seidler for several constructive remarks and for making us know the papers [30, 31, 32, 33, 34].

## References

- [1] Ahmed, N. U. Semigroup Theory with Applications to Systems and Control. No. 246 in Pitman Research Notes in Math. Harlow: Longman, **1991**.
- [2] Da Prato, Giuseppe; Zabczyk, Jerzy. Stochastic Equations in Infinite Dimensions. No. 44 in Encyclopedia of Mathematics and Its Applications. Cambridge: Cambridge University Press, **1992**.
- [3] Bogachev, Vladimir I. Deterministic and stochastic differential equations in infinite-dimensional spaces. Acta Applicandae Mathematicae **1995**. *40*, 25–93.
- [4] Dieudonné, J. Deux exemples singuliers d'équations différentielles. Acta Sci. Math. Szeged **1950**. *12*(Leopoldo Fejer et Frederico Riesz LXX annos natis dedicatus, Pars B), 38–40.
- [5] Godunov, A. N. A counterexample to Peano's theorem in infinite-dimensional Hilbert space. Vestnik Moskov. Univ. Ser. I Mat. Mekh. **1972**. *5*, 31–34. English transl. in Moscow Univ. Math. Bull. 29 (1974).
- [6] Godunov, A. N. Peano's theorem in infinite-dimensional Hilbert space fails even in a weakened formulation. Mat. Zametki **1974**. *15*(5), 476–477. English transl. in Math. USSR Notes 15 (1974).
- [7] Kibenko, A. V.; Krasnosel'skiĭ, M. A.; Mamedov, Ja. D. One-sided estimates in the conditions for the existence of solutions of differential equations in function spaces. Azerbaïdžan. Gos. Univ. Učen. Zap. Ser. Fiz.-Mat. i Him. Nauk **1961**. *1961*(3), 13–19.

- [8] Akhmerov, R. R.; Kamenskiĭ, M. I.; Potapov, A. S.; Rodkina, B. N.; Sadovskii, B. N. Measures of Noncompactness and Condensing Operators. No. 55 in *Operator Theory: Advances and Applications*. Basel, Boston, Berlin: Birkhäuser, **1992**. Revised and translated edition from the Russian 1986 original.
- [9] Ugowski, H. An application of measures of noncompactness to evolution equations. In *Differential equations and applications, I, II* (Russian) (Ruse, 1985). Ruse: ‘Angel Kanchev’ Tech. Univ., **1987**, 979–982.
- [10] Szufła, Stanisław. On the application of measure of noncompactness to existence theorems. *Rend. Sem. Mat. Univ. Padova* **1986**. 75, 1–14.
- [11] Banaś, Józef. Applications of measures of noncompactness to various problems. *Zeszyty Nauk. Politech. Rzeszowskiej Mat. Fiz.* **1987**. (5), 115.
- [12] Obukhovskiy, V. V. Semilinear functional-differential inclusions in a Banach space and controlled parabolic systems. *Avtomatika* **1991**. (3), 73–81, 97.
- [13] Hu, Shouchuan; Papageorgiou, Nikolas S. Handbook of multivalued analysis. Vol. II, vol. 500 of *Mathematics and its Applications*. Dordrecht: Kluwer Academic Publishers, **2000**. Applications.
- [14] Skorokhod, A. V. *Studies in the theory of random processes*. Translated from the Russian by Scripta Technica, Inc. Addison-Wesley Publishing Co., Inc., Reading, Mass., **1965**.
- [15] Rodkina, A. E. On existence and uniqueness of solution of stochastic differential equations with heredity. *Stochastics* **1984**. 12(3-4), 187–200.
- [16] Tubaro, L. An estimate of Burkholder type for stochastic processes defined by the stochastic integral. *Stochastic Anal. Appl.* **1984**. 2(2), 187–192.
- [17] Kotelenetz, Peter. A stopped Doob inequality for stochastic convolution integrals and stochastic evolution equations. *Stochastic Anal. Appl.* **1984**. 2(3), 245–265.
- [18] Da Prato, Giuseppe; Zabczyk, Jerzy. A note on stochastic convolution. *Stochastic Anal. Appl.* **1992**. 10(2), 143–153.
- [19] Hausenblas, Erika; Seidler, Jan. A note on maximal inequality for stochastic convolutions. *Czechoslovak Math. J.* **2001**. 51(126)(4), 785–790.
- [20] Da Prato, G.; Frankowska, H. A stochastic Filippov Theorem. *Stochastic Anal. Appl.* **1994**. 12(4), 409–426.
- [21] Jacod, Jean; Mémin, Jean. Weak and strong solutions of stochastic differential equations: existence and stability. In *Stochastic integrals* (Proc. Sympos., Univ. Durham, Durham, 1980), vol. 851 of *Lecture Notes in Math.*. Berlin: Springer, **1981**, 169–212.
- [22] Kamenskiĭ, M. I. On the Peano theorem in infinite-dimensional spaces. *Mat. Zametki* **1972**. 11, 569–576.
- [23] Kamenskiĭ, M. I.; Obukhovskii, Valeri; Zecca, Pietro. Condensing Multivalued Maps and Semilinear Differential Inclusions in Banach Spaces. No. 7 in *de Gruyter Series in Nonlinear Analysis and Applications*. Berlin: de Gruyter, **2001**.

- [24] Castaing, Charles; Valadier, Michel. Convex Analysis and Measurable Multifunctions. No. 580 in Lecture Notes in Math. Berlin: Springer Verlag, **1977**.
- [25] Valadier, Michel. Désintégration d'une mesure sur un produit. C. R. Acad. Sci. Paris Sér. I **1973**. *276*, A33–A35.
- [26] Balder, Erik. On Prohorov's theorem for transition probabilities. Sémin. Anal. Convexe **1989**. *19*, 9.1–9.11.
- [27] Castaing, Charles; Raynaud de Fitte, Paul; Valadier, Michel. Young measures on Topological Spaces. With Applications in Control Theory and Probability Theory. Dordrecht: Kluwer Academic Publishers, **2004**.
- [28] Valadier, Michel. A course on Young measures. Rendiconti dell'istituto di matematica dell'Università di Trieste **1994**. *26, suppl.*, 349–394. Workshop di Teoria della Misura et Analisi Reale Grado, 1993 (Italia).
- [29] Balder, Erik J. Lectures on Young measure theory and its applications in economics. Rend. Istit. Mat. Univ. Trieste **2000**. *31, suppl.*, 1–69. Workshop di Teoria della Misura et Analisi Reale Grado, 1997 (Italia).
- [30] Yamada, Toshio. On the successive approximation of solutions of stochastic differential equations. J. Math. Kyoto Univ. **1981**. *21(3)*, 501–515.
- [31] Manthey, Ralf. Convergence of successive approximation for parabolic partial differential equations with additive white noise. Serdica **1990**. *16(3-4)*, 194–200.
- [32] Taniguchi, Takeshi. Successive approximations to solutions of stochastic differential equations. J. Differential Equations **1992**. *96(1)*, 152–169.
- [33] Barbu, D. Local and global existence for mild solutions of stochastic differential equations. Portugal. Math. **1998**. *55(4)*, 411–424.
- [34] Barbu, Dorel; Bocşan, Gheorghe. Approximations to mild solutions of stochastic semilinear equations with non-Lipschitz coefficients. Czechoslovak Math. J. **2002**. *52(127)(1)*, 87–95.
- [35] Hartman, Philip. Ordinary differential equations. S. M. Hartman, Baltimore, Md., **1973**. Corrected reprint.
- [36] Pellaumail, Jean. Solutions faibles et semi-martingales. In Azéma, Jacques; Yor, Marc, eds., Séminaire de Probabilités XV, 1979/80. Université de Strasbourg, vol. 850 of *Lecture Notes in Math.*. Berlin: Springer Verlag, **1981**, 561–586.
- [37] Pellaumail, J. Weak solutions for semimartingales. Canad. J. Math. **1981**. *33(5)*, 1165–1181.
- [38] Dudley, R. M. Real Analysis and Probability. Cambridge: Cambridge University Press, **2002**.
- [39] Michael, Ernest. Topologies on spaces of subsets. Trans. Amer. Math. Soc. **1951**. *71*, 152–182.

- [40] Christensen, J. P. R. *Topology and Borel structure*. Amsterdam, London: North-Holland, **1974**.
- [41] Aldous, David J. Stopping times and tightness. *Ann. Probab.* **1978**. *6*, 335–340.
- [42] Jacod, Jean. Théorèmes limite pour les processus. In Hennequin, P. L., ed., *École d'Été de Probabilités de Saint-Flour XIII-1983*, no. 1117 in *Lecture Notes in Math.* Berlin: Springer Verlag, **1985**, 298–409.
- [43] Balder, E. J. New sequential compactness results for spaces of scalarly integrable functions. *J. Math. Anal. Appl.* **1990**. *151*, 1–16.
- [44] Kucia, Anna. Some results on Carathéodory selections and extensions. *J. Math. Anal. Appl.* **1998**. *223*(1), 302–318.
- [45] Schneider, Rolf. On Steiner points of convex bodies. *Israel J. Math.* **1971**. *9*, 241–249.
- [46] Saint Pierre, Jean. Point de Steiner et sections lipschitziennes. *Sém. Anal. Convexe* **1985**. *15*, 7.1–7.42.
- [47] Vitale, Richard A. The Steiner point in infinite dimensions. *Israel J. Math.* **1985**. *52*(3), 245–250.
- [48] Przesławski, Krzysztof. Centres of convex sets in  $L^p$  metrics. *J. Approx. Theory* **1996**. *85*(3), 288–296.
- [49] Posicel'skiĭ, E. D. Lipschitzian mappings in the space of convex bodies. *Optimizacija* **1971**. (4(21)), 83–89. (in Russian).
- [50] Przesławski, Krzysztof; Yost, David. Continuity properties of selectors and Michael's theorem. *Michigan Math. J.* **1989**. *36*(1), 113–134.