DISTRIBUTION-DEPENDENT STOCHASTIC DIFFERENTIAL DELAY EQUATIONS IN FINITE AND INFINITE DIMENSIONS

RICO HEINEMANN*

Abstract. We prove that distribution dependent (also called McKean–Vlasov) stochastic delay equations of the form

$$\mathrm{d}X(t) = b(t, X_t, \mathcal{L}_{X_t})\mathrm{d}t + \sigma(t, X_t, \mathcal{L}_{X_t})\mathrm{d}W(t)$$

have unique (strong) solutions in finite as well as infinite dimensional state spaces if the coefficients fulfill certain monotonicity assumptions.

Introduction

The aim of this paper is to study the existence and uniqueness of Distribution-Dependent Stochastic Differential Delay Equations (DDSDDE's) in finite and infinite dimensional state spaces in the variational framework. A DDSDDE has the form

$$\mathrm{d}X(t) = b(t, X_t, \mathcal{L}_{X_t})\mathrm{d}t + \sigma(t, X_t, \mathcal{L}_{X_t})\mathrm{d}W(t),$$

where W is a standard $\mathbb{R}^d$-valued Wiener process in the finite dimensional case and a cylindrical Q-Wiener process with $Q = I$ in a separable Hilbert space in the infinite dimensional case. X_t denotes the *delay* or *segment* of X at time t. X_t takes values in a path-space and is defined as $X_t(\theta) := X(t + \theta)$, $\theta \in [-r_0, 0]$, whereby $r_0 > 0$ is fixed. $\mathcal{L}_{X_t}$ denotes the *law* of X_t.
Recently there has been an increasing interest in this type of equations as well as in classical distribution-dependent SDE's (DDSDE's) - also referred to as McKean-Vlasov SDEs - , i.e. equations of the form

$$\mathrm{d}X(t) = b(t, X(t), \mathcal{L}_{X(t)})\mathrm{d}t + \sigma(t, X(t), \mathcal{L}_{X(t)})\mathrm{d}W(t),$$

see for instance [7], [10], [11], [12], [13], [17], [18], [20], [3], [25] or [26] as well as the references therein. Clearly, SDDE's can be viewed as a sub-class of DDSDDE's.
A first existence and uniqueness result under monotonicity conditions for distribution-dependent SDE's without delay was published by Wang in 2018 (see reference [26]). Wang's idea was carried over to the case with delay by Huang, Röckner and Wang in [12]. [12] and [26] are the main reference for the second chapter of this paper.

2010 *Mathematics Subject Classification.* 60H10, 35K55.
Key words and phrases. Nonlinear PDE for probability measures, McKean-Vlasov SDEs in infinite dimensions, Wasserstein distance, stochastic delay equations, variational approach.
*Faculty of Mathematics, Bielefeld University, Universitätsstrasse 25, 33615 Bielefeld, Germany
Email: rico.heinemann@t-online.de.

A main motivation to study solutions of DDSDDE's is their relation to solutions of *non-linear Fokker-Planck Kolmogorov equations* (FPKE's). Whenever coefficients b and σ are given one can define the following differential operator from $C_0^\infty(\mathbb{R}^d)$ to the set of all Borel-measurable real-valued functions on $\mathcal{C}$:

$$\left(L_{t,\mu}f\right)(\xi) := \sum_{i=1}^{d} b_i(t,\xi,\mu)(\partial_i f)(\xi(0)) + \frac{1}{2}\sum_{i,j=1}^{d}(\sigma\sigma^*)_{i,j}(t,\xi,\mu)(\partial_i\partial_j f)(\xi(0)),$$

$t \geq 0$, $\mu \in \mathcal{P}_2(\mathcal{C})$, $f \in C_0^\infty(\mathbb{R}^d)$ and $\xi \in \mathcal{C} := C([-r_0,0];\mathbb{R}^d)$, where $\mathcal{P}_2(\mathcal{C})$ denotes the set of all probability measues on $\mathbb{R}^d$ with finite second moments. By Itô's formula one can show that if $(X(t))_{t\geq -r_0}$ is a solution of our DDSDDE, $\mu_t := \mathcal{L}_{X_t}$ solves the corresponding FPKE

$$\partial\mu(t) := L_{t,\mu_t}^*\mu_t,$$

where $\mu(t) := \mathcal{L}_{X(t)}$ = law of $X(t)$. Here, we call a continuous mapping $\mu : \mathbb{R}_+ \to \mathcal{P}_2(\mathcal{C})$ a solution of the FPKE, if

$$\int_0^t \int_{\mathcal{C}} |L_{s,\mu_s}f| \mathrm{d}\mu_s \mathrm{d}s < \infty,$$

and

$$\int_{\mathbb{R}^d} f \mathrm{d}\mu(t) = \int_{\mathbb{R}^d} f \mathrm{d}\mu(0) + \int_0^t \int_{\mathcal{C}} \left(L_{\mu_s,s}f\right) \mathrm{d}\mu_s \mathrm{d}s$$

for all $t \geq 0$ and $f \in C_0^\infty(\mathbb{R}^d)$. For more details of the relation between Fokker-Planck equations and DDSDDE's see for instance [3] or [12]. Since this paper focuses on the existence and uniqueness of DDSDDE's, we are not going to further investigate this relation and just note that an existence and uniqueness result for FPKE's could be deduced from our existence and uniqueness result in Chapter 2 as in [12, Chapter 2]. For more information about FPKE see for instance [6].

A first result for the existence and uniqueness of solutions to DDSDDE's in finite dimensions was proved by Huang, Röckner and Wang in 2017 (see [12]). The main novelty of this paper, compared to [12], is that we also prove an existence and uniqueness result in infinite dimensions (Theorem 3.1.3), i.e. we replace $\mathbb{R}^d$ in [12] by a separable real Hilbert space. To be able to do this we prove another finite dimensional result (Theorem 2.1.6) under assumptions which are better suited for the generalization to infinite dimensions as the conditions presented in [12]. Moreover, our proof of the finite dimensional result replaces the iteration procedure in [12] by a fixed point argument, which turns out to be technically easier and more conceptual.

Next let us give a brief overview of the content of this paper. The first chapter introduces some tools and notations, which are necessary to understand this paper. In the second chapter, we prove that for every initial condition $\psi \in \mathcal{C} := C([-r_0,0];\mathbb{R}^d)$ a DDSDDE has a unique solution, if certain monotonicity, coercivity, growth and continuity assumptions are fulfilled (Theorem 2.1.6). To be able to do this, we define precisely what a *solution* is (see Definition 2.1.1 and Definition 2.1.3). While the main idea of our proof is similar to the proof in [12], i.e. we deduce existence of a solution to DDSDDE's from a result for stochastic differential delay equations (SDDE's), we assume different conditions on the coefficients (see (H1) to (H4) in Chapter 3), which are better suited for the infinite dimensional case and use Banach's fixed-point theorem instead of *iterating in distribution*, i.e. approximating the solution to the DDSDDE by solutions of SDDE's. The conditions on b and σ are chosen in such a way, that given any fixed continuous, adapted, $\mathbb{R}^d$-valued process $(X(t))_{t\geq -r_0}$ with $\mathbb{E}\left[\sup_{t\in[-r_0,T]}|X(t)|^2\right]<\infty$ for all $T>0$, the classical SDDE

$$\mathrm{d}\left(\Lambda X\right)(t)=b(t,(\Lambda X)_t,\mathcal{L}_{X_t})\mathrm{d}t+\sigma(t,(\Lambda X)_t,\mathcal{L}_{X_t})\mathrm{d}W(t),$$

has a unique solution ΛX, fulfilling the initial condition $\Lambda X(0)=\psi$ for given $\psi\in\mathcal{C}$. It is clear, that X is a solution of our DDSDDE, if $\Lambda X=X$. Using the Banach fixed-point theorem, we show that there exists exactly one X, such that $\Lambda X=X$ (see Lemma 2.2.1 and Lemma 2.2.2). In addition to existence and path-wise uniqueness, we also prove weak uniqueness. The weak uniqueness is derived from the Yamada-Watanabe Theorem for SDDE's.

The third chapter contains the main novelty of this paper, as we prove an existence and uniqueness result for DDSDDE's in infinite dimensions. That is, we replace $\mathbb{R}^d$ by a separable Hilbert space H, more precisely an appropriate Gelfand triple (V,H,V^*). Chapter 3 is an extension of the fourth chapter in [16] to DDSDDE's, i.e. we work in the *variational framework* and use a *Galerkin approximation* to deduce the infinite dimensional result from the finite dimensional result.

1. Preliminaries

This chapter introduces some notations and results needed for the formulation and understanding of the rest of this paper, like the Wasserstein distance and an existence and uniqueness result for SDDE's. All results are given without proof, since they are not the actual topic of this paper.
In addition to contents of this chapter, knowledge about measure and integration theory (c.f. [4]), functional analysis (c.f. [1] or [27]), probability theory (c.f. [5], [9] or [22]), stochastic integration theory (c.f. [16] or [24]) as well as stochastic differential equations (c.f. [8] or [16]), is necessary to understand this paper.

1.1. **Notations.** First of all, let us fix some notations. As usual we denote $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$ for the set of all natural, rational and real numbers, respectively. For $d\in\mathbb{N}$, $\mathbb{R}^d$ denotes the d-dimensional euclidean space, $\langle\cdot,\cdot\rangle$ the inner product and $|\cdot|$ the corresponding norm. If $m\in\mathbb{N}$ is another natural number, $\mathbb{R}^{d\times m}$ denotes the space of all $d\times m$-matrices. If A is an arbitrary set, we write 1_A for its indicator function. For $s,t\in\mathbb{R}$ we define $s\vee t:=\max(s,t)$ and $s\wedge t:=\min(s,t)$. Like usual, for $a,b\in\bar{\mathbb{R}}:=\mathbb{R}\cup\{-\infty\}\cup\{+\infty\}$, $(a,b):=\{x\in\bar{\mathbb{R}}:\ a<x<b\}$ denotes the open interval, $[a,b]$ denotes the closed interval, $[a,b)$ denotes the left-closed interval and $(a,b]$ denotes the left-open interval. We call $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq -r_0},P)$ a *stochastic basis*, if $(\Omega,\mathcal{F},P)$ is a *complete probability space* and $(\mathcal{F}_t)_{t\geq -r_0}$ is a *normal filtration*. If

$(X, \|\cdot\|_X)$ is a Banach space, we denote X^* for the dual space of X and $\mathcal{B}(X)$ for the Borel sigma-algebra on X. For $x \in X$ and $x^* \in X^*$ we define ${}_{X^*}\langle x^*, x\rangle_X = x^*(x)$ as the *dualization* between X and X^*. I_X or, if it is clear on which space we are working, I denotes the identity operator. If X and Y are Banach spaces, we denote $L(X, Y)$ for the Banach space of all bounded linear operators from X to Y equipped with the standard operator norm. Moreover, if U and H are Hilbert spaces we denote $L_2(U, H)$ for the space of all Hilbert-Schmidt operators from U to H equipped with the usual norm $\|T\|_{\mathrm{HS}} = \left(\sum_{n=1}^{\infty} \|Te_n\|_H^2\right)^{\frac{1}{2}}$.

1.1.1. *Path spaces.* As we will see in section 1.3 as well in chapter 2 and 3, the coefficients of a stochastic delay equation are defined on path spaces, i.e. spaces of functions. This subsection introduces the spaces of functions which are needed in this paper.
Throughout this paper, whenever $(X, \|\cdot\|_X)$ is a Banach space, $p \geq 2$ and $r_0 > 0$ fixed, we use the following notations:
If (E, d) is a metric space, $C(E; X)$ denotes, like usual, the set of all continuous functions from E to X. If $(S, \mathcal{A}, \mu)$ is a measure space, $L^p(S, \mathcal{A}, \mu; X)$ denotes the usual L^p-space (c.f. [4], [27, Chapter V.5] or [16, Appendix A]). If it is clear which sigma algebra $\mathcal{A}$ or which measure μ is used, we might for simplicity just denote $L^p(S, \mu; X)$, $L^p(S, \mathcal{A}; X)$ or $L^p(S; X)$, respectively. In the case that $S \in \mathcal{B}(\mathbb{R}^d)$, it is always $\mathcal{A} = \mathcal{B}(S)$ and μ is the Lebesgue measure. $\mathcal{C}(X) := C([-r_0, 0]; X)$ equipped with the uniform norm $\|\xi\|_\infty := \sup_{\theta \in [-r_0, 0]} \|\xi(\theta)\|_X$, $\mathcal{C}_\infty(X) = C([-r_0, \infty); X)$, equipped with the metric $d(\xi, \eta) := \sum_{k \in \mathbb{N}} 2^{-k} \big(\sup_{t[-r_0, k]} \|\xi(t) - \eta(t)\|_X \wedge 1\big)$, $L^p_X := L^p([-r_0, 0]; X)$ equipped with the standard L^p-norm $\|\xi\|^p_{L^p_X} := \int_{-r_0}^0 \|\xi(z)\|_X^p \mathrm{d}z$. In the case $X = \mathbb{R}^d$, for some $d \in \mathbb{N}$, we just write $\mathcal{C}, \mathcal{C}_\infty$ and L^p, respectively.

Moreover, whenever $(Y, \|\cdot\|_Y)$ is another Banach space with $X \subset Y$ continuously and $I \subset \mathbb{R}$ is an interval, we define

$$C(I;Y) \cap L^p(I;X) := \Big\{\xi \in C(I;Y) : \exists \bar{\xi} : I \to X \ \mathcal{B}(I)/\mathcal{B}(X)\text{-measurable such that } \bar{\xi} = \xi \ \mathrm{d}t - a.e. \text{ and } \int_I \left\|\bar{\xi}(t)\right\|^p \mathrm{d}t < \infty\Big\}$$

and

$$C(I;Y) \cap L^p_{loc}(I;X) := \Big\{\xi \in C(I;Y) : \exists \bar{\xi} : I \to X \text{ such that } \bar{\xi} = \xi \ \mathrm{d}t - a.e. \text{ and } \int_{I'} \left\|\bar{\xi}(t)\right\|^p \mathrm{d}t < \infty \ \forall I' \subset I \text{ compact.}\Big\}.$$

Obviously $C(I;Y) \cap L^p(I;X) = C(I;Y) \cap L^p_{loc}(I;X)$ if I is compact and $E := L^p(I, Y)$, $C(I;Y) \cap L^p(I;X)$ is a Banach space under the norm $\|\cdot\|_{C(I;Y) \cap L^p(I;X)} := \|\cdot\|_{C(I;Y)} + \|\cdot\|_{L^p(I;X)}$, in the case that I is compact.

1.1.2. *Segments of functions.* The main difference between stochastic delay differential equations and classical SDE's is - as the name already suggests - that the coefficients depend on the *delay* of X at time t instead of the value of X at time t. Therefore we have to define precisely what the *segment* or *delay* of a function is. We do this similar to similar to [12, Chapter 2]. Let X be a Banach space and $r_0 > 0$ fixed. For $t \geq 0$ define the map

$$\pi_t \colon \mathcal{C}_\infty(X) \longrightarrow \mathcal{C}(X)$$

by $(\pi_t f)(\theta) = f(t+\theta)$, $\theta \in [-r_0, 0]$. In the following we will denote $f_t := \pi_t f$, for $t \geq 0$.

Remark 1.1.1. Note that $[0,\infty) \ni t \mapsto f_t$ is an element in $C([0,\infty), \mathcal{C}(X))$.

1.2. **p-th order probability measures and Wasserstein distance.** Since we want to study stochastic differential equations where the coefficients also depend on the distribution of the solution, we need to have a measure for the distance between to probability measures to be able to formulate monotonicity assumptions on our coefficients. The *Wasserstein distance* is the most important tool to do that. The main references for this section are [2] and [23]. Throughout this section, let $(X, \|\cdot\|_X)$ be a separable Banach space, $\mathcal{B}(X)$ the Borel sigma-algebra on X and $\mathcal{P}(X)$ the set of all probability measures on $(X, \mathcal{B}(X))$.

Definition 1.2.1. Let $p \geq 1$. *The class of probability measures of p-th order* is defined as

$$\mathcal{P}_p(X) := \left\{ \mu \in \mathcal{P}(X) : \quad \mu(\|\cdot\|_X^p) := \int_X \|x\|_X^p \mu(\mathrm{d}x) < \infty \right\}.$$

On $\mathcal{P}_p(X)$ we can define the following metric:

Definition 1.2.2. For $\mu, \nu \in \mathcal{P}_p(X)$ define *the p-th Wasserstein distance* as

$$\mathbb{W}_p^X(\mu,\nu) := \inf_{\gamma \in \Gamma(\mu,\nu)} \left(\int_{X\times X} \|x-y\|_X^p \gamma(\mathrm{d}x, \mathrm{d}y) \right)^{\frac{1}{p}}.$$

Here $\Gamma(\mu,\nu)$ denotes the set of all couplings of μ and ν, e.g.

$$\Gamma(\mu,\nu) := \left\{ \gamma \in \mathcal{P}(X \times X) : \quad \gamma \circ \pi_x^{-1} = \mu \text{ and } \gamma \circ \pi_y^{-1} = \nu \right\},$$

where $\pi_x(x,y) := x$ and $\pi_y(x,y) := y$, $(x,y) \in X \times X$, are the standard projections. (Here $X \times X$ is equipped with the σ-field generated by the projections.)

Proposition 1.2.3. $(\mathcal{P}_p(X), \mathbb{W}_p^X)$ *is a polish space, e.g. a separable, complete metric space.*

Proof. See [2, Proposition 7.1.5] □

1.3. **Spaces of measure-valued functions.** As we will see in the next chapter, the coefficients of our stochastic equation are defined on a set of probability measures. Therefore, to be able to formulate the conditions on the coefficients in the next chapters, we need to introduce spaces of measure-valued functions.
Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ and $(E, \|\cdot\|_E)$ be Banach spaces such that

$$X \subset Y \subset E$$

continuously and densely. For $p \geq 2$ we define the following spaces of measures and measure-valued functions:

$$\mathcal{P}_2(\mathcal{C}(Y)) \cap \mathcal{P}_p(L^p_X) := \Big\{\mu \in \mathcal{P}(L^p_E) : \ \mu(\mathcal{C}) = \mu(L^p_X) = 1 \text{ and } \mu\left(\|\cdot\|^2_\infty\right), \mu\big(\|\cdot\|^p_{L^p_X}\big) < \infty\Big\}.$$

Note that this set is well defined since $\mathcal{C}(Y) \subset L^p_E$ and $L^p_X \subset L^p_E$ continuous and hence by Kuratovski's theorem ([14, Theorem 15.1] or [19]) $\mathcal{C}(Y), L^p_X \in \mathcal{B}\left(L^p_E\right)$. Clearly $\mathcal{P}_2(\mathcal{C}(Y)) \cap \mathcal{P}_p(L^p_X)$ is a metric space with respect to the metric $d := \mathbb{W}_2^{\mathcal{C}(Y)} + \mathbb{W}_p^{L^p_X}$. Define

$$C\left([0,\infty); \mathcal{P}_2(\mathcal{C})\right) \cap L^p_{\text{loc}}\left([0,\infty); \mathcal{P}_p\left(L^p\right)\right) := \Big\{\mu : [0,\infty) \to \mathcal{P}_2(\mathcal{C}(Y)) \cap \mathcal{P}_p(L^p_X) : \mu : [0,\infty) \to \mathcal{P}_2(\mathcal{C}(Y)) \text{ is continuous and } \int_0^t \mu_s(\|\cdot\|^p_{L^p})\mathrm{d}s < \infty \ \forall t \geq 0\Big\}.$$

Note that if $\mu \in C\left([0,\infty); \mathcal{P}_2(\mathcal{C})\right) \cap L^p_{\text{loc}}\left([0,\infty); \mathcal{P}_p\left(L^p\right)\right)$, then μ is $\mathcal{B}([0,\infty))/\mathcal{B}\left(\mathcal{P}_2(\mathcal{C}(Y)) \cap \mathcal{P}_p(L^p_X)\right)$-measurable, since $\mu : [0,\infty) \to \mathcal{P}_2(\mathcal{C}(Y)) \supset \mathcal{P}_2(\mathcal{C}(Y)) \cap \mathcal{P}_p(L^p_X)$ is continuous and thereby $\mathcal{B}([0,\infty))/\mathcal{B}(\mathcal{P}_2\left(\mathcal{C}(Y)\right)) \cap \mathcal{P}_p(L^p_X)$-measurable and $\mathcal{B}(\mathcal{P}_2\left(\mathcal{C}(Y)\right)) \cap \mathcal{P}_p(L^p_X) = \mathcal{B}(\mathcal{P}_2\left(\mathcal{C}(Y)\right) \cap \mathcal{P}_p(L^p_X))$ by Kuratowski's theorem. Define

$$C\left([0,T]; \mathcal{P}_2(\mathcal{C})\right) \cap L^p\left([0,T]; \mathcal{P}_p\left(L^p\right)\right) := \Big\{\mu : [0,T]) \to \mathcal{P}_2(\mathcal{C}(Y)) \cap \mathcal{P}_p(L^p_X) : \mu : [0,T] \to \mathcal{P}_2(\mathcal{C}(Y)) \text{ is continuous and } \int_0^T \mu_s(\|\cdot\|^p_{L^p})\mathrm{d}s < \infty\Big\},$$

where $T > 0$ is fixed. With the same argument as above, $\mu \in C\left([0,T]; \mathcal{P}_2(\mathcal{C})\right) \cap L^p\left([0,T]; \mathcal{P}_p\left(L^p\right)\right)$ is $\mathcal{B}([0,T])/\mathcal{B}\left(\mathcal{P}_2(\mathcal{C}(Y)) \cap \mathcal{P}_p(L^p_X)\right)$-measurable.

2. Distribution-Dependent SDE's with delay in finite dimensions

The aim of this chapter is to solve the following delay-distribution dependent SDE in $\mathbb{R}^d$:

$$\mathrm{d}X(t) = b(t, X_t, \mathcal{L}_{X_t})\mathrm{d}t + \sigma(t, X_t, \mathcal{L}_{X_t})\mathrm{d}W(t), \tag{2.1}$$

where $W = (W(t))_{t \geq 0}$ is a d-dimensional Brownian motion, $\in \mathbb{N}$, defined on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq -r_0}, P)$, with $r_0 > 0$ fixed and

$$b \colon [0,\infty) \times \mathcal{C} \times (\mathcal{P}_2(\mathcal{C}) \cap \mathcal{P}_p(L^p)) \longrightarrow \mathbb{R}^d;$$
$$\sigma \colon [0,\infty) \times \mathcal{C} \times (\mathcal{P}_2(\mathcal{C}) \cap \mathcal{P}_p(L^p)) \longrightarrow \mathbb{R}^{d \times d}$$

$\mathcal{B}([0,\infty) \otimes \mathcal{B}(\mathcal{C}) \otimes \mathcal{B}(\mathcal{P}_2(\mathcal{C}) \cap \mathcal{P}_p(L^p))$-measurable, whereby $p \geq 2$ is fixed.

The main difficulty, compared to the well-known, classical SDE's (c.f. [15] or [16]), that has to be overcome to get an existence and uniqueness result, is to deal with the delay and the distribution dependence. To achieve such a result we first formulate certain conditions on the coefficients b and σ and define precisely what a *solution* of (2.1) is. Afterwards we are going to prove existence and uniqueness of solutions to (2.1). The main inspiration for our proof comes from [12], i.e. the existence of solutions to (2.1) is derived from an existence and uniqueness result about SDDE's. But unlike in [12], we use the Banach fixed point theorem instead of an *iteration in distribution* and use [21, Theorem 4.2] instead of [25, Corollary 4.1.2] to the show the existence and uniqueness of SDDE's, because our conditions on the coefficients differ from those in [12].

2.1. **Conditions on the coefficients and main result.** To show existence and uniqueness of solutions to (2.1), we fix $p \geq 2$ and assume that the coefficients b and σ fulfill the following conditions. For simplicity we write $\mathbb{W}_2$ instead of $\mathbb{W}_2^{\mathcal{C}}$ for the Wasserstein distance on $\mathcal{P}_2(\mathcal{C})$ and use the notations introduced in 1.1 and 1.2.

(H1) (Continuity) For every $t \geq 0$, $b(t, \cdot, \cdot)$ and $\sigma(t, \cdot, \cdot)$ are continuous on $\mathcal{C} \times (\mathcal{P}_2(\mathcal{C}) \cap \mathcal{P}_p(L^p))$.

(H2) (Coercivity) There exists $\alpha\colon \mathbb{R}_+ \mapsto \mathbb{R}_+$ non-decreasing such that

$$\int_0^t 2\langle b(s, \xi_s, \mu_s), \xi(s)\rangle \mathrm{d}s \leq -\frac{1}{2}\int_0^t |\xi(s)|^p \mathrm{d}s + \alpha(t)\|\xi_0\|_{L^p}^p + \alpha(t)\int_0^t \left(1 + \|\xi_s\|_\infty^2 + \mu_s(\|\cdot\|_\infty^2)\right) \mathrm{d}s,$$

for all $t \geq 0$, $\xi \in \mathcal{C}_\infty$ and $\mu \in C\left([0,\infty); \mathcal{P}_2(\mathcal{C})\right) \cap L^p_{\mathrm{loc}}\left([0,\infty); \mathcal{P}_p\left(L^p\right)\right)$.

(H3) (Monotonicity) There exists $\beta\colon \mathbb{R}_+ \mapsto \mathbb{R}_+$, non-decreasing, such that

$$\int_0^t 2\langle b(s, \xi_s, \mu_s) - b(s, \eta_s, \nu_s), \xi(s) - \eta(s)\rangle \mathrm{d}s \leq \beta(t)\int_0^t \|\xi_s - \eta_s\|_\infty^2 + \mathbb{W}_2(\mu_s, \nu_s)^2 \mathrm{d}s + \beta(t)\|\xi_0 - \eta_0\|_{L^p}^p$$

and

$$\int_0^t \|\sigma(s, \xi_s, \mu_s) - \sigma(s, \eta_s, \nu_s)\|_{\mathrm{HS}}^2 \mathrm{d}s \leq \beta(t)\int_0^t \|\xi_s - \eta_s\|_\infty^2 + \mathbb{W}_2(\mu_s, \nu_s)^2 \mathrm{d}s + \beta(t)\|\xi_0 - \eta_0\|_{L^p}^p,$$

for all $t \geq 0$; $\xi, \eta \in \mathcal{C}_\infty$ and $\mu, \nu \in C\left([-r_0,\infty); \mathcal{P}_2(\mathcal{C})\right) \cap L^p_{\mathrm{loc}}\left([-r_0,\infty); \mathcal{P}_p\left(L^p\right)\right)$.

(H4) (Growth) b is bounded on bounded sets in $[0,\infty) \times \mathcal{C} \times (\mathcal{P}_2(\mathcal{C}) \cap \mathcal{P}_p(L^p))$, and there exists a non-decreasing function $\gamma\colon \mathbb{R}_+ \mapsto \mathbb{R}_+$ and some $q_0 \in \mathbb{N}$ such that

$$\int_0^t |b(s, \xi_s, \mu_s)|^{\frac{p}{p-1}} \mathrm{d}s \leq \gamma(t)\left(\int_0^t |\xi(s)|^p + \mu_s(\|\cdot\|_{L^p}^p)\mathrm{d}s + \|\xi_0\|_{L^p}^p\right)^{q_0} + \gamma(t)\left(1 + \sup_{s\in[0,t]}\|\xi_s\|_\infty^{2q_0} + \sup_{s\in[0,t]}\mu_s\left(\|\cdot\|_\infty^2\right)^{q_0}\right)$$

and

$$\|\sigma(t,\xi_t,\mu_t)\|_{\mathrm{HS}}^2 \leq \gamma(t)\left(1+\|\xi_t\|_\infty^2+\mu_t(\|\cdot\|_\infty^2)\right),$$

for all $t\geq 0$, $\xi\in\mathcal{C}_\infty$ and $\mu\in C\left([-r_0,\infty);\mathcal{P}_2(\mathcal{C})\right)\cap L^p_{\mathrm{loc}}\left([-r_0,\infty);\mathcal{P}_p\left(L^p\right)\right)$.

Let us briefly comment on these conditions. First of all, these conditions look similar to standard monotonicity and coercivity conditions, like they were for example formulated in [15] or [16]. The main difference is that, in order to deal with the delay and the distribution dependence, the sup-norm and the Wasserstein metric appear on the right hand side. Another difference is that the conditions are in integrated form, which, as we are going to discuss in further detail in section 2.3.2 of this chapter, will be helpful for the generalization to infinite dimensions in the next chapter. Moreover, in the case that b and σ are distribution independent, i.e. $b(t,\xi,\mu)=\bar{b}(t,\xi)$ and $\sigma(t,\xi,\mu)=\bar{\sigma}(t,\xi)$, the existence of a solution to (2.1) is ensured by [21, Theorem 4.2], because, as we will see in the proof of Lemma 2.2.1, for those b and σ (H1)-(H4) imply (H1)-(H5) in [21]. Note that the measurability of $s\mapsto b(s,\xi_s,\mu_s)$ and $s\mapsto\sigma(s,\xi_s,\mu_s)$, with ξ and μ as in the conditions is ensured by Remark 1.1.1 and the assumptions on ξ and μ. By (H4) all integrals in (H1)-(H3) are well-defined.

In the following we introduce different notions of solution to (2.1) and uniqueness of solutions.

Definition 2.1.1. A pair (X,W), where $X=(X(t))_{t\geq -r_0}$ is an $(\mathcal{F}_t)$-adapted, $\mathbb{R}^d$-valued process with continuous sample paths and W is a $\mathbb{R}^d$-valued, $(\mathcal{F}_t)$-Wiener process on a stochastic basis $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq -r_0},P)$ is called a *weak solution* of (2.1) with initial condition $\psi\in\mathcal{C}$ iff

(i)
$$\mathbb{E}[\|X_t\|_\infty^2]+\int_{-r_0}^t \mathbb{E}[|X(s)|^p]\mathrm{d}s<\infty, \tag{2.2}$$
for all $t\geq 0$;

(ii)
$$X(t)=X(0)+\int_0^t b(s,X_s,\mathcal{L}_{X_s})\mathrm{d}s+\int_0^t \sigma(s,X_s,\mathcal{L}_{X_s})\mathrm{d}W(s), \tag{2.3}$$
for all $t\geq 0$ P-a.s.; and

(iii)
$$X(t)=\psi(t), \tag{2.4}$$
for all $t\in[-r_0,0]$ P-a.s.

Remark 2.1.2. Note that (2.2) implies that for every weak solution $(X(t))_{t\geq -r_0}$ we have

$$\mathbb{E}\left[\int_0^T \|X_t\|_{L^p}^p\mathrm{d}t+\sup_{t\in[0,T]}\|X_t\|_\infty^2\right]<\infty\quad\forall\, T\geq 0$$

and $(X_t)_{t\in[0,T]}$ is a continuous $\mathcal{C}$-valued process. This, together with Lebesgues theorem, implies that $[0,\infty)\ni t\mapsto\mathcal{L}_{X_t}$ is a continuous map from $[0,\infty)$ to $(\mathcal{P}_2(\mathcal{C}),\mathbb{W}_2)$.

By Kuratowski's theorem and (2.2), this implies that $[0,\infty) \ni t \mapsto \mathcal{L}_{X_t}$ is $\mathcal{B}([0,\infty))/\mathcal{B}(\mathcal{P}_2(\mathcal{C}) \cap \mathcal{P}_p(L^p))$-measurable.
In particular, $(t,\omega) \mapsto b(t, X_t(\omega), \mathcal{L}_{X_t})$ and $(t,\omega) \mapsto \sigma(t, X_t(\omega), \mathcal{L}_{X_t})$ are progressively measurable maps. Thus the integrals on the right-hand side of (2.3) are well-defined.

Definition 2.1.3. We say (2.1) has a *(strong) solution* if for every stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq -r_0}, P)$ with a given $\mathbb{R}^d$-valued, $(\mathcal{F}_t)$-Wiener process W and given initial condition $\psi \in \mathcal{C}$, there exists a $(\mathcal{F}_t)$-adapted, continuous $\mathbb{R}^d$-valued process X such that X fulfills (2.2)-(2.4) in Definition 2.1.1.

The next definitions recall different notions of uniqueness (c.f. [21, Appendix E]).

Definition 2.1.4. We say that *weak uniqueness* holds for (2.7) if whenever (X, W) and $(\tilde{X}, \tilde{W})$ are weak solutions with stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq -r_0}, P)$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t\geq -r_0}, \tilde{P})$ such that

$$X_0 = \tilde{X}_0 = \psi,$$

for some $\psi \in \mathcal{C}$, then

$$P \circ X^{-1} = \tilde{P} \circ \tilde{X}^{-1}$$

as measures on $(\mathcal{C}_\infty, \mathcal{B}(\mathcal{C}_\infty))$.

Definition 2.1.5. We say that *path-wise uniqueness* holds for (2.7), if whenever (X, W) and $(\tilde{X}, W)$ are two weak solutions on the same stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq -r_0}, P)$ and with the same Wiener process W on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq -r_0}, P)$ such that $X_0 = \tilde{X}_0$ P-a.s., then

$$X(t) = \tilde{X}(t),$$

for all $t \geq 0$ P-a.s.

The next Theorem is the main result of this chapter and shows the existence of a unique strong solution as well as weak uniqueness.

Theorem 2.1.6. *Assume (H1)-(H4).*

(a) *For any $\psi \in \mathcal{C}$, (2.1) has a (pathwise) unique (strong) solution $(X(t))_{t\geq -r_0}$, fulfilling $X_0 = \psi$. Moreover*

$$\mathbb{E}\left[\sup_{t\in[-r_0,T]} |X(t)|^{2q}\right] < \infty, \tag{2.5}$$

for all $T > 0$ and $q \in \mathbb{N}$.

(b) *Whenever (X, W) and (Y, W) are weak solutions of (2.7) on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq -r_0}, P)$, we have*

(i)

$$\mathbb{E}\left[\sup_{t\in[-r_0,T]}|X(t)-Y(t)|^2\right] \tag{2.6}$$
$$\leq \inf_{\epsilon\in(0,1)}\left\{\left(\frac{\mathbb{E}\left[\|X_0-Y_0\|_\infty^2\right]}{1-\epsilon}+2\beta(t)\left(\frac{\epsilon+6}{(1-\epsilon)\epsilon}\right)\mathbb{E}\left[\|X_0-Y_0\|_{L^p}^p\right]\right)\right.$$
$$\left.\cdot\exp\left(4\beta(t)\left(\frac{\epsilon+3}{(1-\epsilon)\epsilon}\right)t\right)\right\}.$$

(ii)

$$\mathbb{E}\left[\sup_{r\in[-r_0,T]}|X(r)|^2+\int_0^T|X(s)|^p\,ds\right]$$
$$\leq H(T)\left(1+\mathbb{E}\left[\|X_0\|_\infty^2\right]+\mathbb{E}\left[\|X_0\|_{L^p}^p\right]\right),$$

for all $T>0$ and some non-decreasing function $H:\mathbb{R}_+\to\mathbb{R}_+$.

(c) *(2.1) has weak uniqueness.*

2.2. **Proof of the main result.** We are going to prove the main result by using the Banach fixed-point theorem. Fix a stochastic basis $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq -r_0},P)$, a d-dimensional $(\mathcal{F}_t)$-Brownian motion $(W(t))_{t\geq -r_0}$ and an initial condition $\psi\in\mathcal{C}$. For $T>0$ and $q\in\mathbb{N}$ define

$$E^q(T):=\big\{X\in L^{2q}(\Omega,\mathcal{F},P;C([-r_0,T];\mathbb{R}^d)):(X(t))_{t\in[-r_0,T]}\text{ is a }$$
$$(\mathcal{F}_t)_{t\in[-r_0,T]}-\text{adapted, continuous process}\big\}.$$

Clearly, $E^q(T)$ is a Banach space with respect to the norm

$$\|X\|_{E^q(T)}^{2q}:=\mathbb{E}\left[\sup_{t\in[-r_0,T]}|X(t)|^{2q}\right].$$

Moreover define

$$E^q:=\left\{X:\Omega\times[-r_0,\infty)\to\mathbb{R}^d:\ X\,|_{\Omega\times[-r_0,T]}\in E^q(T)\ \forall T>0\right\}.$$

Next, solve for any $X\in E^q$ the classical path-dependent SDE

$$\begin{cases}\mathrm{d}Y(t)=b(t,Y_t,\mu_t)\mathrm{d}t+\sigma(t,Y_t,\mu_t)\mathrm{d}W(t),\ t\geq 0,\\ Y_0=\psi,\end{cases}\tag{2.7}$$

where $\mu_t:=\mathcal{L}_{X_t}$.

Before we can prove Theorem 2.1.6 we need the following two lemma. The first lemma deals with the existence of solutions to (2.7).

Lemma 2.2.1. *Assume (H1)-(H4). Then for any $X\in E^q$, $q\geq\frac{p}{2}$ and any initial condition $\psi\in\mathcal{C}$, (2.7) has a unique solution $Y\in E^q$, i.e. there exists a unique* *continuous, adapted,* $\mathbb{R}^d$*-valued processes* $(Y(t))_{t\geq -r_0}$ *which fulfills (2.7). Moreover, for all* $T>0$,

$$\mathbb{E}\left[\sup_{t\in[-r_0,T]}|Y(t)|^{2q}\right]<\infty.$$

Proof. Define $\bar{b}(t,\xi):=b(t,\xi,\mu_t)$ and $\bar{\sigma}(t,\xi):=\sigma(t,\xi,\mu_t)$, $(t,\xi)\in[0,\infty)\times\mathcal{C}$. Now it is easy to see that $\bar{b}$ and $\bar{\sigma}$ fulfill (H1)-(H5) in [21, Theorem 4.2]. □

Now take $T>0$ and $q\in\mathbb{N}$ with $q\geq\frac{p}{2}$ fixed but arbitrary. For $X\in E^q(T)$ define $\Lambda X\in E^q(T)$ as the unique solution to (2.7) up to time T and $\Lambda X\in E^q(T)$. $\Lambda: E^q(T)\to E^q(T)$ is a well-defined mapping, since we can extend every $X\in E^q(T)$ to an element $\tilde{X}\in E^q$ by setting $\tilde{X}(t):=X(T)$ for $t>T$ and apply Lemma 2.2.1 to $\tilde{X}$ in order to get a solution up to infinity and therefore up to time $T>0$. The path-wise uniqueness up to time T can be proved as in [21] or as in the proof of Theorem 2.1.6 (ii) below.
If $X\in E^q(T)$ is a fixed-point of Λ, i.e. $\Lambda X=X$, we have for $t\in[0,T]$ that

$$\begin{aligned}X(t)=\Lambda X(t)&=\Lambda X(0)+\int_0^t b(s,(\Lambda X)_s,\mathcal{L}_{X_s})\mathrm{d}s+\int_0^t\sigma(s,(\Lambda X)_s,\mathcal{L}_{X_s})\mathrm{d}W(s)\\&=X(0)+\int_0^t b(s,X_s,\mathcal{L}_{X_s})\mathrm{d}s+\int_0^t\sigma(s,X_s,\mathcal{L}_{X_s})\mathrm{d}W(s)\quad P-\text{a.s.}\end{aligned}$$

and

$$X(t)=\Lambda X(t)=\psi(t)$$

for all $t\in[-r_0,0]$ P-a.s. Thus X is a solution of (2.7) up to time T. Therefore our next step is to show that Λ fulfills the conditions of the generalized Banach fixed-point theorem.

Lemma 2.2.2. *There exists* $K_q:\mathbb{R}_+\to\mathbb{R}_+$ *non-decreasing such that for all* $X,Y\in E^q(T)$ *and* $n\in\mathbb{N}$

$$\mathbb{E}\left[\sup_{t\in[-r_0,T]}|\Lambda^nX(t)-\Lambda^nY(t)|^{2q}\right]\leq K_q(T)^n\frac{T^n}{n!}\mathbb{E}\left[\sup_{t\in[-r_0,T]}|X(t)-Y(t)|^{2q}\right].\quad(2.8)$$

(Whereby Λ^nX *means, that* Λ *is applied* n*-times to* X*.)*

Proof. For $n\in\mathbb{N}_0$ define $X^{(n)}:=\Lambda^nX$ and $Y^{(n)}:=\Lambda^nY$. Moreover define $\mu_t^{(n)}:=\mathcal{L}_{X_t^{(n)}}$ and $\nu_t^{(n)}:=\mathcal{L}_{Y_t^{(n)}}$, $t\in[0,T]$. By the definition of Λ we have for $n\geq1$, that $X^{(n)}=\Lambda(X^{(n-1)})$ solves

$$\begin{cases}\mathrm{d}X^{(n)}(t)=b(t,X_t^{(n)},\mu_t^{(n-1)})\mathrm{d}t+\sigma(t,X_t^{(n)},\mu_t^{(n-1)})\mathrm{d}W(t),\ t\in[0,T],\\ X_0^{(n)}=\psi\end{cases}$$

and that $Y^{(n)}=\Lambda(Y^{(n-1)})$ solves

$$\begin{cases}\mathrm{d}Y^{(n)}(t)=b(t,Y_t^{(n)},\nu_t^{(n-1)})\mathrm{d}t+\sigma(t,Y_t^{(n)},\nu_t^{(n-1)})\mathrm{d}W(t),\ t\in[0,T],\\ X_0^{(n)}=\psi.\end{cases}$$

Applying Itô's formula to $|X^{(n)}(t)-Y^{(n)}(t)|^2$ and using (H3) one can prove as in [12, Lemma 3.2] that we have for $t\in[0,T]$ and $n\in\mathbb{N}$:

$$\begin{aligned}&\mathbb{E}\bigg[\sup_{r\in[-r_0,t]}|X^{(n)}(r)-Y^{(n)}(r)|^{2q}\bigg]\\ &\quad\le 2(C_q(T)+\tilde{C}_q(T))\mathbb{E}\bigg[\int_0^t\sup_{r\in[-r_0,s]}|X^{(n)}(r)-Y^{(n)}(r)|^{2q}\mathrm{d}s\bigg]\\ &\qquad+2(C_q(T)+\tilde{C}_q(T))\mathbb{E}\bigg[\int_0^t\sup_{r\in[-r_0,s]}|X^{(n-1)}(r)-Y^{(n-1)}(r)|^{2q}\mathrm{d}s\bigg],\end{aligned}$$

with C_q, $\tilde{C}_q:\mathbb{R}_+\to\mathbb{R}_+$ non-decreasing. Now (2.8) follows from Gronwall's Lemma. □

2.2.1. *Proof of Theorem 2.1.6.* Now we can prove Theorem 2.1.6.

Proof. (a): By Lemma 2.2.2 we have for all $q\in\mathbb{N}$ with $q\ge\frac{p}{2}$ and $T>0$

$$\|\Lambda^nX-\Lambda^nY\|_{E^q(T)}\le\left(\frac{(K_q(T)T)^n}{n!}\right)^{\frac{1}{2q}}\|X-Y\|_{E^q(T)},$$

for all $n\in\mathbb{N}$ and $X,Y\in E^q(T)$. Thus, by the generalized Banach fixed-point theorem, Λ has a unique fixed-point $X\in E^q(T)$. As discussed above, this means that X is a a solution of (2.1) up to time T. Since $T>0$ was taken arbitrarily and the (pathwise) uniqueness is ensured by (b), this implies that (2.1) has a unique solution up to every time $T>0$. Hence (2.1) has solution up to infinity.
Since $q\in\mathbb{N}$ with $q\ge\frac{p}{2}$ was taken arbitrarily, X, fulfills (2.5) for all $T>0$ and $q\in\mathbb{N}$.

(b): (i): Let (X,W) and (Y,W) be two weak solutions of (2.1) defined on a stochastic basis $(\Omega,\mathcal{F},(\mathcal{F})_{t\ge-r_0},P)$. By Itô's formula and (H3) we have for all $t\ge0$

$$\begin{aligned}&|X(t)-Y(t)|^2\\ =&|X(0)-Y(0)|^2+\int_0^t2\left\langle b(s,X_s,\mu_s)-b(s,Y_s,\nu_s),X(s)-Y(s)\right\rangle\mathrm{d}s\\ &+\int_0^t\|\sigma(s,X_s,\mu_s)-\sigma(s,Y_s,\nu_s)\|_{\mathrm{HS}}^2\mathrm{d}s\\ &+2\int_0^t\left\langle X(s)-Y(s),\{\sigma(s,X_s,\mu_s)-\sigma(s,Y_s,\nu_s)\}\,\mathrm{d}W(s)\right\rangle\\ \le&\|X_0-Y_0\|_\infty^2+2\beta(t)\|X_0-Y_0\|_{L^p}^p+2\beta(t)\int_0^t\|Y_s-Y_s\|_\infty^2+\mathbb{W}_2(\mu_s,\nu_s)^2\mathrm{d}s\\ &+2\sup_{r\in[0,t]}\left|\int_0^r\left\langle X(s)-Y(s),\{\sigma(s,X_s,\mu_s)-\sigma(s,Y_s,\nu_s)\}\,\mathrm{d}W(s)\right\rangle\right|,\end{aligned}$$

where $\mu_t := \mathcal{L}_{X_t}$ and $\nu_t := \mathcal{L}_{Y_t}$, $t \geq 0$.
Obviously this estimate is also true if $t \in [0, -r_0]$. Thus

$$\begin{aligned}
&\sup_{r\in[-r_0,t]} |X(r)-Y(r)|^2\\
\leq& \|X_0-Y_0\|_\infty^2 + 2\beta(t)\|X_0-Y_0\|_{L^p}^p + 2\beta(t)\int_0^{t\vee 0} \|Y_s-Y_s\|_\infty^2 + \mathbb{W}_2(\mu_s,\nu_s)^2 \mathrm{d}s\\
&+2\sup_{r\in[0,t\vee 0]}\left|\int_0^r \langle X(s)-Y(s), \{\sigma(s,X_s,\mu_s)-\sigma(s,Y_s,\nu_s)\}\,\mathrm{d}W(s)\rangle\right|.
\end{aligned}$$

By the BDG, Young's inequality and (H3) and we have for all $\epsilon \in (0,1)$ and $t \geq 0$

$$\begin{aligned}
&2\mathbb{E}\left[\sup_{r\in[0,t]}\left|\int_0^r \langle X(s)-Y(s), \{\sigma(s,X_s,\mu_s)-\sigma(s,Y_s,\nu_s)\}\,\mathrm{d}W(s)\rangle\right|\right]\\
\leq& 6\mathbb{E}\left[\left(\int_0^t |X(s)-Y(s)|^2\|\sigma(s,X_s,\mu_s)-\sigma(s,Y_s,\nu_s)\|_{\mathrm{HS}}^2\mathrm{d}s\right)^{\frac{1}{2}}\right]\\
\leq& 6\mathbb{E}\left[\sup_{s\in[0,t]}|X(s)-Y(s)|\left(\int_0^t \|\sigma(s,X_s,\mu_s)-\sigma(s,Y_s,\nu_s)\|_{\mathrm{HS}}^2\right)^{\frac{1}{2}}\right]\\
\leq& \epsilon\mathbb{E}\left[\sup_{s\in[0,t]}|X(s)-Y(s)|^2\right] + \frac{6}{\epsilon}\mathbb{E}\left[\left(\int_0^t \|\sigma(s,X_s,\mu_s)-\sigma(s,Y_s,\nu_s)\|_{\mathrm{HS}}^2\mathrm{d}s\right)\right]\\
\leq& \epsilon\mathbb{E}\left[\sup_{s\in[-r_0,t]}|X(s)-Y(s)|^2\right]\\
&\quad + \frac{6\beta(t)}{\epsilon}\mathbb{E}\left[\int_0^t \|X_s-Y_s\|_\infty^2 + \mathbb{W}_2(\mu_s,\nu_s)^2\mathrm{d}s + \|X_0-Y_0\|_{L^p}^p\right].
\end{aligned}$$

Thus, using $\mathbb{W}_2(\mu_r,\nu_r)^2 \leq \mathbb{E}\left[\|X_s-Y_s\|_\infty^2\right] \leq \mathbb{E}\left[\sup_{r\in[-r_0,s]}|X(r)-Y(r)|^2\right]$,

$$\begin{aligned}
&\mathbb{E}\left[\sup_{s\in[-r_0,t]}|X(s)-Y(s)|^2\right]\\
\leq& \mathbb{E}\left[\|X_0-Y_0\|_\infty^2\right] + 2\beta(t)\left(\frac{\epsilon+3}{\epsilon}\right)\mathbb{E}\left[\|X_0-Y_0\|_{L^p}^p\right] + \epsilon\mathbb{E}\left[\sup_{s\in[-r_0,t]}|X(s)-Y(s)|^2\right]\\
&+4\beta(t)\left(\frac{\epsilon+3}{\epsilon}\right)\int_0^t \mathbb{E}\left[\sup_{r\in[-r_0,s]}|X(r)-Y(r)|^2\right]\mathrm{d}s.
\end{aligned}$$

Thus

$$\begin{aligned}
\mathbb{E}\left[\sup_{s\in[-r_0,t]}|X(s)-Y(s)|^2\right] \leq& \frac{\mathbb{E}\left[\|X_0-Y_0\|_\infty^2\right]}{1-\epsilon} + 2\beta(t)\left(\frac{\epsilon+3}{(1-\epsilon)\epsilon}\right)\mathbb{E}\left[\|X_0-Y_0\|_{L^p}^p\right]\\
&+4\beta(t)\left(\frac{\epsilon+3}{(1-\epsilon)\epsilon}\right)\int_0^t \mathbb{E}\left[\sup_{r\in[-r_0,s]}|X(r)-Y(r)|^2\right]\mathrm{d}s.
\end{aligned}$$

Hence, by Gronwall,

$$\mathbb{E}\bigg[\sup_{s\in[-r_0,t]}|X(s)-Y(s)|^2\bigg] \leq \bigg(\frac{\mathbb{E}\left[\|X_0-Y_0\|_\infty^2\right]}{1-\epsilon}+2\beta(t)\left(\frac{\epsilon+3}{(1-\epsilon)\epsilon}\right) \cdot \mathbb{E}\left[\|X_0-Y_0\|_{L^p}^p\right]\bigg)\exp\left(4\beta(t)\left(\frac{\epsilon+3}{(1-\epsilon)\epsilon}\right)t\right).$$

Since $\epsilon\in(0,1)$ was chosen arbitrarily, (2.6) follows.

(ii): By Itô's formula, (H2) and (H4) we have for all $t\geq 0$

$$\begin{aligned}
&|X(t)|^2\\
=&|X(0)|^2+\int_0^t 2\left\langle b(s,X_s,\mu_s),X(s)\right\rangle+\|\sigma(s,X_s,\mu_s)\|_{\mathrm{HS}}^2\mathrm{d}s\\
&\quad+2\int_0^t\left\langle X(s),\sigma(s,X_s,\mu_s)\mathrm{d}W(s)\right\rangle\\
\leq&\|X_0\|_\infty^2-\frac{1}{2}\int_0^t|X(s)|^p\mathrm{d}s+(\alpha(t)+\gamma(t))\int_0^t\left(1+\|X_s\|_\infty^2+\mu_s\left(\|\cdot\|_\infty^2\right)\right)\mathrm{d}s\\
&\quad+\alpha(t)\|X_0\|_{L^p}^p+2\sup_{r\in[0,t]}\left|\int_0^r\left\langle X(s),\sigma(s,X_s,\mu_s)\mathrm{d}W(s)\right\rangle\right|,
\end{aligned}$$

where $\mu_t:=\mathcal{L}_{X_t}$, $t\geq 0$. Obviously this estimate also holds true for $t\in[-r_0,0]$. By the BDG, Young's inequality and (H4) one can prove in the same way as in (i) that

$$\begin{aligned}
&2\mathbb{E}\bigg[\sup_{r\in[0,t]}\left|\int_0^r\left\langle X(s),\sigma(s,X_s,\mu_s)\mathrm{d}W(s)\right\rangle\right|\bigg]\\
&\qquad\leq\frac{1}{2}\mathbb{E}\bigg[\sup_{r\in[-r_0,t]}|X(r)|^2\bigg]+\gamma(t)\int_0^t\left(1+\|X_s\|_\infty^2+\mu_s\left(\|\cdot\|_\infty^2\right)\right)\mathrm{d}s.
\end{aligned}$$

Now fix $T>0$. Taking expectation, the two estimates above imply for all $t\in[0,T]$

$$\begin{aligned}
\mathbb{E}\bigg[\sup_{r\in[-r_0,t]}|X(r)|^2\bigg]\leq&\ \mathbb{E}\left[\|X_0\|_\infty^2\right]+\alpha(T)\mathbb{E}\left[\|X_0\|_{L^p}^p\right]\\
&+(\alpha(T)+2\gamma(T))\int_0^t\left(1+\mathbb{E}\left[\|X_s\|_\infty^2\right]+\mu_s\left(\|\cdot\|_\infty^2\right)\right)\mathrm{d}s\\
&-\frac{1}{2}\int_0^t\mathbb{E}\left[|X(s)|^p\right]\mathrm{d}s+\frac{1}{2}\mathbb{E}\bigg[\sup_{r\in[-r_0,t]}|X(r)|^2\bigg].
\end{aligned}$$

Thus, using that $\mu_s\left(\|\cdot\|_\infty^2\right)=\mathbb{E}\left[\|X_s\|_\infty^2\right]$ for all $s\geq 0$ by the general transformation rule,

$$\begin{aligned}
&\mathbb{E}\bigg[\sup_{r\in[-r_0,t]}|X(r)|^2\bigg]+\int_0^t\mathbb{E}\left[|X(s)|^p\right]\mathrm{d}s\\
\leq&2\mathbb{E}\left[\|X_0\|_\infty^2\right]+2\alpha(T)\mathbb{E}\left[\|X_0\|_{L^p}^p\right]+2(\alpha(T)+2\gamma(T))\int_0^t\left(1+2\mathbb{E}\left[\|X_s\|_\infty^2\right]\right)\mathrm{d}s\\
\leq&2\mathbb{E}\left[\|X_0\|_\infty^2\right]+2\alpha(T)\mathbb{E}\left[\|X_0\|_{L^p}^p\right]\\
&\qquad+2(\alpha(T)+2\gamma(T))\int_0^t\left(1+2\mathbb{E}\bigg[\sup_{r\in[-r_0,s]}\|X(r)\|_\infty^2\bigg]+\int_0^s\mathbb{E}\left[|X(r)|^p\right]\mathrm{d}r\right)\mathrm{d}s.
\end{aligned}$$

Thus, by the Gronwall lemma, there exists $H:\mathbb{R}_+\to\mathbb{R}_+$ non-decreasing such that

$$\mathbb{E}\bigg[\sup_{r\in[-r_0,T]}|X(r)|^2+\int_0^T|X(s)|^p\mathrm{d}s\bigg]\leq H(T)\left(1+\mathbb{E}\left[\|X_0\|_\infty^2\right]+\mathbb{E}\left[\|X_0\|_{L^p}^p\right]\right).$$

Since T was taken arbitrarily, this estimate holds for all $T>0$.

(c): Same as in [12, Theorem 3.1 (3)]. □

Remark 2.2.3. Let $\psi\in\mathcal{C}$ be arbitrary. Define $X^{(0)}(t):=\psi(t\wedge 0)$, $t\geq -r_0$ and $X^{(n)}:=\Lambda X^{(n-1)}$, $n\in\mathbb{N}$, whereby Λ is defined as before. By the definition of Λ, $X^{(n)}$ solves

$$\begin{cases}\mathrm{d}X^{(n)}(t)=b(t,X_t^{(n)},\mathcal{L}_{X_t^{(n-1)}})\mathrm{d}t+\sigma(t,X_t^{(n)},\mathcal{L}_{X_t^{(n-1)}})\mathrm{d}W(t)\\ X_0^{(n)}=\psi.\end{cases}$$

As we know from the Banach fixed point theorem and Lemma 2.2.2, $X^{(n)}=\Lambda^nX^{(0)}\to X$ in $E^q(T)$ as $n\to\infty$, for all $T>0$. In [12] the authors prove the convergence of the $X^{(n)}$ directly and show that the limit is a solution to (2.1). Therefore the *iteration in distribution*, which is used in [12], is contained in our proof.

3. Distribution-Dependent SDEs with delay in infinite dimensions

The goal of this chapter is to obtain a result for the existence and uniqueness of solutions of distribution-dependent SDE's with delay in infinite dimensions. We achieve this by following the idea of [16, Chapter 4], i.e. we approximate with solutions of finite dimensional distribution-dependent SDE's with delay (Galerkin approximation).

Throughout this chapter we fix a Gelfand triple (V,H,V^*), a stochastic basis $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq -r_0},P)$, $r_0>0$, $T>0$, $p\geq 2$ and $p^*:=\frac{p}{p-1}$.
Since

$$V\subset H\subset V^*$$

continuous and densely, we have

$$L_V^p\subset L_H^2\subset L_{V^*}^{p^*}$$

continuous and densely. By Kuratowski's theorem we have $L_V^p \in \mathcal{B}(L_H^2)$, $L_H^2 \in \mathcal{B}(L_{V^*}^{p^*})$ and $\mathcal{B}(L_V^p) = \mathcal{B}(L_H^2) \cap L_V^p$, $\mathcal{B}(L_H^2) = \mathcal{B}(L_{V^*}^{p^*}) \cap L_H^2$. Hence

$$\mathcal{P}_p\left(L_V^p\right) \subset \mathcal{P}_2\left(L_H^2\right) \subset \mathcal{P}_{p^*}\left(L_{V^*}^{p^*}\right),$$

continuously. Therefore we can define $\mathcal{P}_2(\mathcal{C}(H)) \cap \mathcal{P}_p(L_V^p)$ and $C\left([0,T];\mathcal{P}_2(\mathcal{C}(H))\right) \cap L^p\left([0,T];\mathcal{P}_p\left(L_V^p\right)\right)$ as in section 1.1. The aim of this chapter is to solve the following path-distribution dependent SDE on H:

$$\mathrm{d}X(t) = A(t, X_t, \mathcal{L}_{X_t})\mathrm{d}t + B(t, X_t, \mathcal{L}_{X_t})\mathrm{d}W(t), \tag{3.1}$$

with $W = (W(t))_{t\in[0,T]}$, a cylindrical Q-Wiener process with $Q = I$, defined on another separable Hilbert space $(U, \langle\cdot,\cdot\rangle_U)$ and with B taking values in $L_2(U,H)$, but with A taking values in the larger space V^*.
By our definition of solution (see below), X will, however, take values in H again.

3.1. **Conditions on the coefficients and main result.** In this section, the conditions on the coefficients and the main result are presented.
For the rest of this chapter let $\mathbb{W}_2 := \mathbb{W}_2^{L_H^2}$.
Throughout the rest of this chapter, we assume that A and B fulfill the following conditions:

(H1) (Continuity)

$$\begin{aligned} &A\colon [0,T] \times (\mathcal{C}(H) \cap L_V^p) \times (\mathcal{P}_2(\mathcal{C}(H)) \cap \mathcal{P}_p(L_V^p)) \mapsto V^*;\\ &B\colon [0,T] \times (\mathcal{C}(H) \cap L_V^p) \times (\mathcal{P}_2(\mathcal{C}(H)) \cap \mathcal{P}_p(L_V^p)) \mapsto L_2(U,H)\end{aligned}$$

are $\mathcal{B}([0,T]) \otimes \mathcal{B}(\mathcal{C}(H) \cap L_V^p) \otimes \mathcal{B}(\mathcal{P}_2(\mathcal{C}(H)) \cap \mathcal{P}_p(L_V^p))$-measurable. In addition for all $t \in [0,T]$ and $v \in V$ and $u \in U$ the maps

$$(\mathcal{C}(H) \cap L_V^p) \times (\mathcal{P}_2(\mathcal{C}(H)) \cap \mathcal{P}_p(L_V^p)) \ni (\xi,\mu) \mapsto {}_{V^*}\langle A(t,\xi,\mu), v\rangle_V$$

and

$$(\mathcal{C}(H) \cap L_V^p) \times (\mathcal{P}_2(\mathcal{C}(H)) \cap \mathcal{P}_p(L_V^p)) \ni (\xi,\mu) \mapsto B(t,\xi,\mu)u$$

are continuous.

(H2) (Coercivity) There exists $\alpha \geq 0$ such that

$$\begin{aligned}&\int_0^t e^{-\lambda s}\left(2_{V^*}\langle A(s,\xi_s,\mu_s), \xi(s)\rangle_V + \|B(s,\xi_s,\mu_s)\|_{L_2(U,H)}^2\right)\mathrm{d}s\\ &\quad\leq \alpha\int_0^t e^{-\lambda s}\left(1 + \|\xi_s\|_{L_H^2}^2 + \mu_s\left(\|\cdot\|_{L_H^2}^2\right)\right)\mathrm{d}s - \frac{1}{2}\int_0^t e^{-\lambda s}\|\xi(s)\|_V^p\mathrm{d}s,\end{aligned}$$

for all $\lambda \geq 0$, $t \in [0,T]$, $\xi \in C([-r_0,T];H) \cap L^p([-r_0,T];V)$ and $\mu \in C\left([0,T];\mathcal{P}_2(\mathcal{C}(H))\right) \cap L^p\left([0,T];\mathcal{P}_p\left(L^p\right)\right)$.

(H3) (Monotonicity) There exists $\beta \geq 0$ such that

$$\int_0^t e^{-\lambda s} 2_{V^*}\langle A(s,\xi_s,\mu_s) - A(s,\eta_s,\nu_s), \xi(s) - \eta(s)\rangle_V$$
$$\leq \beta \int_0^t e^{-\lambda s} \left(\|\xi_s - \eta_s\|^2_{L^2_H} + \mathbb{W}_2(\mu_s,\nu_s)^2 \right) \mathrm{d}s$$

and

$$\int_0^t e^{-\lambda s} \|B(s,\xi_s,\mu_s) - B(s,\eta_s,\nu_s)\|^2_{L_2(U,H)} \mathrm{d}s$$
$$\leq \beta \int_0^t e^{-\lambda s} \left(\|\xi_s - \eta_s\|^2_{L^2_H} + \mathbb{W}_2(\mu_s,\nu_s)^2 \right) \mathrm{d}s,$$

for all $\lambda \geq 0$, $t \in [0,T]$, $\xi, \eta \in C([-r_0,T];H) \cap L^p([-r_0,T];V)$ and $\mu, \nu \in C([0,T];\mathcal{P}_2(\mathcal{C}(H))) \cap L^p([0,T];\mathcal{P}_p(L^p))$.

(H4) (Growth) For all $v \in V$ ${}_{V^*}\langle A(\cdot,\cdot,\cdot), v\rangle_V$ is bounded on bounded sets in $[0,T] \times (\mathcal{C}(H) \cap L^p_V) \times (\mathcal{P}_2(\mathcal{C}(H)) \cap \mathcal{P}_p(L^p_V))$. Moreover there exist $\gamma \geq 0$ such that

$$\int_0^t \|A(s,\xi_s,\mu_s)\|^{\frac{p}{p-1}}_{V^*} \mathrm{d}s \leq \gamma \int_0^t \left(1 + \|\xi(s)\|^p_V + \mu_s\big(\|\cdot\|^p_{L^p_V}\big)\right) \mathrm{d}s + \gamma \|\xi_0\|^p_{L^p_V}$$

and

$$\|B(t,\xi_t,\mu_t)\|^2_{L_2(U,H)} \leq \gamma \Big(1 + \|\xi_t\|^2_{L^2_H} + \mu_t\big(\|\cdot\|^2_{L^2_H}\big)\Big),$$

for all $\lambda \geq 0$, $t \in [0,T]$, $\xi \in C([-r_0,T];H) \cap L^p([-r_0,T];V)$ and $\mu \in C([0,T];\mathcal{P}_2(\mathcal{C}(H))) \cap L^p([0,T];\mathcal{P}_p(L^p))$.

Let us briefly comment on these conditions. As we will see in Lemma 3.3.2, (H1)-(H4) were chosen in such the way, that in the case $V = H = V^* = \mathbb{R}^d$ for some $d \in \mathbb{N}$, they imply (H1)-(H4) in Chapter 2. The factor $e^{-\lambda s}$ is necessary, because in order to prove our main result, Itô's product rule will be applied to a term of the form $e^{-\lambda t}\|X(t)\|^2_H$ so that the factor $e^{-\lambda s}$ will appear under an integral on the right hand-side, which we want to estimate with our conditions.
Note that the measurability of $s \mapsto A(s,\xi_s,\mu_s)$ and $s \mapsto B(s,\xi_s,\mu_s)$ is ensured by (H1) and the assumptions on ξ and μ. The existence of all integrals, which appear in the conditions, is ensured by (H4).

Next we define precisely what a solution of (3.1) is.

Definition 3.1.1. A continuous H-valued process, $(\mathcal{F}_t)_{t\in[-r_0,T]}$-adapted $(X(t))_{t\in[-r_0,T]}$ is called a *solution* of (3.1), if it has the following properties:

(i) $\mathbb{E}\left[\|X_t\|^2_{\mathcal{C}(H)}\right] < \infty$ for all $t \in [0,T]$;

(ii) For its $\mathrm{d}t \otimes P$-equivalence class $\widehat{X}$ we have $\widehat{X} \in L^p([-r_0,T] \times \Omega, \mathrm{d}t \otimes P; V)$ (Whereby the $\mathrm{d}t \otimes P$-equivalence class $\widehat{X}$ of X consists of all $\tilde{X} : [-r_0,T] \times \Omega \to V^*$, $\mathcal{B}([0,T]) \otimes \mathcal{F}/\mathcal{B}(V^*)$-measurable such that $X = \tilde{X}$ $\mathrm{d}t \otimes P$-a.e.)

(iii)

$$X(t) = X(0) + \int_0^t A(s, \bar{X}_s, \mathcal{L}_{\bar{X}_s})\mathrm{d}s + \int_0^t B(s, \bar{X}_s, \mathcal{L}_{\bar{X}_s})\mathrm{d}W(s), \tag{3.2}$$

for every $t \in [0,T]$ P-a.s., where $(\bar{X}_t)_{t\in[0,T]}$ is a progressively measurable, $\mathcal{C}(H)\cap L^p_V$-valued version (Recall that $(\tilde{X}_t)_{t\in[0,T]}$ is $\mathrm{d}t\otimes P$-version of $(X_t)_{t\in[0,T]}$, if $\tilde{X}_t(\omega) = X_t(\omega)$ for $\mathrm{d}t \otimes P$-a.e. $(t,\omega) \in [0,T] \times \Omega$.) of the $\mathcal{C}(H)$-valued process $(X_t)_{t\in[0,T]}$, with the property that $(\bar{X}(t))_{t\in[0,T]} := (\bar{X}_t(0))_{t\in[0,T]}$ is a V-valued, progressively measurable, $\mathrm{d}t \otimes P$-version of $(X(t))_{t\in[0,T]}$.

Remark 3.1.2. (a) Just like in Remark 2.1.2, Definition 3.1.1(i) implies that $\mathbb{E}\left[\sup_{t\in[-r_0,T]} \|X(t)\|_H^2\right] = \mathbb{E}\left[\sup_{t\in[0,T]} \|X_t\|^2_{\mathcal{C}(H)}\right] < \infty$ and thereby that $[0,T] \ni t \mapsto \mathcal{L}_{X_t} \in \mathcal{P}_2(\mathcal{C}(H))$ is continuous. Hence,

$$[0,T] \ni t \mapsto \mathcal{L}_{\bar{X}_t} \in \mathcal{P}_2(\mathcal{C}(H)) \cap \mathcal{P}_p(L^p_V)$$

is $\mathcal{B}([0,T])/\mathcal{B}(\mathcal{P}_2(\mathcal{C}(H)) \cap \mathcal{P}_p(L^p_V))$-measurable and $\mathcal{L}_{\bar{X}_t} = \mathcal{L}_{X_t}$ for all $t \in [0,T]$. In particular $A(\cdot, \bar{X}_\cdot, \mathcal{L}_{\bar{X}_\cdot})$ and $B(\cdot, \bar{X}_\cdot, \mathcal{L}_{\bar{X}_\cdot})$ are progressively measurable.

(b) By (H4) we have

$$\mathbb{E}\left[\int_0^T \|A(s, \bar{X}_s, \mathcal{L}_{\bar{X}_s})\|_{V^*}^{\frac{p}{p-1}} + \|B(s, \bar{X}_s, \mathcal{L}_{\bar{X}_s})\|^2_{\mathrm{HS}}\mathrm{d}s\right] < \infty.$$

This together with (a) and (b) implies that the right-hand site of (3.2) is well-defined.

The next theorem is the main result of this chapter.

Theorem 3.1.3. *Let A, B as above satisfying (H1)-(H4) and let $\psi \in \mathcal{C}(H) \cap L^P_V$. Then there exists a unique solution X to (3.1) in the sense of the definition above which satisfies the initial condition $X_0 = \psi$. Moreover*

$$\mathbb{E}\Big[\sup_{t\in[-r_0,T]} \|X(t)\|_H^2\Big] < \infty.$$

3.2. **Example: A porous medium type equation.** The following example is similar to [21, Example 4.1.11.], but generalized to the delay distribution dependent case. We consider the following equation:

$$\mathrm{d}X(t) = \Delta\psi(t, X_t, \mathcal{L}_{X_t})\mathrm{d}t + \mathrm{d}W(t),$$

whereby $\psi : \mathbb{R}_+ \times (\mathcal{C}(H) \cap L^p_V) \times (\mathcal{P}_2(\mathcal{C}(H)) \cap \mathcal{P}_p(L^p_V)) \to L^{\frac{p}{p-1}}(\Lambda)$, $\Lambda \subset \mathbb{R}^d$, open and bounded, $p \in [2,\infty[$. We set $V := L^p(\Lambda)$ and $H := \left(H_0^{1,2}(\Lambda)\right)^*$ and recall the following:

Lemma 3.2.1. *The map*

$$\Delta : H_0^{1,2}(\Lambda) \to (L^p(\Lambda))^*$$

extends to a linear isometry

$$\Delta : L^{\frac{p}{p-1}}(\Lambda) \to (L^p(\Lambda))^*$$

and for all $u \in L^{\frac{p}{p-1}}(\Lambda)$, $v \in L^p(\Lambda)$

$$_{V^*}\langle -\Delta u, v\rangle_V = {}_{L^{\frac{p}{p-1}}}\langle u, v\rangle_{L^p} = \int u(x)v(x)dx. \tag{3.3}$$

Assume that ψ fulfills the following properties:

(Ψ1) ψ is $\mathcal{B}([0,T]) \otimes \mathcal{B}\left(\mathcal{C}(H) \cap L_V^p\right) \otimes \mathcal{B}\left(\mathcal{P}_2(\mathcal{C}(H)) \cap \mathcal{P}_p(L_V^p)\right)$-measurable and for all $t \in [0,T]$ and $v \in V = L^p(\Lambda)$ the map

$$(\mathcal{C}(H) \cap L_V^p) \times \left(\mathcal{P}_2\left(\mathcal{C}(H)\right) \cap \mathcal{P}_p\left(L_V^p\right)\right) \ni (\xi, \mu) \mapsto \int \psi(t,\xi,\mu)(x)v(x)\mathrm{d}x$$

is continuous.

(Ψ2) There exists $\alpha \geq 0$ such that

$$\begin{aligned}&\int_0^t e^{-\lambda s}\left(2\int \psi(s,\xi_s,\mu_s)(x)\xi(s,x)\mathrm{d}x\right)\mathrm{d}s\\ &\geq -\alpha\int_0^t e^{-\lambda s}\left(1 + \|\xi_s\|_{L_H^2}^2 + \mu_s\left(\|\cdot\|_{L_H^2}^2\right)\right)\mathrm{d}s + \frac{1}{2}\int_0^t e^{-\lambda s}\|\xi(s)\|_V^p\mathrm{d}s,\end{aligned}$$

for all $\lambda \geq 0$, $t \in [0,T]$, $\xi \in C([-r_0,T];H) \cap L^p([-r_0,T];V)$ and $\mu \in C\left([0,T];\mathcal{P}_2(\mathcal{C}(H))\right) \cap L^p\left([0,T];\mathcal{P}_p\left(L^p\right)\right)$.

(Ψ3) There exists $\beta \geq 0$ such that

$$\int_0^t e^{-\lambda s}2\Big(\int \left(\psi(s,\xi_s,\mu_s)(x) - \psi(s,\eta_s,\nu_s)(x)\right)\left(\xi(s,x) - \eta(s,x)\right)\mathrm{d}x\Big)\mathrm{d}s \geq 0$$

for all $\lambda \geq 0$, $t \in [0,T]$, $\xi, \eta \in C([-r_0,T];H) \cap L^p([-r_0,T];V)$ and $\mu, \nu \in C\left([0,T];\mathcal{P}_2(\mathcal{C}(H))\right) \cap L^p\left([0,T];\mathcal{P}_p\left(L^p\right)\right)$.

(Ψ4) For all $v \in V$, $\int \psi(\cdot,\cdot,\cdot)(x)v(x)\mathrm{d}x$ is bounded on bounded sets in $[0,T] \times (\mathcal{C}(H) \cap L_V^p) \times \left(\mathcal{P}_2\left(\mathcal{C}(H)\right) \cap \mathcal{P}_p\left(L_V^p\right)\right)$. Moreover there exist $\gamma \geq 0$ such that

$$\begin{aligned}&\int_0^t \|\psi(s,\xi_s,\mu_s)\|_{L^{\frac{p}{p-1}}(\Lambda)}^{\frac{p}{p-1}}\mathrm{d}s\\ &\qquad \leq \gamma\int_0^t\left(1 + \|\xi(s)\|_V^p + \mu_s\left(\|\cdot\|_{L_V^p}^p\right)\right)\mathrm{d}s + \gamma\|\xi_0\|_{L_V^p}^p.\end{aligned}$$

Now define $A\colon [0,T] \times (\mathcal{C}(H) \cap L_V^p) \times \left(\mathcal{P}_2\left(\mathcal{C}(H)\right) \cap \mathcal{P}_p\left(L_V^p\right)\right) \to V^* = (L^p(\Lambda))^*$ by

$$A(t,\xi,\mu) := \Delta\psi(t,\xi,\mu),\ (t,\xi,\mu) \in [0,T] \times (\mathcal{C}(H) \cap L_V^p) \times \left(\mathcal{P}_2\left(\mathcal{C}(H)\right) \cap \mathcal{P}_p\left(L_V^p\right)\right).$$

By Lemma 3.2.1 A is well-defined and really takes values in V^*. By (3.3) we have for $(t,\xi,\mu) \in [0,T] \times (\mathcal{C}(H) \cap L_V^p) \times \left(\mathcal{P}_2\left(\mathcal{C}(H)\right) \cap \mathcal{P}_p\left(L_V^p\right)\right)$ and $v \in V$

$$_{V^*}\langle A(t,\xi,\mu), v\rangle_V = -\int \psi(t,\xi,\mu)(x)v(x)\mathrm{d}x.$$

Now it is easy to see that (Ψ1)-(Ψ4) imply (H1)-(H4).

3.3. **Proof of the main result.** Let $\{e_i \mid i \in \mathbb{N}\} \subset V$ be an orthonormal basis of H such that $\mathrm{span}\{e_i \mid i \in \mathbb{N}\}$ is dense in V. Define $H_n := \mathrm{span}\{e_1 \cdots e_n\} \subset V$ and $\|\cdot\|_{H_n} := \|\cdot\|_H$. Since H_n is a finite dimensional vector space, $\|\cdot\|_V \leq \|\cdot\|_V$, $\|\cdot\|_V \leq \|\cdot\|_H$ and $\|\cdot\|_V \leq \|\cdot\|_{V^*}$ are equivalent on H_n. Let $P_n \colon V^* \mapsto H_n$ be defined as

$$P_n y := \sum_{i=1}^{n} {}_{V^*}\langle y, e_i\rangle_V e_i, \ y \in V^*.$$

Since ${}_{V^*}\langle y, e_i\rangle_V = \langle y, e_i\rangle_H$ for $y \in H$, the restriction of P_n to H, denoted by $P_n|_H$, is just the orthogonal projection onto H_n in H. Moreover the following Lemma holds true.

Lemma 3.3.1. *Let P_n be as above. Then:*

(i) ${}_{V^*}\langle z, P_n y\rangle_V = {}_{V^*}\langle y, P_n z\rangle_V$ *for all* $y, z \in V^*$,
(ii) ${}_{V^*}\langle P_n y, v\rangle_V = {}_{V^*}\langle y, P_n v\rangle_V$ *for all* $y \in V^*$, $v \in V$.

Let $\{g_i \mid i \in \mathbb{N}\}$ be an orthonormal basis of U and set

$$W^{(n)}(t) := \sum_{i=1}^{n} \langle W(t), g_i\rangle_U g_i.$$

Here we define for $g \in U$

$$\langle W(t), g\rangle_U := \int_0^t \langle g, \cdot\rangle_U \mathrm{d}W(s), \ t \in (0, T],$$

where the stochastic integral is well-defined, since the map $u \mapsto \langle g, u\rangle_U$, $u \in U$, is in $L_2(U, \mathbb{R})$. By the definition of a Q-Wiener process [16, Chapter 2.5], it is clear that $(W^{(n)}(t))_{t\in[0,T]}$ is a n-dimensional Brownian motion on H_n. In addition define $U_n := \mathrm{span}\{g_1, \cdots, g_n\}$ and let $\tilde{P}_n$ is the orthogonal projection onto U_n in U.
Now we consider for each $n \in \mathbb{N}$ the following stochastic equation on H_n:

$$\begin{cases} \mathrm{d}X^{(n)}(t) = P_n A(t, X_t^{(n)}, \mathcal{L}_{X_t^{(n)}})\mathrm{d}t + P_n B(t, X_t^{(n)}, \mathcal{L}_{X_t^{(n)}})\mathrm{d}W^{(n)}(t), \ t \in [0, T] \\ X_0^{(n)} = P_n \psi. \end{cases} \tag{3.4}$$

Lemma 3.3.2. *Assume (H1)-(H4). Then, for every $n \in \mathbb{N}$, there exists a continuous, adapted, H_n-valued process $X^{(n)}$ which solves (3.4).*

Proof. Fix $n \in \mathbb{N}$. It is obvious that Theorem 2.1.6 still holds true, if we replace $\mathbb{R}^d$ with and arbitrary, finite dimensional vector space. Therefore we have to show, that $b(t, \xi, \mu) := P_n A(t, \xi, \mu)$ and $\sigma(t, \xi, \mu) := P_n B(t, \xi, \mu)\tilde{P}_n$, $(t, \xi, \mu) \in [0, T] \times \mathcal{C}(H_n) \times \big(\mathcal{P}_2(\mathcal{C}(H_n)) \cap \mathcal{P}_p(L^p_{H_n})\big) \subset [0, T] \times (\mathcal{C}(H) \cap L^p_V) \times (\mathcal{P}_2(\mathcal{C}(H)) \cap \mathcal{P}_p(L^p_V))$, fulfill (H1)-(H4) in Theorem 2.1.6. $\square$

The following lemma is crucial for the construction of a solution to (3.1). But first we fix the following notations: Let

$$\begin{aligned} J &:= L^2([0,T]\times\Omega, \mathrm{d}t\otimes P; L_2(U,H)),\\ K &:= L^p([0,T]\times\Omega, \mathrm{d}t\otimes P; V),\\ K^* &:= (L^p([0,T]\times\Omega, \mathrm{d}t\otimes P; V))^* \cong L^{\frac{p}{p-1}}([-r_0,T]\otimes\Omega, \mathrm{d}t\times P; V^*). \end{aligned}$$

Lemma 3.3.3. *Under the assumptions of the main theorem, there exists $C\in]0,\infty[$ such that*

$$\begin{aligned} \|X^{(n)}\|_K + \|A(\cdot, X^{(n)}_\cdot, \mathcal{L}_{X^{(n)}_\cdot})\|_{K^*} + \|B(\cdot, X^{(n)}_\cdot, \mathcal{L}_{X^{(n)}_\cdot})\|_J \\ + \sup_{t\in[-r_0,T]} \mathbb{E}[\|X^{(n)}(t)\|_H^2] \le C \end{aligned} \tag{3.5}$$

for all $n\in\mathbb{N}$.

Proof. Since P_n is a the orthonorgal projection of H onto H_n it is well known that $\|P_n\|_{L(H)}\le 1$. Hence for $t\in[-r_0,0]$ it is $\mathbb{E}[\|X^{(n)}(t)\|_H^2] = \|P_n\psi(t)\|_H^2 \le \|\psi\|^2_{\mathcal{C}(H)}$. Therefore we only have to show

$$\begin{aligned} \|X^{(n)}\|_{K_0} + \|A(\cdot, X^{(n)}_\cdot, \mathcal{L}_{X^{(n)}_\cdot})\|_{K^*} + \|B(\cdot, X^{(n)}_\cdot, \mathcal{L}_{X^{(n)}_\cdot})\|_J \\ + \sup_{t\in[0,T]} \mathbb{E}[\|X^{(n)}(t)\|_H^2] \le C \end{aligned}$$

for some $C\in]0,\infty[$.
By (H4) it is even enough to show that

$$\|X^{(n)}\|_K + \sup_{t\in[0,T]} \mathbb{E}[\|X^{(n)}(t)\|_H^2] \le \tilde{C}$$

for some $\tilde{C}\in]0,\infty[$. By the finite-dimensional Itô formula and Lemma 3.3.1 we have P-a.s.

$$\begin{aligned} &\|X^{(n)}(t)\|_H^2\\ &= \|X^{(n)}(0)\|_H^2 + \int_0^t \Big(2_{V^*}\big\langle P_n A\big(s, X_s^{(n)}, \mathcal{L}_{X_s^{(n)}}\big), X^{(n)}(s)\big\rangle_V + \|Z^{(n)}(s)\|^2_{L_2(U_n,H)}\Big)\mathrm{d}s\\ &\quad + M^{(n)}(t)\\ &= \|X^{(n)}(0)\|_H^2 + \int_0^t \Big(2_{V^*}\big\langle A\big(s, X_s^{(n)}, \mathcal{L}_{X_s^{(n)}}\big), X^{(n)}(s)\big\rangle_V + \|Z^{(n)}(s)\|^2_{L_2(U_n,H)}\Big)\mathrm{d}s\\ &\quad + M^{(n)}(t), \end{aligned}$$

for all $t\in[0,T]$, where $Z^{(n)}(s) = P_n B(s, X_s^{(n)}, \mathcal{L}_{X_s^{(n)}})$, $U_n = \mathrm{span}\{g_1,\cdots,g_n\}$ and

$$M^{(n)}(t) := 2\int_0^t \Big\langle X^{(n)}(s), P_n B(s, X_s^{(n)}, \mathcal{L}_{X_s^{(n)}})\mathrm{d}W^{(n)}(s)\Big\rangle_H \quad t\in[0,T],$$

is a local martingale.
Let $(\tau_l)_{l\in\mathbb{N}}$ be $(\mathcal{F}_t)$-stopping times such that $\|X^{(n)}(t\wedge\tau_l)(\omega)\|_V^2$ is uniformly bounded

in $(t,\omega) \in [0,T] \times \Omega$, $M^{(n)}((t \wedge \tau_l)$, $t \in [0,T]$, is a martingale for each $l \in \mathbb{N}$ and $\tau_l \uparrow T$ as $l \to \infty$. Then for all $l \in \mathbb{N}$, $t \in [0,T]$

$$\begin{aligned}
&\mathbb{E}\left[\|X^{(n)}(t \wedge \tau_l)\|_H^2\right] \\
=&\mathbb{E}\left[\|X^{(n)}(0)\|_H^2\right] \\
&+ \int_0^t \mathbb{E}\left[1_{[0,\tau_l]}(s)\left(2_{V^*}\langle A(s, X_s^{(n)}, \mathcal{L}_{X_s^{(n)}}, X^{(n)}(s)\rangle_V + \|Z^{(n)}(s)\|^2_{L_2(U_n,H)}\right)\right] \mathrm{d}s.
\end{aligned}$$

Using the product rule, (H3) and Fubini we obtain for $\lambda \geq 0$

$$\begin{aligned}
&\mathbb{E}\left[e^{-\lambda t}\|X^{(n)}(t \wedge \tau_l)\|_H^2\right] \\
=&\mathbb{E}\left[\|X^{(n)}(0)\|_H^2\right] - \int_0^t \lambda \mathbb{E}\left[\|X^{(n)}(s \wedge \tau_l)\|_H^2\right] e^{-\lambda s}\mathrm{d}s \\
&+ \mathbb{E}\bigg[\int_0^{t\wedge\tau_l} e^{-\lambda s}\Big(2_{V^*}\langle A(s, X_s^{(n)}, \mathcal{L}_{X_s^{(n)}}), X^{(n)}(s)\rangle_V \\
&\qquad\qquad + \|B(s, X_s^{(n)}, \mathcal{L}_{X_s^{(n)}})\|^2_{L_2(U_n,H)}\Big)\mathrm{d}s\bigg] \\
\leq&\mathbb{E}\left[\|X^{(n)}(0)\|_H^2\right] - \int_0^t \lambda \mathbb{E}\left[\|X^{(n)}(s \wedge \tau_l)\|_H^2\right] e^{-\lambda s}\mathrm{d}s \\
&+ \mathbb{E}\bigg[\int_0^{t\wedge\tau_l} e^{-\lambda s}\left(\alpha\left(1 + \|X_s^{(n)}\|^2_{L_H^2} + \mu_s^{(n)}(\|\cdot\|^2_{L_H^2})\right) - \frac{1}{2}\|X^{(n)}(s)\|_V^p\right)\mathrm{d}s\bigg],
\end{aligned}$$

where $\mu_s^{(n)} := \mathcal{L}_{X_s^{(n)}}$. Rearranging the terms yields

$$\begin{aligned}
&\mathbb{E}\left[e^{-\lambda t}\|X^{(n)}(t \wedge \tau_l)\|_H^2\right] + \int_0^t \lambda e^{-\lambda s}\mathbb{E}\left[\|X^{(n)}(s \wedge \tau_l)\|_H^2\right]\mathrm{d}s \\
&\qquad + \frac{1}{2}\int_0^t e^{-\lambda s}\mathbb{E}\left[1_{[0,\tau_l]}(s)\|X^{(n)}(s \wedge \tau_l)\|_V^p\right]\mathrm{d}s \\
\leq&\|\psi\|^2_{\mathcal{C}(H)} + \int_0^t \mathbb{E}\left[2\alpha e^{-\lambda s}\|X_s^{(n)}\|^2_{L_H^2}\right]\mathrm{d}s + \alpha\int_0^t e^{-\lambda s}\mathrm{d}s \\
\leq&K_1\|\psi\|^2_{\mathcal{C}(H)} + K_2\int_0^t e^{-\lambda s}\mathbb{E}\left[\|X^{(n)}(s)\|_H^2\right]\mathrm{d}s + \alpha\int_0^t e^{-\lambda s}\mathrm{d}s,
\end{aligned}$$

where $K_1, K_2 \geq 0$ are constants independent of n and K_2 is independent of λ. To obtain the first estimate it is used that by the definition of $\mu_s^{(n)}$ and $\mu_s^{(n)}(\|\cdot\|^2_{L_H^2})$ it is $\mu_s^{(n)}(\|\cdot\|^2_{L_H^2}) = \mathbb{E}\left[\|X_s\|^2_{L_H^2}\right]$. For the second we used Lemma 3.3.4 in the case $Y = 0$, $B = H$ and $p = 2$. In addition it is used that

$$\|X_0^{(n)}(\theta)\|_H^2 = \|P_n\psi(\theta)\|_H^2 \leq \|\psi(\theta)\|_H^2 \leq \|\psi\|^2_{\mathcal{C}(H)},$$

because $P_n|_H$ is the orthogonal projection onto H_n in H and therefore $\|P_n\|_{L(H)} \leq 1$. Choosing $\lambda = K_2$ (which is possible, because K_2 is independent of λ), taking $l \to \infty$

and applying Fatous's lemma we get

$$\mathbb{E}\Big[e^{-K_2 t}\|X^{(n)}(t)\|_H^2\Big] + \frac{1}{2}\int_0^t e^{-K_2 s}\mathbb{E}\Big[\|X^{(n)}(s)\|_V^p\Big]\mathrm{d}s \leq K_1\|\psi\|_{\mathcal{C}(H)}^2 + \alpha\int_0^t e^{-K_2 s}\mathrm{d}s,$$

for all $t \in [0, T]$. Here we used that by Chapter 3 the subtracted term is finite. Now the assertion follows for the first and fourth summand in (3.5). □

In the proof above we used the following result which is also important for the proof of the main theorem below:

Lemma 3.3.4. *Let B be a Banach space, $p \geq 2$ and X, $Y \in L^p([-r_0, T] \times \Omega, dt \otimes P; B)$ and $\lambda \geq 0$ Then*

$$\mathbb{E}\Big[\int_0^t e^{-\lambda s}\|X_s - Y_s\|_{L_B^p}^p\Big] \leq r_0\mathbb{E}\Big[\int_0^t e^{-\lambda s}\|X(s) - Y(s)\|_B^p\Big] + r_0 e^{\lambda r_0}\mathbb{E}\Big[\|X_0 - Y_0\|_{L_B^p}^p\Big]$$

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

Proof. Use Fubini and the transformation. □

3.3.1. *Proof of Theorem 3.1.3.* Now we can finally prove Theorem 3.1.3.

Proof. By Lemma 3.3.3 and the reflexivity of the spaces K, K^*, J and $L^2([0, T] \times \Omega, \mathrm{d}t \otimes P; H)$ there exist $\tilde{X} \in K$, $Y \in K^*$, $Z \in J$ and a subsequence $n_k \overset{k\to\infty}{\longrightarrow} \infty$ such that:

(i) $X^{(n_k)} \overset{k\to\infty}{\longrightarrow} \tilde{X}$ weakly in K and weakly in $L^2([0, T] \times \Omega, \mathrm{d}t \otimes P; H)$.
(ii) $Y^{(n_k)} := A(\cdot, X_\cdot^{(n_k)}, \mathcal{L}_{X_\cdot^{(n_k)}}) \overset{k\to\infty}{\longrightarrow} Y$ weakly in K^*.
(iii) $Z^{(n_k)} := B(\cdot, X_\cdot^{(n_k)}, \mathcal{L}_{X_\cdot^{(n_k)}}) \overset{k\to\infty}{\longrightarrow} Z$ weakly in J.

Note that $\tilde{X}$, Y and Z are progressively measurable, because the approximants are progressively measurable.
(iii), $P_{n_k} \overset{k\to\infty}{\longrightarrow} I_H$ and $\tilde{P}_{n_k} \overset{k\to\infty}{\longrightarrow} I_U$ implies that $P_{n_k}B(\cdot, X_\cdot^{(n_k)}, \mathcal{L}_{X_\cdot^{(n_k)}})\tilde{P}_{n_k} \overset{k\to\infty}{\longrightarrow} Z$ weakly in J. Therefore, since

$$\int_0^\cdot P_{n_k}B(s, X_s^{(n_k)}, \mathcal{L}_{X_s^{(n_k)}})\mathrm{d}W^{(n_k)}(s) = \int_0^\cdot P_{n_k}B(s, X_s^{(n_k)}, \mathcal{L}_{X_s^{(n_k)}})\tilde{P}_{n_k}\mathrm{d}W(s)$$

and since a bounded linear operator between two Banach spaces is weakly continuous, we obtain:

(iv)

$$\int_0^{\cdot} P_{n_k} B(s, X_s^{(n_k)}, \mathcal{L}_{X_s^{(n_k)}}) \mathrm{d}W^{(n_k)}(s) \overset{k\to\infty}{\longrightarrow} \int_0^{\cdot} Z(s)\mathrm{d}W(s)$$

weakly in $\mathcal{M}_T^2(H)$, which denotes the space of continuous, square integrable martingales $M : [0,T]\times\Omega \to H$ and is equipped with the norm $\|M\|^2_{\mathcal{M}_T^2(H)} := \mathbb{E}\left[\sup_{t\in[0,T]} \|M(t)\|_H^2\right]$.

Now let $v \in \bigcup_{n\geq 1} H_n$ $(\subset V)$ and $\varphi \in L^\infty([0,T]\times\Omega;\mathbb{R})$. Using (i)-(iv), the definition of $X^{(n_k)}$, Fubini and Lemma 3.3.1 we obtain

$$\begin{aligned}
&\mathbb{E}\bigg[\int_0^T {}_{V^*}\langle \tilde{X}(t), \varphi(t)v\rangle_V \mathrm{d}t\bigg] \\
&= \lim_{k\to\infty} \mathbb{E}\bigg[\int_0^T {}_{V^*}\langle X^{(n_k)}(t), \varphi(t)v\rangle_V \mathrm{d}t\bigg] \\
&= \lim_{k\to\infty} \mathbb{E}\bigg[\int_0^T \bigg({}_{V^*}\langle P_{n_k}\psi(0), \varphi(t)v\rangle + {}_{V^*}\Big\langle \int_0^t P_{n_k} Y^{(n_k)}(s)\mathrm{d}s, \varphi(t)v\Big\rangle_V \\
&\qquad\qquad + \Big\langle \int_0^t P_{n_k} Z^{(n_k)}(s)\mathrm{d}W^{(n_k)}(s), \varphi(t)v\Big\rangle_H \mathrm{d}t\bigg)\bigg] \\
&= \lim_{k\to\infty} \bigg(\mathbb{E}\bigg[\langle P_{n_k}\psi(0), v\rangle_H \int_0^T \varphi(t)v\mathrm{d}t\bigg] + \mathbb{E}\bigg[\int_0^T {}_{V*}\Big\langle Y^{(n_k)}(s), \int_s^T \varphi(t)v\mathrm{d}t\Big\rangle_V \mathrm{d}s\bigg] \\
&\qquad + \mathbb{E}\bigg[\int_0^T \Big\langle \int_0^t P_{n_k} Z^{(n_k)}(s)\mathrm{d}W^{(n_k)}(s), \varphi(t)v\Big\rangle_H \mathrm{d}t\bigg]\bigg) \\
&= \mathbb{E}\bigg[\int_0^T {}_{V*}\Big\langle \psi(0) + \int_0^t Y(s)\mathrm{d}s + \int_0^t Z(s)\mathrm{d}W(s), \varphi(t)v\Big\rangle_V \mathrm{d}t\bigg].
\end{aligned}$$

Defining

$$X(t) := \begin{cases} \psi(0) + \int_0^t Y(s)\mathrm{d}s + \int_0^t Z(s)\mathrm{d}W(s), & t\in[0,T] \\ \psi(t), & t\in[-r_0,0]. \end{cases}$$

we have for all $v \in \bigcup_{n\geq 1} H_n$ $(\subset V)$ and $\varphi \in L^\infty([0,T]\times\Omega;\mathbb{R})$

$$\mathbb{E}\bigg[\int_0^T {}_{V^*}\langle \tilde{X}(t), \varphi(t)v\rangle \mathrm{d}t\bigg] = \mathbb{E}\bigg[\int_0^T {}_{V^*}\langle X(t), \varphi(t)v\rangle \mathrm{d}t\bigg].$$

Thus, using that $\bigcup_{n\geq 1} H_n$ is dense in V by the choice of $(e_i)_{i\in\mathbb{N}}$, we have $X(t,\omega) = \tilde{X}(t,\omega)$ $\mathrm{d}t \otimes P$-a.e. $(t,\omega)\in[0,T]\times\Omega$.

This together with $X_0 = \psi \in \mathcal{C}(H) \cap L^p_V$ implies, that for the $\mathrm{d}t \otimes P$ equivalence class $\hat{X}$ of X, we have $\hat{X} \in L^p([-r_0, T] \times \Omega; \mathrm{d}t \otimes P; V)$. [16, Theorem 4.2.5] now implies that X is a continuous H-valued $(\mathcal{F}_t)$-adapted process,

$$\mathbb{E}\left[\sup_{t\in[-r_0,T]} \|X(t)\|_H^2\right] < \infty.$$

Therefore, it remains to verify that

$$A(\cdot, \bar{X}_\cdot, \mathcal{L}_{\bar{X}_\cdot}) = Y, B(\cdot, \bar{X}_\cdot, \mathcal{L}_{\bar{X}_\cdot}) = Z, \mathrm{d}t \otimes P - a.e., \tag{3.6}$$

where $(\bar{X}_t)_{t\in[0,T]}$ is a progressively measurable, $\mathcal{C}(H) \cap L^p_V$-valued version of X, the existence of which can be proved as in [16, Excercise 4.2.3., Part 2]. To prove (3.6) we first take $\rho \in L^\infty([0, T], \mathrm{d}t; \mathbb{R})$, non-negative. Then (i) and Cauchy-Schwartz implies that

$$\begin{aligned}\mathbb{E}\left[\int_0^T \rho(t)\|\tilde{X}(t)\|_H^2 \mathrm{d}t\right] &= \lim_{k\to\infty} \mathbb{E}\left[\int_0^T \langle \rho(t)\tilde{X}(t), X^{(n_k)}(t)\rangle_H \mathrm{d}t\right] \\ &\le \left(\mathbb{E}\left[\int_0^T \rho(t)\|\tilde{X}(t)\|_H^2 \mathrm{d}t\right]\right)^{\frac{1}{2}} \liminf_{k\to\infty} \left(\mathbb{E}\left[\int_0^T \rho(t)\|X^{(n_k)}(t)\|_H^2 \mathrm{d}t\right]\right)^{\frac{1}{2}}.\end{aligned}$$

Since $X = \tilde{X} = \bar{X}$ $\mathrm{d}t \otimes P$-a.e. on $[0, T] \times \Omega$ and $\left(\mathbb{E}\left[\int_0^T \rho(t)\|\bar{X}(t)\|_H^2 \mathrm{d}t\right]\right)^{\frac{1}{2}} < \infty$, this implies

$$\mathbb{E}\left[\int_0^T \rho(t)\|\bar{X}(t)\|_H^2 \mathrm{d}t\right]^{\frac{1}{2}} \le \liminf_{k\to\infty} \left(\mathbb{E}\left[\int_0^T \rho(t)\|X^{(n_k)}(t)\|_H^2 \mathrm{d}t\right]\right)^{\frac{1}{2}}. \tag{3.7}$$

By using Itô's formula for the expected value (c.f. [16, Remark 4.2.8]) the product rule and Fubini we obtain for $\lambda \ge 0$ that

$$\begin{aligned}&\mathbb{E}[e^{-\lambda t}\|X(t)\|_H^2] - \mathbb{E}[\|\psi(0)\|_H^2] \\ &= \mathbb{E}\left[\int_0^t e^{-\lambda s}\left(2_{V^*}\langle Y(s), \bar{X}(s)\rangle_V + \|Z(s)\|^2_{L_2(U,H)} - \lambda\|X(s)\|_H^2\right) \mathrm{d}s\right].\end{aligned} \tag{3.8}$$

Let $\phi \in L^p([-r_0, T] \times \Omega, \mathrm{d}t \otimes \Omega; V)$ such that $\phi(\omega, \cdot) \in C([-r_0, T]; H)$ for P-a.e. $\omega \in \Omega$ and $\mathbb{E}\left[\|\phi_t\|^2_{\mathcal{C}(H)}\right] < \infty$ for all $t \in [0, T]$ (This implies just like in 2.1.2 that $t \mapsto \mathcal{L}_{\phi_t}$ is continuous.). By using Itô's formula for the expected value in the case $V = H = V^* = H_{n_k}$ and defining $\mu_t^{(n)} := \mathcal{L}_{X_t^{(n)}}$, $\nu_t := \mathcal{L}_{\phi_t}$, $t \in [0, T]$, we deduce that

$$
\begin{aligned}
&\mathbb{E}[e^{-\lambda t}\|X^{(n_k)}(t)\|_H^2] - \mathbb{E}[\|P_{n_k}\psi(0)\|_H^2] \\
=&\mathbb{E}\bigg[\int_0^t e^{-\lambda s}\Big(2_{V^*}\langle P_{n_k}A(s,X_s^{(n_k)},\mu_s^{(n_k)}),X^{(n_k)}(s)\rangle_V \\
&\qquad\qquad + \|P_{n_k}B(s,X_s^{(n_k)},\mu_s^{(n_k)})\tilde{P}_{n_k}\|_{L_2(U,H)}^2 - \lambda\|X^{(n_k)}(s)\|_H^2\Big)\mathrm{d}s\bigg] \\
\le&\mathbb{E}\bigg[\int_0^t e^{-\lambda s}\Big(2_{V^*}\langle A(s,X_s^{(n_k)},\mu_s^{(n_k)}),X^{(n_k)}(s)\rangle_V \\
&\qquad\qquad + \|B(s,X_s^{(n_k)},\mu_s^{(n_k)})\|_{L_2(U,H)}^2 - \lambda\|X^{(n_k)}(s)\|_H^2\Big)\mathrm{d}s\bigg] \\
=&\mathbb{E}\bigg[\int_0^t e^{-\lambda s}\bigg(2_{V^*}\langle A(s,X_s^{(n_k)},\mu_s^{(n_k)}) - A(s,\phi_s,\nu_s),X^{(n_k)}(s)-\phi(s)\rangle_V \\
&\qquad\qquad + \|B(s,X_s^{(n_k)},\mu_s^{(n_k)}) - B(s,\phi_s,\nu_s)\|_{L_2(U,H)}^2 \\
&\qquad\qquad - \lambda\|X^{(n_k)}(s)-\phi(s)\|_H^2\bigg)\mathrm{d}s\bigg] \qquad\qquad (3.9)\\
&+\mathbb{E}\bigg[\int_0^t e^{-\lambda s}\Big(2_{V^*}\langle A(s,\phi_s,\nu_s),X^{(n_k)}(s)\rangle_V \\
&\qquad\qquad + 2_{V^*}\langle A(s,X_s^{(n_k)},\mu_s^{(n_k)}) - A(s,\phi_s,\nu_s),\phi(s)\rangle_V \\
&\qquad\qquad - \|B(s,\phi_s,\nu_s)\|_{L_2(U,H)}^2 + 2\langle B(s,X_s^{(n_k)},\mu_s^{(n_k)}),B(s,\phi_s,\nu_s)\rangle_{L_2(U,H)} \\
&\qquad\qquad - 2\lambda\langle X^{(n_k)}(s),\phi(s)\rangle_H + \lambda\|\phi(s)\|_H^2\Big)\mathrm{d}s\bigg].
\end{aligned}
$$

By the definition of the Wasserstein distance, it is $\mathbb{W}_2(\mu_t^{(n)},\nu_t) \le \mathbb{E}\big[\|X_t^{(n_k)} - \phi_t\|_{L_H^2}^2\big]$ for all $t \in [0,T]$. This together with Lemma 3.3.4 and (H3) implies for $\lambda := 2\beta r_0$

$$
\begin{aligned}
&\mathbb{E}\bigg[\int_0^t e^{-\lambda s}\Big(2_{V^*}\langle A(s,X_s^{(n_k)},\mu_s^{(n_k)}) - A(s,\phi_s,\nu_s),X^{(n_k)}(s)-\phi(s)\rangle_V \\
&\qquad + \|B(s,X_s^{(n_k)},\mu_s^{(n_k)}) - B(s,\phi_s,\nu_s)\|_{L_2(U,H)}^2 - \lambda\|X^{(n_k)}(s)-\phi(s)\|_H^2\Big)\mathrm{d}s\bigg] \\
&\le \mathbb{E}\bigg[\beta\int_0^t e^{-\lambda s}\Big(\|X_s^{(n_k)}-\phi_s\|_{L_H^2}^2 + \mathbb{W}_2(\mu_s^{(n)},\nu_s) - \lambda\|X^{(n_k)}(s)-\phi(s)\|_H^2\Big)\mathrm{d}s\bigg] \\
&\le \mathbb{E}\bigg[\int_0^t e^{-\lambda s}\left(2\beta r_0 - \lambda\right)\|X^{(n_k)}(s)-\phi(s)\|_H^2\mathrm{d}s\bigg] + 2\beta r_0 e^{\lambda r_0}\mathbb{E}\left[\|X_0^{(n_k)}-\phi_0\|_{L_H^2}^2\right] \\
&\le 2\beta r_0 e^{\lambda r_0}\mathbb{E}\left[\|X_0^{(n_k)}-\phi_0\|_{L_H^2}^2\right],
\end{aligned}
\tag{3.10}
$$

for all $t \in [0,T]$. By inserting (3.10) in (3.9) and letting $k \to \infty$, we conclude by (i)-(iii), Fubini's theorem and (3.7) that for every non-negative $\rho \in L^\infty([0,T], \mathrm{d}t; \mathbb{R})$

$$\begin{aligned}
&\mathbb{E}\bigg[\int_0^T \rho(t)\{e^{-\lambda t}\|X(t)\|_H^2 - \|\psi(0)\|_H^2\}\mathrm{d}t\bigg] \\
\le&\mathbb{E}\bigg[\int_0^T \rho(t)\bigg\{\int_0^t e^{-\lambda s}\Big(2_{V^*}\langle A(s,\phi_s,\nu_s), \bar{X}(s)\rangle_V + 2_{V^*}\langle Y(s) - A(s,\phi_s,\nu_s), \phi(s)\rangle_V \\
&\quad - \|B(s,\phi_s,\nu_s)\|_{L_2(U,H)}^2 + 2\langle Z(s), B(s,\phi_s,\nu_s)\rangle_{L_2(U,H)} - 2\lambda\langle X(s),\phi(s)\rangle_H \\
&\quad + \lambda\|\phi(s)\|_H^2\Big)\mathrm{d}s\bigg\}\mathrm{d}t\bigg] + \bigg(\int_0^T \rho(t)\mathrm{d}t\bigg)2\beta r_0 e^{\lambda r_0}\mathbb{E}\left[\|X_0 - \phi_0\|_{L_H^2}^2\right].
\end{aligned}$$

Inserting (3.8) for the left-hand site rearranging and defining $L = L(\rho,\beta,r_0,T) := \Big(\int_0^T \rho(t)\mathrm{d}t\Big)2\beta r_0 e^{\lambda r_0}T$ we arrive at

$$\begin{aligned}
&\mathbb{E}\bigg[\int_0^T \rho(t)\bigg\{\int_0^t e^{-\lambda s}\Big(2_{V^*}\langle Y(s) - A(s,\phi_s,\nu_s), \tilde{X}(s) - \phi(s)\rangle_V \\
&\qquad + \|B(s,\phi_s,\nu_s) - Z(s)\|_{L_2(U,H)}^2 - \lambda\|X(s) - \phi(s)\|_H^2\Big)\mathrm{d}s\bigg\}\mathrm{d}t\bigg] \qquad (3.11)\\
&\le L\mathbb{E}\left[\|X_0 - \phi_0\|_{L_H^2}^2\right]
\end{aligned}$$

Taking $\phi = \bar{X}$ and noting that $\bar{X}_0 = \psi = X_0$ P-a.s. we obtain from (3.11) that

$$\mathbb{E}\bigg[\int_0^T \rho(t)\bigg\{\int_0^t e^{-\lambda s}\|B(s,\bar{X}_s,\mathcal{L}_{\bar{X}_s}) - Z(s)\|_{L_2(U,H)}^2\bigg\}\mathrm{d}t\bigg] \le 0$$

and therefore $B(\cdot,\bar{X}_\cdot,\mathcal{L}_{\bar{X}_\cdot}) = Z \quad \mathrm{d}t \otimes P - a.e.$
Finally we take $\phi = \bar{X} - \epsilon\tilde{\phi}v$ for $\epsilon > 0$, $v \in V$ and $\tilde{\phi} \in L^\infty([-r_0,T]\times\Omega, \mathrm{d}t \otimes P; \mathbb{R})$ with $\tilde{\phi}(\omega,\cdot)$ continuous for P-a.e $\omega \in \Omega$, $\mathbb{E}\left[\|\tilde{\phi}_t\|_{\mathcal{C}(\mathbb{R})}^2\right] < \infty$ for all $t \in [0,T]$. (3.11) now implies

$$\begin{aligned}
&\epsilon\Bigg(\mathbb{E}\bigg[\int_0^T \rho(t)\bigg\{\int_0^t e^{-\lambda s}\left(2_{V^*}\langle Y(s) - A(s,\phi_s,\nu_s), \tilde{\phi}(s)v\rangle_V - \epsilon\|\tilde{\phi}(s)v\|_H^2\right)\mathrm{d}s\bigg\}\mathrm{d}t\bigg]\Bigg) \\
&\quad \le \epsilon^2 L\mathbb{E}\left[\|\tilde{\phi}v\|_{L_H^2}^2\right]. \qquad (3.12)
\end{aligned}$$

Dividing (3.12) by ϵ and taking $\epsilon \to 0$ we obtain by Lebesgue's dominated convergence theorem, (H1) and (H4) that

$$\mathbb{E}\bigg[\int_0^T \rho(t)\bigg\{\int_0^t e^{-\lambda s}2_{V^*}\langle Y(s) - A(s,\bar{X}_s,\mathcal{L}_{\bar{X}_s}), \tilde{\phi}(s)v\rangle_V\mathrm{d}s\bigg\}\mathrm{d}t\bigg] \le 0.$$

Replacing $\tilde{\phi}$ with $-\tilde{\phi}$ leads to

$$\mathbb{E}\bigg[\int_0^T \rho(t)\bigg\{\int_0^t e^{-\lambda s}2_{V^*}\langle Y(s) - A(s,\bar{X}_s,\mathcal{L}_{\bar{X}_s}), \tilde{\phi}(s)v\rangle_V\mathrm{d}s\bigg\}\mathrm{d}t\bigg] = 0.$$

By the arbitrariness of ρ, $\tilde{\phi}$ and v we conclude with that $A(\cdot, \bar{X}_\cdot, \mathcal{L}_{\bar{X}_\cdot}) = Y$. This completes the proof of existence. The uniqueness follows directly from the theorem below. □

Theorem 3.3.5. *Consider the situation of Theorem 3.1.3 and let X, Y be two solutions of (3.1) in the sense of Definition 3.1.1. Then for $\beta \geq 0$ as in (H3) and $t \in [0,T]$:*

(i)

$$\mathbb{E}\left[\|X(t) - Y(t)\|_H^2\right] \leq \left(1 + r_0^2 e^{2\beta r_0^2}\right) e^{2\beta r_0 t} \mathbb{E}\left[\|X_0 - Y_0\|_{\mathcal{C}(H)}^2\right],$$

(ii)

$$\mathbb{E}\Bigg[\sup_{s\in[0,t]} \|X_s - Y_s\|_{\mathcal{C}(H)}^2\Bigg] \leq \inf_{\epsilon\in(0,1)} \left(\left(\frac{\mathbb{E}\Big[\|X_0 - Y_0\|_{\mathcal{C}(H)}^2\Big]}{1-\epsilon} \right) \exp\left[\frac{2r_0}{1-\epsilon}\left(1 + \frac{6}{\epsilon}\right)\beta t\right] \right).$$

Proof. (i) By our definition of solution (Definition 3.1.1) we can apply Itô's formula to $X - Y$ and the product rule to obtain for $t \in [0,T]$

$$\begin{aligned}
e^{-2\beta r_0 t}&\mathbb{E}\Big[\|X(t) - Y(t)\|_H^2\Big] \\
&= \mathbb{E}\left[\|X(0) - Y(0)\|_H^2\right] \\
&\quad + \mathbb{E}\Bigg[\int_0^t e^{-2\beta r_0 s} 2_{V^*}\langle A(s, \bar{X}_s, \mathcal{L}_{\bar{X}_s}) - A(s, \bar{Y}_s, \mathcal{L}_{\bar{Y}_s}), \bar{X}(s) - \bar{Y}(s)\rangle_V \\
&\qquad\qquad + \|B(s, \bar{X}_s, \mathcal{L}_{\bar{X}_s}) - B(s, \bar{Y}_s, \mathcal{L}_{\bar{Y}_s})\|_{L_2(U,H)}^2 \mathrm{d}s\Bigg] \\
&\quad - 2\beta r_0 \int_0^t e^{-2\beta r_0 s} \mathbb{E}\left[[\|X(s) - Y(s)\|_H^2\right] \mathrm{d}s \\
&\leq \mathbb{E}\left[\|X(0) - Y(0)\|_H^2\right] + \mathbb{E}\left[\int_0^t e^{-2\beta r_0 s} \|X_s - Y_s\|_{L_H^2}^2 + \mathbb{W}_2(\mathcal{L}_{\bar{X}_s}, \mathcal{L}_{\bar{Y}_s}) \mathrm{d}s\right] \\
&\quad - 2\beta r_0 \int_0^t e^{-2\beta r_0 s} \mathbb{E}\left[[\|X(s) - Y(s)\|_H^2\right] \mathrm{d}s \\
&\leq \mathbb{E}\left[\|X(0) - Y(0)\|_H^2\right] + r_0 e^{2\beta r_0^2} \mathbb{E}\left[\|X_0 - Y_0\|_{L_H^2}^2\right] \\
&\leq \left(1 + r_0^2 e^{2\beta r_0^2}\right) \mathbb{E}\left[\|X_0 - Y_0\|_{\mathcal{C}(H)}^2\right].
\end{aligned}$$

Where we used (H3) to obtain the first estimate. To obtain the second estimate we used that by the definition of the Wasserstein distance $\mathbb{W}_2(\mathcal{L}_{\bar{X}_s}, \mathcal{L}_{\bar{Y}_s}) \leq \mathbb{E}[\|X_s - Y_s\|_{L_H^2}^2]$ together with Lemma 3.3.4.
Multiplying with $e^{2\beta r_0 t}$ yields (i).

(ii) The proof is similar to the proof of Theorem 2.1.6 (b). ☐

Note that in the proof of Theorem 3.3.5 (i) only the first estimate in (H3) is used. In fact this is only needed to show the existence of solutions to the finite dimensional equation (3.4). That means that if it were possible to proof the existence of solutions to finite dimensional path-distribution dependent SDE's without the "Lipschitz-condition"

$$\int_0^t \|\sigma(s,\xi_s,\mu_s)-\sigma(s,\eta_s,\nu_s)\|_{\mathrm{HS}}^2 \le \beta(t)\int_0^t \|\xi_s-\eta_s\|_\infty^2+\mathbb{W}_2(\mu_s,\nu_s)^2\mathrm{d}s$$

in (H3) in chapter 2, the second part of (H3) in this chapter could be dropped and we would have existence and uniqueness of solutions to (3.1) in the sense of Definition 3.1.1 by the arguments presented in this chapter.

By Theorem 3.1.3 we know that $\mathbb{E}[\sup_{t\in[-r_0,T]}\|X(t)\|_H^2]<\infty$. The final proposition of this chapter gives a more precise estimate of $\mathbb{E}[\sup_{t\in[-r_0,T]}\|X(t)\|_H^2]$. The proof is analogous to the proof of Theorem 2.1.6 (b) (ii).

Proposition 3.3.6. *Consider the situation of Theorem 3.1.3 and let X,Y be two solutions of (3.1) in the sense of Definition 3.1.1. Then*

$$\mathbb{E}\left[\sup_{r\in[-r_0,T]}\|X(r)\|_H^2+\int_0^T\|X(s)\|_V^p\,ds\right]\le C\left(1+\mathbb{E}\left[\|X_0\|_{\mathcal{C}(H)}^2\right]\right)$$

for some $C>0$.

Proof. The proof is essentially the same as the proof of Theorem 2.1.6 (b) (ii). ☐

Acknowledgment

I want to thank Prof. Dr. Michael Röckner and Dr. Tatjana Pasurek for fruitful discussion and helpful comments during the preparation of this paper. Support of the DFG through CRC 1283 is also gratefully acknowledged.

References

[1] H. W. Alt. *Lineare Funktionalanalysis.* Springer-Lehrbuch Masterclass. Springer Berlin Heidelberg, Berlin, Heidelberg, sixth edition, 2012.
[2] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.
[3] V. Barbu and M. Röckner. From nonlinear Fokker-Planck equations to solutions of distribution dependent SDE. Eprint: arXiv:1808.10706[math.PR]. To appear in Ann. Prob., 2018.
[4] H. Bauer. *Measure and integration theory*, volume 26 of *De Gruyter Studies in Mathematics.* Walter de Gruyter & Co., Berlin, 2001. Translated from the German by Robert B. Burckel.
[5] H. Bauer. *Wahrscheinlichkeitstheorie.* de Gruyter Lehrbuch. [de Gruyter Textbook]. Walter de Gruyter & Co., Berlin, fifth edition, 2002.
[6] V. I. Bogachev, N. V. Krylov, M. Röckner, and S. V. Shaposhnikov. *Fokker-Planck-Kolmogorov equations*, volume 207 of *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, 2015.

[7] R. Carmona and F. Delarue. *Probabilistic theory of mean field games with applications. I: Mean field FBSDEs, control, and games, II: Mean field games with common noise and master equations*, volume 83 & 84 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2018.

[8] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 152 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 2014.

[9] J. L. Doob. *Stochastic processes*. Wiley series in probability and mathematical statistics : Probability and mathematical statistics. Wiley, New York, seventh edition, 1967.

[10] G. dos Reis, G. Smith, and P. Tankov. Importance sampling for McKean-Vlasov SDEs. Eprint: arXiv:1803.09320[math.PR], 2018.

[11] W. Hammersley, D. Šiška, and L. Szpruch. McKean-Vlasov SDEs under Measure Dependent Lyapunov Conditions. Eprint: arXiv:1802.03974[math.PR], 2018.

[12] X. Huang, M. Röckner, and F.-Y. Wang. Nonlinear Fokker-Planck equations for probability measures on path space and path-distribution dependent SDEs. *Discrete Contin. Dyn. Syst.*, 39(6):3017–3035, 2019.

[13] X. Huang and F.-Y. Wang. Distribution dependent SDEs with singular coefficients. *Stochastic Process. Appl.*, 129(11):4747–4770, 2019.

[14] A. S. Kechris. *Classical descriptive set theory*, volume 156 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.

[15] N. V. Krylov. On Kolmogorov's equations for finite-dimensional diffusions. In *Stochastic PDE's and Kolmogorov equations in infinite dimensions (Cetraro, 1998)*, volume 1715 of *Lecture Notes in Math.*, pages 1–63. Springer, Berlin, 1999.

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

[17] S. Mehri, M. Scheutzow, W. Stannat, and B. Z. Zangeneh. Propagation of chaos for stochastic spatially structured neuronal networks with delay driven by jump diffusions. *Ann. Appl. Probab.*, 30(1):175–207, 2020.

[18] Y. S. Mishura and A. Y. Veretennikov. Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations. Eprint: arXiv:1603.02212[math.PR], 2016.

[19] K. R. Parthasarathy. *Probability measures on metric spaces*. Probability and Mathematical Statistics, No. 3. Academic Press, Inc., New York-London, 1967.

[20] M. Röckner and X. Zhang. Well-posedness of distribution dependent SDEs with singular drifts. Eprint: arXiv:1809.02216[math.PR], 2018.

[21] M. Röckner, R. Zhu, and X. Zhu. Existence and uniqueness of solutions to stochastic functional differential equations in infinite dimensions. *Nonlinear Anal.*, 125:358–397, 2015.

[22] D. W. Stroock. *Probability theory: An analytic view*. Cambridge University Press, Cambridge, second edition, 2011.

[23] C. Villani. *Optimal transport: Old and new*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009.

[24] H. von Weizsäcker and G. Winkler. *Stochastic integrals: An introduction*. Advanced Lectures in Mathematics. Friedr. Vieweg & Sohn, Braunschweig, 1990.

[25] F.-Y. Wang. *Harnack inequalities for stochastic partial differential equations*. SpringerBriefs in Mathematics. Springer, New York, 2013.

[26] F.-Y. Wang. Distribution dependent SDEs for Landau type equations. *Stochastic Process. Appl.*, 128(2):595–621, 2018.

[27] K. Yosida. *Functional analysis*, volume 123 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin-New York, third edition, 1971.