On the analysis of path-dependent functionals of stochastic PDEs

Dissertation
zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt im
Fachbereich Mathematik und Naturwissenschaften der Universität Kassel

Von
Saeed Hadjizadeh

Kassel, April 2023
Datum der Disputation: 25.09.2023

Supervised by Prof. Dr. Felix Lindner

Acknowledgements

First and foremost, I would like to thank my supervisor Prof. Dr. Felix Lindner for his consistent support and guidance during the compilation of this dissertation.

I am really grateful that I have been lucky enough to work with wonderful people at both RPTU Kaiserslautern and the University of Kassel during my work on this thesis. My special thanks to Dr. Markus Antoni who reviewed early drafts of this writing and provided valuable feedback.

To conclude, I cannot forget to thank my family and friends for all the unconditional support and encouragement.

Abstract

Weak approximation methods for stochastic partial differential equations (SPDEs) are concerned with approximating the probability distribution of the solution process rather than the realizations of the solution process itself. In this thesis, we provide new results and methods concerning the weak error analysis of numerical approximations of path-dependent functionals of solution processes of SPDEs. Two separate approaches to analyzing weak approximation errors are considered: the Itô calculus approach and the Malliavin calculus approach. In the context of the Itô calculus approach to weak error analysis, we develop and apply a novel path-dependent mild Itô formula suitable for the analysis of path-dependent functionals of mild solutions of SPDEs and their numerical approximations. In the context of the Malliavin calculus approach to SPDEs, we analyze spectral Galerkin projections of mild solutions of SPDEs with multiplicative noise and establish estimates for the corresponding weak approximation errors for a general class of path-dependent functionals. The considered functionals are defined on the Bochner space of paths that are q-integrable with respect to a given finite Borel measure, for a suitable integrability parameter $q \in (1, 2]$.

Zusammenfassung

Schwache Approximationsverfahren für stochastische partielle Differentialgleichungen dienen der Approximation der Wahrscheinlichkeitsverteilung des Lösungsprozesses, wobei dessen einzelne Realisierungen eine untergeordnete Rolle spielen. Diese Dissertation liefert neue Resultate und Methoden im Hinblick auf die mathematische Analyse schwacher Approximationsfehler für numerische Approximationen pfadabhängiger Funktionale von Lösungsprozessen stochastischer partieller Differentialgleichungen. Dabei werden zwei separate Ansätze zur Analyse schwacher Approximationsfehler verfolgt: ein auf dem Itô-Kalkül basierender Ansatz und ein auf dem Malliavin-Kalkül fußender Ansatz. Im Kontext des auf dem Itô-Kalkül basierenden Ansatzes entwickeln wir eine neuartige, pfadabhängige milde Itô-Formel, die sich für die Analyse pfadabhängiger Funktionale milder Lösungen stochastischer partieller Differentialgleichungen und derer numerischer Approximationen eignet. Im Kontext des auf dem Malliavin-Kalkül fußenden Ansatzes untersuchen wir Galerkin-Projektionen milder Lösungen stochastischer partieller Differentialgleichungen mit multiplikativem Rauschen und leiten Abschätzungen für die entsprechenden schwachen Approximationsfehler für eine allgemeine Klasse pfadabhängiger Funktionale her. Die betrachteten Funktionale sind definiert auf dem Bochner-Raum aller bezüglich eines gegebenen endlichen Borel-Maßes q-integrierbaren Pfade, wobei $q \in (1, 2]$ ein geeigneter Integrabilitätsparameter ist.

Contents

Chapter 1

Introduction

Stochastic evolution equations (SEEs) such as stochastic ordinary differential equations (SODEs) and stochastic partial differential equations of evolutionary type (SPDEs) are frequently used to model stochastic dynamics of real-world systems in finite- and infinite-dimensional state spaces. These systems range from physical phenomena over biological systems to finance. As the solutions of SEEs are usually not known explicitly, numerical approximations are essential with regard to, e.g., the extraction of relevant information from the models or statistical parameter fitting.

Weak approximation methods for SEEs are concerned with approximating the probability distribution of the solution process rather than the realizations of the solution process itself. In this context, the quantity of interest is typically some functional of the solution process. While the majority of research on the analysis of weak approximation errors is focused on functionals which depend only on evaluations of the solution process at a fixed time point, there has been growing interest in path-dependent functionals in recent years. In this thesis, we provide new results and methods concerning the weak error analysis of numerical approximations of path-dependent functionals of solution processes of SPDEs.

In this introductory chapter, we first specify a suitable setting and introduce some notation concerning SEEs and weak approximation errors in Section 1.1. We then briefly sketch two of the main approaches to weak error analysis in the research literature in Section 1.2 and Section 1.3. The main results and the overall structure of this thesis are summarized in Section 1.4.

1.1 Weak approximation of stochastic evolution equations

To set the stage, let us consider the SEE

$$\begin{cases} \mathrm{d}X(t) = \big[AX(t) + F\big(X(t)\big)\big]\mathrm{d}t + B\big(X(t)\big)\mathrm{d}W(t), & t \in [0,T], \\ X(0) = \xi, \end{cases} \tag{1.1.1}$$

where $T \in (0,\infty)$ is a finite time horizon and the solution process $X = (X(t))_{t\in[0,T]}$ takes values in a separable real Hilbert space H. Moreover, $A\colon D(A) \subset H \to H$ is the generator of a strongly continuous semigroup $(e^{tA})_{t\geq 0}$ of bounded linear operators on H, $W = (W(t))_{t\in[0,T]}$ is a cylindrical Id_U-Wiener process, where U is another separable real Hilbert space, and the mappings $F\colon H \to H$ and $B\colon H \to L(U,H)$ are assumed to satisfy certain measurability and regularity properties. Here $L(U,H)$ denotes the space of bounded linear operators from U to H. The initial condition ξ in (1.1.1) is an H-valued random variable that is assumed to satisfy specific measurability and integrability properties as well. In this thesis, we follow the so-called semigroup approach to SEEs, which amounts to reformulating (1.1.1) in the mild form

$$X(t) = e^{tA}\xi + \int_0^t e^{(t-s)A}F(X(s))\,\mathrm{d}s + \int_0^t e^{(t-s)A}B(X(s))\,\mathrm{d}W(s), \quad t \in [0,T]. \tag{1.1.2}$$

A suitably measurable H-valued process $X = (X(t))_{t\in[0,T]}$ satisfying (1.1.2) is called a mild solution of (1.1.1). Note that the precise definitions of all objects and concepts mentioned above are presented in Chapter 2 below.

In the case where H is infinite-dimensional and the linear operator A is unbounded, SEE (1.1.1) is well-suited to describe certain SPDEs in an abstract form. For example, H might be the space $L^2(\mathcal{O})$ of square-integrable real-valued functions on a convex domain $\mathcal{O} \subset \mathbb{R}^2$ with polygonal boundary and $-A$ might be an elliptic second order partial differential operator with Sobolev space domain $D(A) = H^2(\mathcal{O}) \cap H^1_0(\mathcal{O})$. Our standard references for the semigroup approach to SEEs in infinite dimensions are [16, 23, 49].

In the case where $H = \mathbb{R}^d$ is finite-dimensional, SEE (1.1.1) describes an SODE. In particular, the linear operator A in (1.1.1) is always bounded with $D(A) = H$ and can thus be included in the function F. We can therefore set $A = 0$, so that in (1.1.2) we obtain that $e^{tA} = \mathrm{Id}_H$ is the identity operator for all $t \in [0,T]$.

Throughout this introductory chapter, the H-valued process $\tilde{X} = (\tilde{X}(t))_{t\in[0,T]}$ denotes a given numerical approximation of the solution process $X = (X(t))_{t\in[0,T]}$. For example, $\tilde{X}$ might be a time-interpolated solution of an explicit or implicit Euler scheme in the SODE case or a time-interpolated solution of a temporal, spatial, or spatio-temporal discretization scheme in the SPDE case; see, e.g., Example 3.1.3 and Example 3.1.4 in Section 3.2 below. In this thesis we mainly focus on the case where $\tilde{X}$ represents a spatial semi-discretization of X in form of a Galerkin projection onto a finite-dimensional subspace of H, given by $\tilde{X}(t) = P_N X(t)$ with discretization parameter $N \in \mathbb{N}$. Here we essentially assume that A is a diagonal operator with eigenbasis $(e_n)_{n\in\mathbb{N}} \subset H$ and associated decreasing sequence of eigenvalues $(\lambda_n)_{n\in\mathbb{N}} \subset (-\infty, 0)$ and that the projection operator $P_N\colon H \to H$ is given by $P_N(v) = \sum_{n=1}^N \langle e_n, v\rangle_H\, e_n$, $v \in H$. Note that we follow the terminology in [13] and distinguish between Galerkin projections, which are obtained by applying the projection operator P_N to the solution process X of (1.1.2), and Galerkin approximations, which are defined as solutions of projected SEEs that do not involve the unknown solution X of (1.1.2); see, e.g., Example 3.1.4 below. While such partial discretizations are obviously not directly implementable and computable, Galerkin discretizations as well as Galerkin projections play a key role as fundamental building blocks for the construction and analysis of a large class of fully-discrete numerical approximation schemes for SPDEs.

In the context of weak approximation of SEEs, one is typically interested in approximating a functional of the distribution of the state of the solution process at a fixed time point such as $\mathbb{E}[f(X(T))]$, where $f\colon H \to \mathbb{R}$ is a given mapping. For instance, $\mathbb{E}[f(X(T))]$ could represent the expected payoff for a financial derivative in the SODE case or spatial correlations of the solution process of the form $\mathbb{E}[f(X(T))] = \mathbb{E}\big[\,\langle h_1, X(T)\rangle_H \cdot \langle h_2, X(T)\rangle_H\,\big]$ with test functions $h_1, h_2 \in H$ in the SPDE case. If we use the process $\tilde{X} = (\tilde{X}(t))_{t\in[0,T]}$ to approximate the solution $X = (X(t))_{t\in[0,T]}$ of (1.1.2), then the corresponding weak approximation error is given by

$$\big|\mathbb{E}[f(X(T))] - \mathbb{E}[f(\tilde{X}(T))]\big|. \tag{1.1.3}$$

The analysis of weak approximation errors is essential for the assessment of the consistency and the efficiency of a given approximation scheme, and a particular interest lies in the derivation of sharp upper bounds. It is worth noting here that even in the case where $\tilde{X}$ stems from a fully-discrete and directly implementable numerical scheme, the practical approximation of the quantity of interest $\mathbb{E}[f(X(T))]$ further requires the approximation of the expectation operator $\mathbb{E}[\,\cdot\,]$ via, e.g., a suitable Monte Carlo method [25, 46]. This leads to an additional statistical approximation error, the analysis of which does not lie within the scope of this thesis. The analysis of weak errors of the type (1.1.3) is quite far developed in the SODE case; see, e.g. [18, 34, 45, 55] and the references therein. Weak error analysis in the SPDE case has come into focus in research in the last 10 to 15 years; see, e.g., [1, 8, 9, 13, 15, 17, 19, 24, 37, 38, 44]. In both cases it is typically found that for sufficiently regular test functions f the rate of convergence to zero of the weak approximation error (1.1.3) as the discretization is refined is twice the rate of convergence of the strong approximation error

$$\Big(\mathbb{E}\Big[\big\|X(T) - \tilde{X}(T)\big\|_H^p\Big]\Big)^{1/p}, \tag{1.1.4}$$

for certain integrability parameters $p \in [1, \infty)$. Knowledge of both weak and strong convergence rates is important for a proper conceptualization and application of multilevel Monte Carlo methods [25, 46].

As mentioned before, most research concerning the weak error analysis of numerical approximations of SEEs is focused on approximating functionals which depend only on evaluations of the solution process at a fixed time point. In contrast, the analysis of weak approximation errors involving path-dependent functionals of the form $\mathbb{E}[f(X)]$ is not as well established. Here the mapping f is a functional defined on a suitable path space, e.g., the space $C([0,T],H)$ of continuous paths, and the solution X of (1.1.2) is identified with the random variable $\omega \mapsto \big(X(\omega,t)\big)_{t\in[0,T]}$ with values in the path space, where $\omega \in \Omega$ is taken from the sample space of the underlying probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and where $X(\omega,t) = X(t)(\omega)$ denotes a realization of the H-valued random variable $X(t)$. The corresponding path-dependent weak approximation error then reads

$$\big|\mathbb{E}[f(X)] - \mathbb{E}[f(\tilde{X})]\big|. \tag{1.1.5}$$

In this thesis, we provide new results and tools for the analysis of weak approximation errors of type (1.1.5) concerning numerical approximations of path-dependent functionals of mild solutions of SPDEs.

1.2 Itô calculus and weak error analysis

A by now classical approach to weak error analysis for numerical approximations of SEEs is based on Itô calculus in combination with Kolmogorov equations associated to the SEE; see, e.g., [34, 55, 56]. In the sequel, we first sketch a variant of this approach for the well established case of SODEs and weak errors of type (1.1.3), which depend only on evaluations of the solution process at a fixed time point. Thereafter, we present available extensions of the described strategy to the case of SPDEs and weak errors of type (1.1.3) and to the case of SODEs and path-dependent weak errors of type (1.1.5).

Recall that in the SODE case we may assume $e^{tA} = \mathrm{Id}_H$ in (1.1.2), so that the solution process $X = (X(t))_{t\in[0,T]}$ is an Itô process of the form

$$X(t) = \xi + \int_0^t \Psi(s)\,\mathrm{d}s + \int_0^t \Phi(s)\,\mathrm{d}W(s), \quad t \in [0,T], \tag{1.2.1}$$

where the H-valued and $L(U,H)$-valued integrand processes $\Psi = (\Psi(s))_{s\in[0,T]}$ and $\Phi = (\Phi(s))_{s\in[0,T]}$ are given by $\Psi(s) = F(X(s))$ and $\Phi(s) = B(X(s))$, respectively. Itô calculus extends the methods of calculus to Itô processes of type (1.2.1) and a central result in this theory is Itô's change-of-variable formula [28, 29, 33, 54]. The standard Itô formula states that for sufficiently regular functions $g\colon [0,T]\times H\to\mathbb{R}$ and all $t\in[0,T]$ it holds almost surely that

$$\begin{aligned} g\big(t,X(t)\big) &= g(0,\xi) + \int_0^t \big[\partial_1 g\big(s,X(s)\big) + \partial_2 g\big(s,X(s)\big)\Psi(s)\big]\,\mathrm{d}s \\ &\quad + \int_0^t \partial_2 g\big(s,X(s)\big)\Phi(s)\,\mathrm{d}W(s) \\ &\quad + \frac{1}{2}\sum_{u\in\mathcal{U}}\int_0^t \partial_2^2 g\big(s,X(s)\big)\big(\Phi(s)u,\Phi(s)u\big)\,\mathrm{d}s. \end{aligned} \tag{1.2.2}$$

Here and below we denote for every $(s,x)\in[0,T]\times H$ by $\partial_1 g(s,x)$ and $\partial_2 g(s,x)$ the partial derivatives with respect to the first and the second argument of g, i.e., the time derivative and the spatial derivative, and we interpret for every $(s,x)\in[0,T]\times H$ the spatial derivatives $\partial_2 g(s,x)$ and $\partial_2^2 g(s,x)$ of first and second order in the usual way as linear and bilinear mappings from H to $\mathbb{R}$ and from $H\times H$ to $\mathbb{R}$, respectively. Moreover, $\mathcal{U}\subset U$ denotes an arbitrary orthonormal basis of the Hilbert space U associated to the driving Wiener process W. For a large class of numerical approximation schemes for SODEs such as explicit or drift-implicit Euler schemes and for suitable stochastic interpolations between the discrete time points, the time-interpolated approximation process $\tilde X=(\tilde X(t))_{t\in[0,T]}$ can be represented as an Itô process of the form (1.2.1) as well, with different integrand processes $\tilde\Psi$, $\tilde\Phi$ and possibly a different initial condition $\tilde\xi$ in place of Ψ, Φ, and ξ, respectively. Under suitable technical assumptions, formula (1.2.2) then remains valid with $\tilde X$, $\tilde\xi$, $\tilde\Phi$, and $\tilde\Psi$ in place of X, ξ, Φ, and Ψ.

When it comes to analyzing the weak error (1.1.3) in the described SODE setting via Itô calculus, a fruitful idea is to employ the specific function $g(t,x)=\mathbb{E}\big[f(X^x(T-t))\big]$, $(t,x)\in[0,T]\times H$, where $X^x=\big(X^x(t)\big)_{t\in[0,T]}$ denotes the solution of (1.1.2) with initial condition $x\in H$ in place of ξ. It is well-known that under sufficient regularity assumptions on F, B, and f, this particular choice of g represents the solution of the backward Kolmogorov equation with terminal condition f associated to (1.1.2). Moreover, if we assume for simplicity that the given initial condition ξ in (1.1.2) is non-random and thus an element of H, it follows directly from the definition of g that the weak error (1.1.3) can be rewritten as $\big|\mathbb{E}[g(0,\xi)-g(T,\tilde X(T))]\big|$. In combination with an application of Itô's formula to the process $\tilde X=(\tilde X(t))_{t\in[0,T]}$ and the function g, this leads to a useful weak error representation formula which may serve as a starting point for a thorough weak error analysis. For details and further aspects of this approach with regard to the Euler-Maruyama scheme for SODEs we refer to [34, Section 14.1] and [55, Sections 3.1–3.3]. Related results can be found, e.g., in [44, 45].

In the SPDE case, the solution process $X=(X(t))_{t\in[0,T]}$ in (1.1.2) is not a standard Itô process anymore, so that the standard Itô formula cannot be applied. As a consequence, the approach to weak error analysis based on Itô calculus described above cannot directly be extended to the this case. In this context, a useful extension of Itô calculus to mild solutions of SPDEs and their numerical approximations has been introduced in a systematic way in [15], where the authors analyze so-called mild Itô processes of the form

$$X(t) = S(t,0)\xi + \int_0^t S(t,s)\Psi(s)\,\mathrm{d}s + \int_0^t S(t,s)\Phi(s)\,\mathrm{d}W(s), \quad t \in [0,T]. \tag{1.2.3}$$

Here $S = (S(t,s))_{0\le s\le t\le T}$ is a strongly continuous evolution family of bounded linear operators on H and, as before, the integrand processes Ψ and Φ take values in H and $L(U,H)$, respectively, and are assumed to satisfy suitable measurability and integrability conditions; see Definition 2.4.9 in Chapter 2 below for details. Comparing (1.2.3) with (1.1.2), it is clear that the solution process in (1.1.2) is a mild Itô process of the form (1.2.3) with $S(t,s) = e^{(t-s)A}$, $\Psi(s) = F(X(s))$, and $\Phi(s) = B(X(s))$. The so-called mild Itô formula presented in [15] states that for any sufficiently regular mapping $g\colon [0,T]\times H \to \mathbb{R}$ and all $t \in [0,T]$ it holds almost surely that

$$\begin{aligned} g\big(t,X(t)\big) &= g(0,S(t,0)\xi) + \int_0^t \big[\partial_1 g\big(s,S(t,s)X(s)\big) + \partial_2 g\big(s,S(t,s)X(s)\big)S(t,s)\Psi(s)\big]\,\mathrm{d}s \\ &\quad + \int_0^t \partial_2 g\big(s,S(t,s)X(s)\big)S(t,s)\Phi(s)\,\mathrm{d}W(s) \\ &\quad + \frac{1}{2}\sum_{u\in\mathcal{U}}\int_0^t \partial_2^2 g\big(s,S(t,s)X(s)\big)\big(S(t,s)\Phi(s)u, S(t,s)\Phi(s)u\big)\,\mathrm{d}s. \end{aligned} \tag{1.2.4}$$

This generalization of the standard Itô formula involving the evolution operators $S(t,s) \in L(H)$, $0 \le s \le t \le T$, where $L(H) = L(H,H)$ denotes the space of bounded linear operators on H, is well-suited for the analysis of both mild solutions of SPDEs and their numerical approximations. Note that the difference between a general evolution family $S = (S(t,s))_{0\le s\le t\le T}$ and the specific family $(e^{(t-s)A})_{0\le s\le t\le T}$ associated to the operator semigroup $(e^{tA})_{t\ge 0}$ is that the operators of the evolution family depend explicitly on both variables s and t, and not necessarily only on the difference $t-s$ as in the semigroup case. The motivation for this additional technical complexity in the definition (1.2.3) of a mild Itô process is that numerical approximations of the semigroup $(e^{tA})_{t\ge 0}$ often form only an evolution family and not necessarily a semigroup; see, e.g., the discussion of the linear implicit Euler scheme in Example 3.1.3 in Chapter 3 below. In particular, the approximation process $\tilde{X} = (\tilde{X}(t))_{t\in[0,T]}$ can typically also be written in the form (1.2.3), with a different evolution family $\tilde{S}$, different integrand processes $\tilde{\Psi}$, $\tilde{\Phi}$, and possibly a different initial condition $\tilde{\xi}$ in place of S, Ψ, Φ, and ξ, respectively; compare Example 3.1.3 and Example 3.1.4 in Chapter 3 below. Under suitable technical assumptions, formula (1.2.4) then remains valid with $\tilde{X}$, $\tilde{\xi}$, $\tilde{S}$, $\tilde{\Psi}$, and $\tilde{\Phi}$ in place of X, ξ, S, Ψ, and Φ.

The mild Itô formula (1.2.4) has been successfully employed, e.g., in [13] and [32], to extend the Itô calculus approach for the analysis of weak approximation errors of the type (1.1.3) from the SODE case to a large class of semilinear SPDEs. The origins of this extension lie in the pioneering work [19], in which the authors study weak approximation errors for linear SPDEs with additive noise using Hilbert space-valued Itô calculus and Kolmogorov equations; compare also the closely related articles [24, 36–38, 41]. In the case of semilinear SPDEs, weak approximation errors have also been analyzed using a combination of the approach based on Itô calculus and Kolmogorov equations and the integration-by-parts formula from Malliavin calculus; see, e.g., [3, 8, 17]. It is worth noting here that in the special case where the approximation $\tilde{X} = (\tilde{X}(t))_{t\in[0,T]}$ of the solution process $X = (X(t))_{t\in[0,T]}$ is given by a Galerkin projection, i.e., $\tilde{X}(t) = P_N X(t)$ for some $N \in \mathbb{N}$, neither Kolmogorov equations nor Malliavin integration-by-parts are needed to derive sharp estimates for the weak approximation error (1.1.3) under standard assumptions on the semilinear SPDE (1.1.2) and the function f. Instead, in this simplified case it turns out to be sufficient to apply the mild Itô formula (1.2.4) directly with $g(t,x) = f(x)$, $(t,x) \in [0,T] \times H$; compare [13, Section 2]. Let us further remark that, besides the mild Itô formula (1.2.4), further generalizations of the standard Itô formula to solution processes of SPDEs have been presented in the research literature in the context of the variational approach to SPDEs; see, e.g. [26, 27, 45]. However, these generalizations are typically concerned only with specific functions g such as the squared norm $g(x) = \|x\|_H^2$, $x \in H$.

Next we return to the SODE case and consider path-dependent weak approximation errors of type (1.1.5). The standard Itô calculus approach to weak error analysis for SODEs is obviously not directly applicable in this case. Nevertheless, independently from questions concerning the analysis of weak approximation errors, a useful extension of Itô's calculus to path-dependent functionals has been established in [5, 11, 12, 20]. The key ingredient of this so-called functional Itô calculus is a path-dependent Itô formula, which we present in a slightly simplified form suitable for our purpose: Given an H-valued Itô process $X = (X(t))_{t\in[0,T]}$ with continuous sample paths of the form (1.2.1), we follow the notation in [5] and denote for every $t \in [0,T]$ by $X_t = (X(r \wedge t))_{r\in[0,T]}$ the process stopped at time t, i.e., $X_t(r) = X(r \wedge t)$ for all $r \in [0,T]$. Let $C([0,T],H)$ and $D([0,T],H)$ be the spaces of H-valued continuous paths and of H-valued càdlàg (right continuous with left limits) paths, respectively, both endowed with the uniform norm, and note that for every $t \in [0,T]$ we may identify the stopped process X_t with the $C([0,T],H)$-valued random variable $\omega \mapsto \big(X(\omega, r \wedge t)\big)_{r\in[0,T]}$, where $\omega \in \Omega$ is taken from the sample space of the underlying probability space and where $X(\omega, r \wedge t) = X(r \wedge t)(\omega)$ denotes a realization of the H-valued random variable $X(r \wedge t)$. In particular, observe that the family $(X_t)_{t\in[0,T]}$ thus represents a $C([0,T],H)$-valued stochastic process. In this setting, the so-called functional Itô formula states that for sufficiently regular mappings $g\colon [0,T] \times D([0,T],H) \to \mathbb{R}$

and all $t \in [0,T]$ it holds almost surely that

$$\begin{aligned} g(t, X_t) = g(0, X_0) &+ \int_0^t \partial_1 g(s, X_s)\,\mathrm{d}s \\ &+ \int_0^t \partial_2 g(s, X_s)\big(\mathbb{1}_{[s,T]}(\cdot)\Psi(s)\big)\,\mathrm{d}s \\ &+ \int_0^t \partial_2 g(s, X_s)\big(\mathbb{1}_{[s,T]}(\cdot)\Phi(s)\big)\,\mathrm{d}W(s) \\ &+ \frac{1}{2}\sum_{u\in\mathcal{U}} \int_0^t \partial_2^2 g(s, X_s)\big(\mathbb{1}_{[s,T]}(\cdot)\Phi(s)u, \mathbb{1}_{[s,T]}(\cdot)\Phi(s)u\big)\,\mathrm{d}s. \end{aligned} \tag{1.2.5}$$

Here for every $(s,x) \in [0,T] \times D([0,T],H)$ the spatial derivatives $\partial_2 g(s,x)$ and $\partial_2^2 g(s,x)$ of first and second order are assumed to be Fréchet derivatives and are interpreted as linear and bilinear mapping from $D([0,T],H)$ to $\mathbb{R}$ and from $D([0,T],H) \times D([0,T],H)$ to $\mathbb{R}$, respectively. Moreover, for every $s \in [0,T]$, $u \in \mathcal{U}$ the terms $\mathbb{1}_{[s,T]}(\cdot)\Psi(s)$ and $\mathbb{1}_{[s,T]}(\cdot)\Phi(s)u$ appearing in (1.2.5) denote the $D([0,T],H)$-valued random variables $\omega \mapsto \big(\mathbb{1}_{[s,T]}(r)\Psi(\omega,s)\big)_{r\in[0,T]}$ and $\omega \mapsto \big(\mathbb{1}_{[s,T]}(r)\Phi(\omega,s)u\big)_{r\in[0,T]}$, respectively. The functional Itô calculus introduced in [5, 11, 12, 20] has been generalized and modified in the SODE case in various directions, e.g., in [14, 40, 48].

It has been shown in [35] that functional Itô calculus can be employed to extend the classical Itô calculus approach to weak error analysis of numerical approximations of SODEs to a large class of sufficiently smooth path-dependent functionals. A generalization of functional Itô calculus to Itô processes with values in possibly infinite-dimensional Hilbert spaces has been presented in [52]. However, this work is restricted to standard Itô processes of the form (1.2.1) without the involvement of an operator semigroup or an evolution family, so that SPDEs are not covered. To the best of our knowledge, there exists no path-dependent Itô formula for solution processes of SPDEs and their numerical approximations so far. Consequently, the Itô calculus approach to weak error analysis has not yet been extended to path-dependent weak approximation errors for SPDEs. It is one of the contributions of this thesis to partially fill the gap.

1.3 Malliavin calculus and weak error analysis

Another useful toolbox for the analysis of weak errors of numerical approximations of SEEs is provided by Malliavin calculus, which is a stochastic calculus of variations that extends the classical calculus of variations from deterministic functions to stochastic processes [43, 47]. Malliavin calculus has been used in various different ways in the context of the analysis of weak approximation errors of type (1.1.3), often in combination with the approach based on Itô calculus and Kolmogorov equations; see, e.g., [3, 6–8, 17].

In this thesis we focus on a duality approach that has been introduced in [10, 39] and does not rely on Kolmogorov equations associated to the SEE. Instead, the basic ansatz is to start with a linearization of the weak error (1.1.3) in terms of a simple application of the mean value theorem,

$$\mathbb{E}\big[f(X(T)) - f(\tilde{X}(T))\big] = \int_0^1 \mathbb{E}\Big\langle f'\big(\theta X(T) + (1-\theta)\tilde{X}(T)\big),\, X(T) - \tilde{X}(T)\Big\rangle_H \,\mathrm{d}\theta. \qquad (1.3.1)$$

Here we assume again that $X = (X(t))_{t\in[0,T]}$ is the solution of an SEE (SODE or SPDE) of type (1.1.2), that $\tilde{X} = (\tilde{X}(t))_{t\in[0,T]}$ is a given approximation of X, and that $f\colon H \to \mathbb{R}$ is a sufficiently regular function. Moreover, for every $x \in H$ the Fréchet derivative $f'(x)$ of f at x, originally defined as a bounded linear functional $f'(x)\colon H \to \mathbb{R}$, is interpreted as an element $f'(x) \in H$ by means of the Riesz isomorphism. The crucial idea in the duality approach to weak error analysis is to exploit the Malliavin regularity of the H-valued random variable $f'\big(\theta X(T) + (1-\theta)\tilde{X}(T)\big)$ appearing in the integral on the right hand side of (1.3.1) and to formally consider the integrand term $\mathbb{E}\,\langle\cdots,\cdots\rangle_H$ on the right hand side of (1.3.1) as a duality bracket between a Malliavin-Sobolev space with positive smoothness parameter and the dual space thereof. The derivation of an upper bound for the weak approximation error (1.1.3) thus essentially boils down to the derivation of an upper bound of the norm of $X(T)-\tilde{X}(T)$ in a dual space of a Malliavin-Sobolev space, which is a weaker norm than the corresponding L^p-norm defining the strong approximation error (1.1.4). This abstract point of view is, however, not necessary for our purpose.

In order to indicate the duality approach to weak error analysis in some more detail, assume that the solution process X is given as a mild Itô process of the form (1.2.3) and let the approximation process $\tilde{X}$ be given by

$$\tilde{X}(t) = \tilde{S}(t,0)\tilde{\xi} + \int_0^t \tilde{S}(t,s)\tilde{\Psi}(s)\,\mathrm{d}s + \int_0^t \tilde{S}(t,s)\tilde{\Phi}(s)\,\mathrm{d}W(s), \quad t \in [0,T],$$

where $\tilde{S} = (\tilde{S}(t,s))_{0\le s\le t\le T} \subset L(H)$ is an approximation of the evolution family $S = (S(t,s))_{0\le s\le t\le T} \subset L(H)$ and where the integrand processes $\tilde{\Psi}$ and $\tilde{\Phi}$ take values in H and $L(U,H)$, respectively, and are assumed to satisfy suitable measurability and integrability conditions. Observe that the difference $X(T) - \tilde{X}(T)$ appearing in the integral on the right hand side of (1.3.1) can now be rewritten as

$$\begin{aligned} X(T) - \tilde{X}(T) = S(T,0)\xi - \tilde{S}(T,0)\tilde{\xi} + \int_0^T \Big[S(T,s)\Psi(s) - \tilde{S}(T,s)\tilde{\Psi}(s)\Big]\mathrm{d}s \\ + \int_0^T \Big[S(T,s)\Phi(s) - \tilde{S}(T,s)\tilde{\Phi}(s)\Big]\mathrm{d}W(s), \end{aligned}$$

and that the stochastic integral term therein represents the bottleneck when it comes to deriving sharp weak error estimates. Given that X and $\tilde{X}$ are sufficiently regular in a Malliavin sense, the Malliavin integration-by-parts formula in Lemma 4.2.3 in Section 4.2 below implies that the corresponding part of the weak error (1.3.1) can be simplified as

$$\begin{aligned}&\mathbb{E}\left[\left\langle F, \int_0^T \Big[S(T,s)\Phi(s) - \tilde{S}(T,s)\tilde{\Phi}(s)\Big]\mathrm{d}W(s)\right\rangle_H\right]\\ &\quad = \mathbb{E}\left[\int_0^T \Big\langle D_s F,\, S(T,s)\Phi(s) - \tilde{S}(T,s)\tilde{\Phi}(s)\Big\rangle_{\mathrm{HS}(U,H)}\,\mathrm{d}s\right]\end{aligned} \tag{1.3.2}$$

with $F = f'\big(\theta X(T) + (1-\theta)\tilde{X}(T)\big)$. Here we denote by $\mathrm{HS}(U,H)$ the space of Hilbert-Schmidt operators from U to H with inner product $\langle\cdot,\cdot\rangle_{\mathrm{HS}(U,H)}$ and, loosely speaking, the Malliavin derivative $DF = \big(D_t F\big)_{t\in[0,T]}$ of F constitutes an $\mathrm{HS}(U,H)$-valued stochastic process; see Section 2.2 and Section 4.2 below for details. The application of the integration-by-parts formula (1.3.2) often allows for a proper handling of the crucial stochastic integral terms involved in estimates of the weak approximation error (1.3.1).

The described duality approach to weak error analysis has been suggested in [10] for SODEs and independently in [39] for linear SPDEs with additive noise. Extension to semilinear SPDEs with additive noise can be found, e.g., in [1, 2, 4]. While the approach has originally been applied to weak approximation errors of the type (1.1.3) involving functionals that depend only on evaluations of the stochastic process at a fixed time point, the authors in [10] remark that path-dependent functionals can potentially be treated as well. In the SPDE case, a specific class of path-dependent functions of the form $f(x) = \prod_{i=1}^K \phi_i\big(\int_0^T x(t)\,\mu_i(\mathrm{d}t)\big)$ and the corresponding path-dependent weak error (1.1.5) have been analyzed in [1], with ϕ_i and μ_i being sufficiently smooth functions from H to $\mathbb{R}$ and finite Borel-measures on $[0,T]$, respectively. A similar class of specific path-dependent functionals has been considered in [4]. So far the duality approach to weak error analysis has not been extended to more general classes of path-dependent functionals. One of the contributions of this thesis is a new result in this context.

In the SODE case, Malliavin calculus methods have also been used to analyze weak errors of type (1.1.3) involving non-smooth functions $f\colon H \to \mathbb{R}$ such as indicator functions of certain subsets of H; see, e.g., [6, 7, 10]. The corresponding arguments are not directly applicable to the SPDE case, in which only sufficiently smooth functions f have been treated in the literature; a standard assumption is that f is at least twice continuously differentiable. Accordingly, non-smooth functions f do not lie within the scope of this thesis.

To complete the picture, let us add that a further approach to weak error analysis different from the approaches described in Section 1.2 and Section 1.3 can be found in [9]. Here the authors analyze weak approximation errors of type (1.1.5) for a relatively general class of path-dependent functionals and spectral Galerkin approximations of semilinear SPDEs. The approach makes use of the regularity of an underlying Itô map and heavily relies on the fact that only additive noise is considered.

1.4 Contributions and structure of the thesis

In this section we briefly outline the main results concerning the analysis of path-dependent functionals of SPDEs presented in this thesis.

In the context of the Itô calculus approach to weak error analysis reviewed in Section 1.2 above, we establish in Theorem 3.2.2 in Section 3.2 below a novel path-dependent mild Itô formula suitable for the analysis of path-dependent functionals of mild solutions of SPDEs and their numerical approximations. To this end, we combine the ideas behind the mild Itô formula (1.2.4) presented in [15] in an SPDE setting and the path-dependent Itô formulas of type (1.2.5) introduced in [5, 11, 12, 20] in an SODE framework. For the sake of a better compatibility with (1.2.4) and (1.2.5) we present here a slightly different version of our path-dependent mild Itô formula than the one formulated in Theorem 3.2.2; see Remark 3.2.5 for details. Let $X = (X(t))_{t\in[0,T]}$ be a mild Itô process of the form (1.2.3) with continuous H-valued sample paths, strongly continuous evolution family $S = (S(t,s))_{0\le s\le t\le T} \subset L(H)$, and integrand processes Ψ and Φ taking values in H and $\mathrm{HS}(U,H)$, respectively, satisfying the technical conditions summarized in Assumption 3.1.1 in Section 3.1 below. For all $s,t \in [0,T]$ with $s \le t$ let as before $X_t = X(\cdot \wedge t)$ be the $C([0,T],H)$-valued random variable given by the process stopped at time t and let further $S_{t,s}\colon D([0,T],H) \to D([0,T],H)$ be the bounded linear operator defined by

$$\begin{aligned}\big(S_{t,s}x\big)(r) &= \mathbb{1}_{[0,s)}(r)x(r) + \mathbb{1}_{[s,t)}(r)S(r,s)x(r) + \mathbb{1}_{[t,T]}(r)S(t,s)x(r)\\ &= \mathbb{1}_{[0,s)}(r)x(r) + \mathbb{1}_{[s,T]}(r)S(r\wedge t,s)x(r),\end{aligned}$$

$x \in D([0,T],H)$, $r \in [0,T]$. Moreover, suppose that the mapping $g\colon [0,T] \times D([0,T],H) \to \mathbb{R}$, $(t,x) \mapsto g(t,x)$ is continuously differentiable in t and two times continuously Fréchet differentiable in x and satisfies the regularity conditions formulated in Assumption 3.1.5 below with g and $\mathbb{R}$ in place of f and V. In this situation our path-dependent mild Itô formula ensures that for all $t \in [0,T]$ it holds almost surely that

$$\begin{aligned}g(t,X_t) &= g(0,S_{t,0}X_0) + \int_0^t \partial_1 g(s,S_{t,s}X_s)\,\mathrm{d}s\\ &\quad + \int_0^t \partial_2 g(s,S_{t,s}X_s)S_{t,s}\big(\mathbb{1}_{[s,T]}(\cdot)\Psi(s)\big)\,\mathrm{d}s\\ &\quad + \int_0^t \partial_2 g(s,S_{t,s}X_s)S_{t,s}\big(\mathbb{1}_{[s,T]}(\cdot)\Phi(s)\big)\,\mathrm{d}W(s)\\ &\quad + \frac{1}{2}\int_0^t \sum_{u\in\mathcal{U}} \partial_2^2 g(s,S_{t,s}X_s)\Big(S_{t,s}\big(\mathbb{1}_{[s,T]}(\cdot)\Phi(s)u\big), S_{t,s}\big(\mathbb{1}_{[s,T]}(\cdot)\Phi(s)u\big)\Big)\,\mathrm{d}s.\end{aligned} \tag{1.4.1}$$

Note that our path-dependent mild Itô formula extends the mild Itô formula (1.2.4) to path-dependent functionals f and extends the path-dependent Itô formula (1.2.5) to mild Itô processes

X.

The main idea of the proof of (1.4.1) is to employ appropriate discretizations of the transformed $C([0,T],H)$-valued process $(S_{t,s}X_s)_{s\in[0,t]}$ in combination with a Taylor expansion of g. These discretizations constitute a uniform approximation of $(S_{t,s}X_s)_{s\in[0,t]}$ over the time interval $[0,t]$. One of several technical difficulties involved in the proof lies in the well-known fact that the Banach spaces $C([0,T],H)$ and $D([0,T],H)$ are not suited as state spaces for the construction of Banach space-valued stochastic integrals, so that we cannot work, e.g., with an underlying $C([0,T],H)$-valued SEE for the $C([0,T],H)$-valued process $(S_{t,s}X_s)_{s\in[0,t]}$. As a first application of our functional mild Itô formula we investigate in Section 3.4.1 below the weak order of convergence of spectral Galerkin projections of SPDEs with multiplicative noise for the approximation of spatio-temporal covariances of the solution process. An example for possible further applications is outlined in Section 3.4.2.

In the context of the Malliavin calculus approach to weak error analysis reviewed in Section 1.3 above, we establish in Theorem 4.4.3 below a weak error estimate concerning path-dependent weak approximation errors of type (1.1.5) for the mild solutions X of SPDEs with multiplicative noise of type (1.1.1), spectral Galerkin projections $\tilde{X}$ of X, and a general class of path-dependent functionals $f\colon L^q(\mu;H)\to\mathbb{R}$ defined on the space $L^q(\mu;H)=L^q\big([0,T],\mathcal{B}([0,T]),\mu;H\big)$ of paths that are q-integrable w.r.t. a given finite Borel-measure μ on $[0,T]$ for a suitable integrability parameter $q\in(1,2]$. Here we assume that the linear operator $A\colon D(A)\subset H\to H$ in (1.1.1) is a diagonal operator with eigenbasis $(e_n)_{n\in\mathbb{N}}\subset H$ and associated sequence of eigenvalues $(\lambda_n)_{n\in\mathbb{N}}$ with $\sup_{n\in\mathbb{N}}\lambda_n<0$, that the mappings F and B in (1.1.1) are Lipschitz continuous from H to $H_{-\vartheta}$ and from H to the space of Hilbert-Schmidt operators $\mathrm{HS}(U,H_{-\vartheta/2})$, respectively for some $\vartheta\in[0,1)$, and that the initial condition ξ in (1.1.1) satisfies $\xi\in H_\rho$ for some $\rho\in[0,1-\vartheta)$. By H_r, $r\in\mathbb{R}$ we denote a family of interpolation spaces associated to $-A$ such that $H_r=D((-A)^r)$ for all $r\geq 0$; see Section 2.2 below for details. The spectral Galerkin projections $P_NX=(P_NX(t))_{t\in[0,T]}$, $N\in\mathbb{N}$, of X are defined by the orthogonal projection operators $P_N\colon H\to H$ given by $P_Nh=\sum_{n=1}^N\langle h,e_n\rangle_H\,e_n$, $h\in H$, and the path-dependent test function f is essentially assumed to be a twice continuously Fréchet differentiable mapping from $L^q(\mu;H)$ to $\mathbb{R}$ with bounded first and second Fréchet derivatives, where $q\in(1,2]$ satisfies $q<\frac{1}{\rho+\vartheta/2}$. In this situation, Theorem 4.4.3 below establishes that there exists a constant $M\in(0,\infty)$ such that

$$\big|\mathbb{E}\left[f(X)-f(P_NX)\right]\big|\leq M\left\|\mathrm{Id}-P_N\right\|_{L(H,H_{-\rho})}. \tag{1.4.2}$$

Note that the term $\|\mathrm{Id}-P_N\|_{L(H,H_{-\rho})}$ in the inequality above is finite. In particular, if the sequence $(\lambda_n)_{n\in\mathbb{N}}$ of eigenvalues of A is decreasing with $\lim_{n\to\infty}\lambda_n=-\infty$, we have

$$\|\mathrm{Id}-P_N\|_{L(H,H_{-\rho})}=(-\lambda_{N+1})^{-\rho}.$$

It turns out that the rate of weak convergence is as expected twice the rate of strong convergence; compare Remark 4.4.4 below.

Compared with the path-dependent weak error results obtained within the duality approach in [1, 4], the strategy of the proof of our weak error result in Theorem 4.4.3 allows to cover a more general class of path-dependent functionals, but we focus on Galerkin projections instead of implementable discretizations. The starting point of our proof is to replace the ansatz (1.3.1) by

$$\mathbb{E}\big[f(X) - f(\tilde{X})\big] = \int_0^1 \mathbb{E}\left\langle f'\big(\theta X + (1-\theta)\tilde{X}\big),\, X - \tilde{X}\right\rangle_{L^2(\mu;H)} \mathrm{d}\theta,$$

with $\tilde{X} = P_N X$. Here for every $x \in L^2(\mu;H)$ the Fréchet derivative $f'(x)\colon L^q(\mu;H) \to \mathbb{R}$, is interpreted as an element $f'(x) \in L^2(\mu;H)$ by means of the embedding $L\big(L^q(\mu;H),\mathbb{R}\big) \subset L\big(L^2(\mu;H),\mathbb{R}\big)$ and an application of the Riesz isomorphism. A challenging task is then to derive suitable Malliavin regularity results for X to be considered as a random variable in the path space. For this we extend the Malliavin regularity results in [22] in a way that fits our purpose. Compared with the weak error analysis in [9], Theorem 4.4.3 covers a less general class of functions but allows for multiplicative noise instead of only additive noise.

The structure of this thesis is as follows: In Chapter 2 we give a short review of mathematical definitions and preliminary results from the literature. This chapter is divided in the topics mathematical analysis, linear operators, Gâteaux and Fréchet derivatives, and stochastic analysis.

Chapter 3 contains our path-dependent mild Itô formula. This chapter begins with our setting and assumptions. In Section 3.2 we state our path-dependent mild Itô formula in Theorem 3.2.2. Then we discuss a comparison with related Itô-type formulas from the literature. In Section 3.3, we prove our path-dependent mild Itô formula. As an exemplary application, we show in Section 3.4.1 an upper bound for the weak error of approximations of spatio-temporal covariances of the solution process of a semilinear SPDE with multiplicative noise. In Section 3.4.2, we discuss a possible further application of our path-dependent mild Itô formula to linear SPDEs and the associated Kolmogorov equations.

In Chapter 4 we study the weak convergence rate for Galerkin projections of sample paths of mild solutions of nonlinear SPDEs with multiplicative Gaussian noise via Malliavin calculus. The approach applied in this chapter is inspired by [2] and [1]. Our setting in this chapter is mostly based on [22, Section 3]. After a short review of Malliavin calculus in Hilbert spaces in Section 4.2, we investigate the Malliavin regularity of mild solutions of SPDEs in Section 4.3. In Section 4.4 we state our weak error result in Theorem 4.4.3.

Chapter 2

Preliminaries

In this chapter we state the necessary notations, definitions and preparatory results that we use later in this thesis. We present the auxiliary results concerning mathematical analysis in Section 2.1 and concerning linear functions in Section 2.2 . In Section 2.3, we review the definitions of Gâteaux and Fréchet derivatives and we develop our setting regarding stochastic analysis in infinite dimensions in Section 2.4.

2.1 Some concepts from mathematical analysis

For a set A we denote its power set by $\mathcal{P}(A)$. The Borel σ-algebra associated to a topological space $(E, \mathcal{E})$ is denoted by $\mathcal{B}(E)$. If $n \in \mathbb{N}$ and $A \in \mathcal{B}(\mathbb{R}^n)$, we denote by $\lambda^n\big|_A : \mathcal{B}(A) \to [0, \infty]$ the Lebesgue-Borel measure on A. For the case $n = 1$ we simply use λ instead of λ^1. A measure μ on $(A, \mathcal{B}(A))$ is called finite Borel measure if $\mu(A) < \infty$.

For Banach spaces $(E, \|\cdot\|_E)$ and $(K, \|\cdot\|_K)$ we denote by $C(E, K)$ the space of continuous mappings from E to K. Recall that $C(E, K)$ equipped with the uniform-norm, i.e., $\|f\|_\infty = \sup_{x \in E} \|f(x)\|_K$, $f \in C(E, K)$, is a Banach space. We denote by $\mathrm{Lip}(E, K)$ the space of all continuous mappings $f : E \to K$ with

$$|f|_{\mathrm{Lip}(E,K)} := \sup_{x,y \in E,\, x \neq y} \frac{\|f(x) - f(y)\|_K}{\|x - y\|_E} < \infty.$$

We also use the norm $\|f\|_{\mathrm{Lip}(E,K)} = \|f(0)\|_K + |f|_{\mathrm{Lip}(E,K)}$.

To define the integral of operator-valued functions, we first need the definition of strong measurability. Roughly speaking, strongly measurable functions are those functions which can be pointwise approximated by simple functions. Recall that simple functions are functions with finite image set. For details see, e.g., Section 2.3 in [30].

Definition 2.1.1. We call a real Banach space $(E, \|\cdot\|_E)$ separable if there exists countable set $A \subset E$ such that the closure of A is equal to E.

Definition 2.1.2. Let $(E, \|\cdot\|_E)$ and $(K, \|\cdot\|_K)$ be real Banach spaces and let $(\Omega, \mathcal{A})$ be a measurable space. A function $f\colon \Omega \to E$ is called $\mathcal{A}$-strongly measurable if it is $\mathcal{A}$-$\mathcal{B}(E)$-measurable and $f(\Omega) \subset E$ is separable. Moreover, an operator-valued function $F\colon \Omega \to L(E, K)$ is called E-strongly measurable if for all $x \in E$, the K-valued mapping $F(\cdot)x\colon \Omega \to K$, $\omega \mapsto F(\omega)x$, is $\mathcal{A}$-strongly measurable.

The procedure of defining an integral for Banach space-valued functions, also called the Bochner integral, is similar to the real-valued case, however, the measurability is replaced by strong measurability. For details see, e.g., Appendix A in [42]. In the following we introduce the set of Bochner-integrable functions.

Definition 2.1.3. Let $(E, \|\cdot\|_E)$ be a real Banach space, let $(\Omega, \mathcal{A}, \mu)$ be a measure space and let $p \in (0, \infty)$. Then we denote by $\mathscr{L}^p\big(\Omega, \mathcal{A}, \mu; E\big)$ the space of $\mathcal{A}$-strongly measurable mappings $f\colon \Omega \to E$ with the property that

$$\int_\Omega \|f(x)\|_E^p \; \mu(\mathrm{d}x) < \infty.$$

Moreover, the space of equivalence classes of mappings $f \in \mathscr{L}^p\big(\Omega, \mathcal{A}, \mu; E\big)$, which coincide μ-almost everywhere is denoted by $L^p\big(\Omega, \mathcal{A}, \mu; E\big)$ or simply by $L^p\big(\mu; E\big)$ if it causes no confusion. Recall that for $p \in [1, \infty)$ the space $L^p\big(\mu; E\big)$ equipped with the norm

$$\|f\|_{L^p(\mu;E)} = \Big(\int_\Omega \|f(x)\|_E^p \; \mu(\mathrm{d}x) \Big)^{1/p}, \quad f \in L^p\big(\mu; E\big)$$

is a Banach space.

2.2 Linear functions in Banach spaces

Let $(E, \|\cdot\|_E)$ and $(K, \|\cdot\|_K)$ be real Banach spaces. We denote the space of linear operators from $U \subset E$ to K by $\mathrm{Lin}(U, K)$ and for the domain of a linear operator A we write $D(A)$. Moreover, we denote the space of bounded linear operators from E to K by $L(E, K)$. We write $L(E)$ for $L(E, E)$ and Id_E (or simply Id if no confusion arises) for the identity operator on E.

This section is mainly based on [30].

Definition 2.2.1. The mapping $B\colon E \times E \to K$ is called bilinear if for every $x \in E$ the mappings $B(x, \cdot)\colon E \to K$ and $B(\cdot, x)\colon E \to K$ are linear. Moreover, B is called bounded if

$$\|B\| = \sup_{x_1, x_2 \in E, \|x_1\|_E = 1 = \|x_2\|_E} \|B(x_1, x_2)\|_K < \infty,$$

so that

$$\|B(x_1, x_2)\|_K \leq \|B\| \cdot \|x_1\|_E \cdot \|x_2\|_E \quad \forall x_1, x_2 \in E.$$

The space of all bounded bilinear mappings from $E \times E$ to K is denoted by $L^{(2)}(E, K)$.

Definition 2.2.2. The space $E^* = L(E, \mathbb{R})$ of all real-valued bounded operators on E is called the dual space of E.

We usually identify, via Riesz-isomorphism, the dual space of a real Hilbert space $\left(U, \|\cdot\|_U, \langle\cdot,\cdot\rangle_U\right)$ with U itself.

Definition 2.2.3. Let $\left(U, \|\cdot\|_U, \langle\cdot,\cdot\rangle_U\right)$ be a real Hilbert space and $A\colon D(A) \subset U \to U$ be a linear operator. We call the operator A symmetric if for all $u, v \in D(A)$ it holds that

$$\langle Au, v\rangle_U = \langle u, Av\rangle_U .$$

Definition 2.2.4. Let $\left(U, \|\cdot\|_U, \langle\cdot,\cdot\rangle_U\right)$ and $\left(H, \|\cdot\|_H, \langle\cdot,\cdot\rangle_H\right)$ be real Hilbert spaces and $A \in L(U, H)$. The Hilbert-adjoint operator of A, denoted by $A^* \in L(H, U)$, is the bounded linear operator that satisfies

$$\langle Au, h\rangle_H = \langle u, A^* h\rangle_U \quad \forall u \in U, h \in H.$$

Definition 2.2.5. Let $\left(U, \|\cdot\|_U, \langle\cdot,\cdot\rangle_U\right)$ be a real Hilbert space and $A\colon D(A) \subset U \to U$ be an linear operator. We call the operator A non-negative if $\langle Au, u\rangle_U \geq 0$ for all $u \in D(A)$.

Definition 2.2.6. Let $\left(U, \|\cdot\|_U, \langle\cdot,\cdot\rangle_U\right)$ be a real Hilbert space and let $\mathcal{U}$ be an orthonormal basis of U. Then we define the trace of $A \in L(U)$ by $\operatorname{tr} A = \sum_{u \in \mathcal{U}} \langle Au, u\rangle_U$, given that the series is convergent and its value does not depend on the choice of the orthonormal basis.

The definition of trace is not necessarily independent of the choice of the orthonormal basis. In [51, Remark B.0.4], it is shown that the definition of trace for a special class of operators is well-defined and independent of the choice of the orthonormal basis. This class of operators is called the set of *nuclear operators*.

Definition 2.2.7. Let $\left(U, \|\cdot\|_U, \langle\cdot,\cdot\rangle_U\right)$ and $\left(H, \|\cdot\|_H, \langle\cdot,\cdot\rangle_H\right)$ be separable real Hilbert spaces. An operator $A \in L(U, H)$ is called nuclear if there exist sequences $(a_j)_{j\in\mathbb{N}} \subset H$ and $(b_j)_{j\in\mathbb{N}} \subset U$ such that

$$Ax = \sum_{j=1}^{\infty} a_j \langle b_j, x\rangle_U \quad \forall x \in U,$$

and

$$\sum_{j=1}^{\infty} \|a_j\|_H \cdot \|b_j\|_U < \infty.$$

Definition 2.2.8. Let $(U, \|\cdot\|_U, \langle\cdot,\cdot\rangle_U)$ and $(H, \|\cdot\|_H, \langle\cdot,\cdot\rangle_H)$ be separable real Hilbert spaces and $\mathcal{U}$ be an orthonormal basis of U. A bounded linear operator $A \in L(U, H)$ is called a Hilbert-Schmidt operator if $\sum_{u\in\mathcal{U}} \|Au\|_H^2 < \infty$. The space of Hilbert-Schmidt operators from U to H is denoted by

$$\mathrm{HS}(U, H) = \Big\{A \in L(U, H) \colon \sum_{u\in\mathcal{U}} \|Au\|_U^2 < \infty\Big\}.$$

The Hilbert-Schmidt norm on $\mathrm{HS}(U, H)$ is then defined by

$$\|A\|_{\mathrm{HS}(U,H)} = \Big(\sum_{u\in\mathcal{U}} \|Au\|_H^2\Big)^{1/2} \quad \forall A \in \mathrm{HS}(U, H).$$

Note that the Hilbert-Schmidt norm does not depend on the choice of the orthonormal basis; see, e.g., [51, Remark B.0.6]. Furthermore, it is well-known (see, e.g., [51, Proposition B.0.7]) that $\mathrm{HS}(U, H)$ equipped with the inner product $\langle L, T\rangle_{\mathrm{HS}(U,H)} = \big(\sum_{u\in\mathcal{U}} \langle Lu, Tu\rangle_U^2\big)^{1/2}$, with $L, T \in \mathrm{HS}(U, H)$, is a separable real Hilbert space.

Diagonal operators

Next we introduce a class of linear operators and associated interpolation spaces, which will be used later in our weak error rate results.

Definition 2.2.9. Let $(H, \|\cdot\|_H, \langle\cdot,\cdot\rangle_H)$ be a separable real Hilbert space. We call the linear operator $A \colon D(A) \subset H \to H$ a diagonal linear operator on H if there exist a set $\mathcal{I} \subset \mathbb{N}$, a family of real numbers $(\lambda_i)_{i\in\mathcal{I}}$, and an orthonormal basis $(e_i)_{i\in\mathcal{I}}$ of H such that

$$Av = \sum_{i\in\mathcal{I}} \lambda_i \langle e_i, v\rangle_H \, e_i,$$

for all $v \in D(A) = \big\{v \in H \colon \sum_{i\in\mathcal{I}} |\lambda_i|^2 |\langle e_i, v\rangle_H|^2 < \infty\big\}$.

Definition 2.2.10. Let $(H, \|\cdot\|_H, \langle\cdot,\cdot\rangle_H)$ be a separable real Hilbert space, let A be a diagonal operator on H, let $(\lambda_i)_{i\in\mathcal{I}}$ with $\inf_{i\in\mathcal{I}} \lambda_i > 0$ be the family of eigenvalues of A as in the definition above, and let $(H_r, \langle\cdot,\cdot\rangle_{H_r}, \|\cdot\|_{H_r})$, $r \in \mathbb{R}$, be real Hilbert spaces with the following properties

- $\forall r, s \in \mathbb{R}, r \geq s \colon H_r \subset H_s \subset \overline{H_r}^{H_s}$,
- $\forall r \in [0, \infty) \colon \big(D(A^r), \langle A^r(\cdot), A^r(\cdot)\rangle_H, \|A^r(\cdot)\|_H\big) = \big(H_r, \langle\cdot,\cdot\rangle_{H_r}, \|\cdot\|_{H_r}\big)$, and
- $\forall r \in (-\infty, 0], v \in H \colon \|v\|_{H_r} = \|A^r v\|_H$.

Then we call $(H_r, \langle\cdot,\cdot\rangle_{H_r}, \|\cdot\|_{H_r})$, $r \in \mathbb{R}$, a family of interpolation spaces associated to A.

Note that the functional powers $A^r\colon D(A^r) \subset H \to H$, $r \in [0,\infty)$, of the operator A appearing in Definition 2.2.10 above are defined in the usual way; see, e.g., [30, Section 3.6.3]. The existence of interpolation spaces defined in Definition 2.2.10 above is shown in, e.g., [30, Theorem 3.6.29].

Semigroups of bounded linear operators

Definition 2.2.11. Let $S\colon [0,\infty) \to L(E)$ be a mapping with the property that for all $t_1, t_2 \in [0,\infty)$ it holds that

$$S(0) = \mathrm{Id}_E, \quad \text{and} \quad S(t_1)\, S(t_2) = S(t_1 + t_2).$$

Then we call S a semigroup of bounded linear operators on E.

There are different types of semigroups of bounded linear operators. For this work, however, we need only the following type of semigroups.

Definition 2.2.12. A semigroup S of bounded linear operators on E is called a strongly continuous semigroup or C_0-semigroup if for all $v \in E$ it holds that $\lim_{t\to 0} S(t)v = v$.

Definition 2.2.13. Let S be a strongly continuous semigroup on E. A linear operator $A\colon D(A) \subset E \to E$ is called the (infinitesimal) generator of S if

$$D(A) = \left\{ v \in E\colon \lim_{t \searrow 0} \frac{S(t)v - v}{t} \text{ exists} \right\},$$

and it holds for all $v \in D(A)$ that

$$Av = \lim_{t \searrow 0} \frac{S(t)v - v}{t}.$$

The proof of the lemma below can be found in, e.g., [30, Theorem 4.8.2] and [30, Theorem 4.8.5].

Lemma 2.2.14. Let $\left(H, \|\cdot\|_H, \langle\cdot,\cdot\rangle_H\right)$ be a separable real Hilbert space, let $l \subset \mathbb{N}$, let $(e_i)_{i\in\mathcal{I}} \subset H$ be an orthonormal basis of H, let A be a diagonal operator on H, and let $(\lambda_i)_{i\in\mathcal{I}}$ be a family of numbers with $\sup_{i\in\mathcal{I}} \lambda_i < 0$ and with the property that

$$D(A) = \left\{ v \in H\colon \sum_{i\in\mathcal{I}} |\lambda_i|^2 |\langle e_i, v\rangle_H|^2 < \infty \right\}$$

and such that for all $v \in D(A)$ it holds that $Av = \sum_{i\in\mathcal{I}} \lambda_i \langle e_i, v\rangle_H e_i$. Furthermore, let $\left(H_r, \|\cdot\|_{H_r}, \langle\cdot,\cdot\rangle_{H_r}\right)$, $r \in \mathbb{R}$, be a family of interpolation spaces associated to $-A$. Then

(i) it holds for all $r \in [0, \infty)$ that

$$\sup_{t \in (0,\infty)} \left\| \left(-tA \right)^r e^{tA} \right\|_{L(H)} \leq \left[\frac{r}{e} \right]^r < \infty,$$

(ii) it holds that A is the generator of the strongly continuous semigroup $(e^{tA})_{t \geq 0}$ on H and it holds for all $v \in \cup_{r \in \mathbb{R}} H_r$ and $t \in (0, \infty)$ that

$$e^{tA} v = \sum_{i \in \mathcal{I}} e^{t\lambda_i} \left\langle e_i, v \right\rangle_H e_i.$$

2.3 Differentiation of nonlinear functions on Banach spaces

Let $(E, \|\cdot\|_E)$ and $(K, \|\cdot\|_K)$ be real Banach spaces. In this section, we review, mainly based on [57], two types of derivatives defined for functions on Banach spaces.

Definition 2.3.1. We call a mapping $f \colon E \to K$ Gâteaux differentiable on E if for each $x \in E$ there exists an operator $\mathcal{T}_x \in L(E, K)$ such that

$$\lim_{h \to 0} \frac{f(x + hv) - f(x)}{h} = \mathcal{T}_x v \qquad \forall v \in E. \tag{2.3.1}$$

We write $f \in \mathscr{G}^1(E, K)$ if f is continuous, Gâteaux differentiable and the mapping $\mathcal{T}_{\cdot} \colon E \to L(E, K)$, $x \mapsto \mathcal{T}_x$, is strongly continuous, i.e., for every $v \in E$ the mapping $\mathcal{T}_{\cdot} v \colon E \to K$, $x \mapsto \mathcal{T}_x v$, is continuous.

Definition 2.3.2. We call a mapping $f \colon E \to K$ Fréchet differentiable at point $x \in E$ if the convergence in (2.3.1) is uniformly on $\{v \in E \colon \|v\|_E \leq 1\}$. We call the mapping $f \colon E \to K$ Fréchet differentiable on E if it is Fréchet differentiable for all $x \in E$. In this case we denote its Fréchet derivative by $f' \colon E \to L(E, K)$. If f' is Fréchet differentiable as well, then for every $x \in E$ the operator $(f')'(x) \colon E \to L(E, K)$ is a bounded linear operator, i.e., $(f')'(x) \in L(E, L(E, K))$. Therefore, for all $v_1, v_2 \in E$ it holds that $(f')'(x)(v_1) \in L(E, K)$ and that

$$\left\| (f')'(x)(v_1)(v_2) \right\|_K \leq \left\| (f')'(x)(v_1) \right\|_{L(E,K)} \cdot \|v_2\|_E \leq \left\| (f')'(x) \right\|_{L(E, L(E,K))} \cdot \|v_1\|_E \cdot \|v_2\|_E .$$

Hence the mapping $(v_1, v_2) \mapsto (f')'(x)(v_1)(v_2)$ is a bounded bilinear mapping from $E \times E$ to K, i.e., $(f')' \in L^{(2)}(E, K)$. We use the notation f'' instead of $(f')'$ for the second Fréchet derivative of f and the identification

$$(f')'(x)(v_1)(v_2) = f''(x)(v_1, v_2) \quad \forall x, v_1, v_2 \in E.$$

For more details see [58, Section 4.6]. By $C^2(E, K)$ we denote the set of all mappings $f \colon E \to K$, which are 2-times continuously Fréchet differentiable.

Remark 2.3.3. If $f \in C^2(E, K)$, then its second Fréchet derivative f'' is a symmetric bounded bilinear operator and by symmetry we mean that

$$f''(x)(v_1, v_2) = f''(x)(v_2, v_1) \quad \forall x, v_1, v_2 \in E.$$

A proof can be found in [58, Problem 4.10].

The following lemma states a second-order Taylor formula for Fréchet differentiable functions on Banach spaces. A proof can be found in, e.g., [58, Theorem 4.C].

Lemma 2.3.4. Let $f \in C^2(E, K)$. Then it holds for all $x, v \in E$ that

$$f(x+v) = f(x) + f'(x)(v) + \frac{1}{2} f''(x)(v,v) + \int_0^1 (1-\theta)\big[f''(x+\theta v) - f''(x)\big](v,v)\,\mathrm{d}\theta.$$

2.4 Stochastic analysis in infinite dimensions

If $T \in (0,\infty)$ and $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\big)$ is a filtered probability space, we call the σ-field

$$\begin{aligned}\mathcal{P}_T =& \sigma\Big(\big\{F_s \times (s,t] \colon 0 \le s < t \le T, F_s \in \mathcal{F}_s\big\} \cup \big\{F_0 \times \{0\} \colon F_0 \in \mathcal{F}_0\big\}\Big)\\ =& \sigma\Big(Y \colon \Omega \times [0,T] \to \mathbb{R} \colon Y \text{ is left-continuous and adapted to } (\mathcal{F}_t)_{t\in[0,T]}\Big).\end{aligned}$$

the predictable σ-field on $\Omega \times [0,T]$ and its elements are called predictable sets. A $\mathcal{P}_T$-measurable process will be called predictable.

2.4.1 Conditional expectation in Banach spaces

In [51, Section 2.2], it is shown that, analogously to real-valued random variables, the conditional expectation for Banach space-valued random variables are well-defined as well because the real-valued conditional expectation operator is a positive operator.

Lemma 2.4.1. Let $(E, \|\cdot\|_E)$ be a separable real Banach space, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\mathcal{G} \subset \mathcal{F}$ be a σ-algebra and let X be a Bochner integrable E-valued random variable. Then there exists a unique, up to a $\mathbb{P}$-null set, Bochner integrable E-valued random variable Z, measurable with respect to $\mathcal{G}$, such that

$$\int_A X \,\mathrm{d}\mathbb{P} = \int_A Z \,\mathrm{d}\mathbb{P} \quad \forall A \in \mathcal{G}.$$

The random variable Z is denoted by $\mathbb{E}\big[X\big|\mathcal{G}\big]$ and is called the conditional expectation of X given $\mathcal{G}$. Furthermore, it holds $\mathbb{P}$-almost surely that

$$\big\|\mathbb{E}\big[X\big|\mathcal{G}\big]\big\|_E \leq \mathbb{E}\big[\, \|X\|_E \,\big|\mathcal{G}\big].$$

Proof. See [51, Proposition 2.2.1]. □

The following result is taken from Proposition 2.2.2 in [51].

Lemma 2.4.2. Let $(M_1, \mathcal{A}_1)$ and $(M_2, \mathcal{A}_2)$ be two measurable spaces, let $\Psi\colon M_1 \times M_2 \to \mathbb{R}$ be a bounded measurable mapping, let X_1 and X_2 be two random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $(M_1, \mathcal{A}_1)$ and $(M_2, \mathcal{A}_2)$, respectively and let $\mathcal{G} \subset \mathcal{F}$ be an σ-algebra. Assume that X_1 is $\mathcal{G}$-measurable and X_2 is independent of $\mathcal{G}$, then it holds that

$$\mathbb{E}\big[\Psi(X_1, X_2)\big|\mathcal{G}\big] = \hat{\Psi}(X_1),$$

where

$$\hat{\Psi}(x) = \mathbb{E}\left[\Psi(x, X_2)\right] \quad \forall x \in M_1.$$

2.4.2 Wiener processes and stochastic integration in Hilbert spaces

To develop our stochastic calculus, we review some definitions mainly from [51]. Our aim is to introduce the stochastic integral with respect to a cylindrical Wiener process and then introduce mild solutions of stochastic evolution equations. To do this, we first need to state the definition of the stochastic integral with respect to a standard Q-Wiener process.

In this section we assume that $T \in (0, \infty)$, $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\big)$ is a filtered probability space, and $(U, \langle\cdot,\cdot\rangle_U, \|\cdot\|_U)$ and $(H, \langle\cdot,\cdot\rangle_H, \|\cdot\|_H)$ are infinite-dimensional separable real Hilbert spaces. The finite-dimensional cases are then particular implications of the infinite-dimensional setting.

Standard Q-Wiener process

Definition 2.4.3. Let $Q\colon U \to U$ be a non-negative and symmetric linear operator with finite trace and let $(u_k)_{k\in\mathbb{N}}$ be an orthonormal basis of U consisting of eigenvectors of Q corresponding to real eigenvalues $(\lambda_k)_{k\in\mathbb{N}}$. We call the U-valued $(\mathcal{F}_t)$-adapted stochastic process $(W(t))_{t\in[0,T]}$ a standard Q-Wiener process if:

- $W(0) = 0$,
- for all $t_1, t_2 \in [0, T]$ with $t_1 \leq t_2$, $\sigma\big(W(t_2) - W(t_1)\big)$ and $\mathcal{F}_{t_1}$ are independent,

- for all $t_1, t_2 \in [0,T]$ with $t_1 \leq t_2$ the random variable $W(t_2)-W(t_1)$ has Gaussian distribution with mean 0 and covariance operator $(t_2-t_1)Q$, i.e., $W(t_2)-W(t_1) \sim \mathcal{N}(0,(t_2-t_1)Q)$.

A representation of a standard Q-Wiener process is given in [51, Proposition 2.1.10].

Lemma 2.4.4. Let $Q\colon U \to U$ be a non-negative and symmetric linear operator with finite trace and let $(u_k)_{k\in\mathbb{N}}$ be an orthonormal basis of U consisting of eigenvectors of Q corresponding to real eigenvalues $(\lambda_k)_{k\in\mathbb{N}}$. Then a U-valued stochastic process $(W(t))_{t\in[0,T]}$ is a standard Q-Wiener process on U if and only if

$$W(t) = \sum_{k\in\mathbb{N}} \sqrt{\lambda_k}\,\beta_k(t)u_k, \tag{2.4.1}$$

where $(\beta_k(t))_{t\in[0,T]}$, $k = 1,2,\ldots$, are independent real-valued standard Brownian motions. The series in (2.4.1) converges in $L^2(\mathbb{P}; C([0,T],U))$ with the uniform norm on $C([0,T],U)$.

Stochastic integration with respect to a standard Q-Wiener Process

Note that for every non-negative symmetric $L \in L(U)$, there exists a unique bounded linear operator $L^{1/2}\colon U \to U$ with $L = L^{1/2} \circ L^{1/2}$. For more details see, e.g., [51, Proposition 2.3.4].

A predictable $\mathrm{HS}(Q^{1/2}(U),H)$-valued stochastic process $\big(\Phi(t)\big)_{t\in[0,T]}$ is integrable with respect to a standard Q-Wiener process $(W(t))_{t\in[0,T]}$ if $\mathbb{P}\big(\int_0^T \|\Phi(s)\|^2_{\mathrm{HS}(Q^{1/2}(U),H)}\,\mathrm{d}s < \infty\big) = 1$. Note that the inner product of the subspace $U_0 := Q^{1/2}(U)$ is given by $\langle u_0, v_0\rangle_{U_0} = \big\langle Q^{-1/2}u_0, Q^{-1/2}v_0\big\rangle_U$, $u_0, v_0 \in U_0$, where $Q^{-1/2}$ is the pseudo inverse of $Q^{1/2}$ in the case that Q is not one to one. According to [21, Proposition C.3], $(U_0, \langle\cdot,\cdot\rangle_{U_0})$ is a separable real Hilbert space. The construction of the stochastic integral is explained in, e.g., [51, Section 2.3.2].

The following result is the aforementioned Itô isometry and is taken from [30, Theorem 6.3.29].

Lemma 2.4.5. Let $Q\colon U \to U$ be a non-negative and symmetric linear operator with finite trace, let $(W(t))_{t\in[0,T]}$ be a standard Q-Wiener process defined on the filtered probability space $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\big)$ and let $\Phi\colon \Omega\times[0,T] \to \mathrm{HS}(Q^{1/2}(U),H)$ be predictable with $\mathbb{E}\int_0^T \|\Phi(s)\|^2_{\mathrm{HS}(Q^{1/2}(U),H)}\,\mathrm{d}s < \infty$. Then it holds for all $t \in [0,T]$ that

$$\mathbb{E}\left[\left\|\int_0^t \Phi(s)\,\mathrm{d}W(s)\right\|_H^2\right] = \mathbb{E}\left[\int_0^t \|\Phi(s)\|^2_{\mathrm{HS}(Q^{1/2}(U),H)}\,\mathrm{d}s\right],$$

and $\mathbb{E}\left[\int_0^t \Phi(s)\,\mathrm{d}W(s)\right] = 0$.

Cylindrical Wiener process

In many situations one is interested in cases where Q does not have finite trace, for example the identity operator. In such cases, the series in (2.4.1) is not convergent in $L^2(\mathbb{P}; C([0,T],U))$ any more. This problem can be solved by introducing the concept of cylindrical Wiener processes.

Definition 2.4.6. Let $Q\colon U \to U$ be a non-negative and symmetric linear operator, let $U_0 = Q^{1/2}U$ and let $(e_k)_{k\in\mathbb{N}}$ be an orthonormal basis of U_0. Suppose that there exists a further Hilbert space $(U_1, \langle\cdot,\cdot\rangle_{U_1}, \|\cdot\|_{U_1})$ and a Hilbert-Schmidt embedding $J\colon U_0 \to U_1$. The operator $Q_1 = JJ^*$ is then non-negative and symmetric with finite trace. Then by a cylindrical Q-Wiener process we mean actually the standard Q_1-Wiener process on U_1, and it can be represented (see, e.g., [51, Proposition 2.5.2]) by

$$W(t) = \sum_{k\in\mathbb{N}} \beta_k(t)\, Je_k, \tag{2.4.2}$$

with convergence in $L^2(\mathbb{P}; C([0,T],U_1))$. Note that $(W(t))_{t\in[0,T]}$ takes values in the larger space U_1 instead of U.

Moreover, note that $(U_1, \langle\cdot,\cdot\rangle_{U_1}, \|\cdot\|_{U_1})$ and J in the Definition 2.4.6 above always exist. For instance, choose $U_1 = U$ and $(\alpha_k)_{k\in\mathbb{N}}$ such that $\sum_{k\in\mathbb{N}} \alpha_k^2 < \infty$ and define $J(u) = \sum_{k\in\mathbb{N}} \alpha_k \langle u, e_k\rangle_{U_0} e_k$, for all $u \in U_0$; see [51, Remark 2.5.1].

Stochastic integration with respect to a cylindrical Q-Wiener Process

Assume the setting in Definition 2.4.6. We say a predictable $\mathrm{HS}(Q_1^{1/2}(U_1), H)$-valued stochastic process $\big(\Phi(t)\big)_{t\in[0,T]}$ is integrable with respect to the cylindrical Q-Wiener process $(W(t))_{t\in[0,T]}$ if $\mathbb{P}\big(\int_0^T \|\Phi(s)\|^2_{\mathrm{HS}(Q_1^{1/2}(U_1),H)}\,\mathrm{d}s < \infty\big) = 1$, and we basically integrate with respect to the standard Q_1-Wiener process defined on U_1.

It is important to know that the definition of the stochastic integral $\int_0^\cdot \Phi(s)\mathrm{d}W(s)$ is independent of the choice of U_1 and J, see [51, Remark 2.5.3].

For this thesis, the cylindrical Id_U-Wiener process is of most interest. In this case, $U_0 = \mathrm{Id}_U^{1/2}(U) = U$ and the representation in (2.4.2) reads

$$W(t) = \sum_{k\in\mathbb{N}} \beta_k(t)\, Ju_k, \tag{2.4.3}$$

with $(u_k)_{k\in\mathbb{N}}$ being an orthonormal basis of U and the series converging in $L^2(\mathbb{P}; C([0,T],U_1))$. The stochastic integral $\int_0^\cdot \Phi(s)\mathrm{d}W(s)$ is then well-defined for all $\Phi\colon \Omega\times[0,T] \to \mathrm{HS}(U,H)$ such that Φ is predictable and $\mathbb{P}\big(\int_0^T \|\Phi(s)\|^2_{\mathrm{HS}(U,H)}\,\mathrm{d}s < \infty\big) = 1$. Indeed, by [16, Proposition 4.7] it holds that $Q_1^{1/2}(U_1) = U$.

The representation in (2.4.3) and the definition in [51, Equation (2.5.2)] ensure that there exists a family of independent real-valued Brownian motions $(\beta(t))_{t\in[0,T]}$, $k \in \mathbb{N}$ such that for all predictable $\big(\Phi(t)\big)_{t\in[0,T]}$ with $\mathbb{P}\big(\int_0^T \|\Phi(s)\|^2_{\mathrm{HS}(U,H)}\,\mathrm{d}s < \infty\big) = 1$ and all $t \in [0,T]$ it holds that

$$\int_0^t \Phi(s)\mathrm{d}W(s) = \sum_{k\in\mathbb{N}} \int_0^t \Phi(s)u_k\,\mathrm{d}\beta_k(s), \tag{2.4.4}$$

where the series above converges in $L^2(\mathbb{P}; C([0,T],H))$. If further $\mathbb{E}[\int_0^T \|\Phi(s)\|^2_{\mathrm{HS}(U,H)}\,\mathrm{d}s] < \infty$ then, analogously to Lemma 2.4.5, it holds that

$$\mathbb{E}\left[\left\|\int_0^t \Phi(s)\,\mathrm{d}W(s)\right\|_H^2\right] = \mathbb{E}\left[\int_0^t \|\Phi(s)\|^2_{\mathrm{HS}(U,H)}\,\mathrm{d}s\right] \quad \forall t \in [0,T],$$

and

$$\mathbb{E}\left[\int_0^t \Phi(s)\,\mathrm{d}W(s)\right] = 0 \quad \forall t \in [0,T].$$

2.4.3 A class of stochastic evolution equations

Assumption 2.4.7. Let $T \in (0,\infty)$, let $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\big)$ be a filtered probability space, let the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ satisfy the usual conditions, and let $\big(U, \|\cdot\|_U, \langle\cdot,\cdot\rangle_U\big)$ and $\big(H, \|\cdot\|_H, \langle\cdot,\cdot\rangle_H\big)$ be separable real Hilbert spaces. Furthermore, let the stochastic process $W = (W(t))_{t\in[0,T]}$ be an $(\mathcal{F}_t)$-adapted cylindrical Id_U-Wiener process, let $A\colon D(A) \subset H \to H$ be the generator of strongly continuous semigroup $(e^{tA})_{t\geq 0} \subset L(H)$ and let $F\colon H \to H$ and $B\colon H \to L(U,H)$ be measurable and strongly measurable mappings, respectively.

This section is based on Chapter 2 in [21] and Appendix G in [42] and we introduce different types of solutions of the following stochastic evolution equation

$$\begin{cases} \mathrm{d}X(t) = \big[AX(t) + F(X(t))\big]\mathrm{d}t + B(X(t))\,\mathrm{d}W(t), & t \in [0,T] \\ X(0) = \xi \in L^2\big(\Omega, \mathcal{F}_0, \mathbb{P}; H\big). \end{cases} \tag{2.4.5}$$

Definition 2.4.8. Let Assumption 2.4.7 be fulfilled. We call an H-valued predictable stochastic process $X = (X(t))_{t\in[0,T]}$

(i) an (analytically) strong solution of (2.4.5) if

(a) it holds $\mathbb{P}(\mathrm{d}\omega) \otimes \mathrm{d}t$-almost everywhere that $X(\omega,t) \in D(A)$,

(b) $(AX(t))_{t\in[0,T]}$ and $(F(X(t)))_{t\in[0,T]}$ are $\mathbb{P}$-a.s. Bochner-integrable and $(B(X(t)))_{t\in[0,T]}$ is integrable with respect to the cylindrical Id_U-Wiener process $(W(t))_{t\in[0,T]}$,

(c) for every $t \in [0,T]$, it holds $\mathbb{P}$-a.s. that

$$X(t) = \xi + \int_0^t \big[AX(s) + F(X(s))\big] \mathrm{d}s + \int_0^t B(X(s)) \, \mathrm{d}W(s),$$

(ii) an (analytically) weak solution of (2.4.5) if for every $h \in D(A^*)$ and $t \in [0,T]$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned}\langle X(t), h\rangle_H = \langle \xi, h\rangle_H + \int_0^t \big[\, \langle X(s), A^*h\rangle_H + \langle F(X(s)), h\rangle_H \,\big] \mathrm{d}s \\ + \int_0^t \langle h, B(X(s)) \, \mathrm{d}W(s)\rangle_H \, .\end{aligned} \tag{2.4.6}$$

In particular, the integrals in (2.4.6) above have to be well-defined.

(iii) a mild solution of (2.4.5) if for all $t \in [0,T]$ it holds $\mathbb{P}$-a.s. that

$$X(t) = e^{tA}\xi + \int_0^t e^{(t-s)A} F(X(s)) \, \mathrm{d}s + \int_0^t e^{(t-s)A} B(X(s)) \, \mathrm{d}W(s). \tag{2.4.7}$$

In particular, the integrals in (2.4.7) above have to be well-defined.

It is remarkable that every strong solution of (2.4.5) is also a weak solution. Moreover, if $(X(t))_{t\in[0,T]}$ is a weak solution of (2.4.5), $B(X(t))_{t\in[0,T]}$ takes values in $\mathrm{HS}(U,H)$ and

$$\mathbb{P}\Big(\int_0^T \|X(s)\|_H + \|F(X(s))\|_H + \|B(X(s))\|^2_{\mathrm{HS}(U,H)} \, \mathrm{d}s < \infty \Big) = 1,$$

then the process is also a mild solution. For details see, e.g., [42, Proposition G.0.5].

Definition 2.4.9. Let T, $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\big)$, U, H, and $W = (W(t))_{t\in[0,T]}$ be taken from Assumption 2.4.7. Additionally, let $S\colon \{(\tau,\sigma) \in [0,T]^2 \colon \sigma \le \tau\} \to L(H)$ be a strongly measurable mapping satisfying $S(t,s)S(s,r) = S(t,r)$ for all $0 \le r \le s \le t \le T$ and let $\Psi\colon [0,T] \times \Omega \to H$ and strongly measurable $\Phi\colon [0,T] \times \Omega \to L(U,H)$ be two predictable stochastic processes with

$$\forall t \in [0,T]\colon \quad \int_0^t \Big(\|S(t,s)\Psi(s)\|_H + \|S(t,s)\Phi(s)\|^2_{\mathrm{HS}(U,H)} \Big) \, \mathrm{d}s < \infty \quad \mathbb{P}\text{-a.s.}$$

Then a predictable stochastic process $X\colon [0,T] \times \Omega \to H$ satisfying

$$\forall t \in [0,T]\colon \quad X(t) = S(t,0)X(0) + \int_0^t S(t,s)\Psi(s) \mathrm{d}s + \int_0^t S(t,s)\Phi(s) \mathrm{d}W(s) \quad \mathbb{P}\text{-a.s.},$$

is called a mild Itô process.

Chapter 3

On path-dependent Itô calculus in infinite-dimensional Hilbert spaces

This chapter contains our path-dependent mild Itô formula and begins with our setting and assumptions. In Section 3.2 we state our path-dependent mild Itô formula in Theorem 3.2.2. Then we discuss a comparison with related Itô-type formulas from the literature. In Section 3.3, we prove our path-dependent mild Itô formula. As an exemplary application, we show in Section 3.4.1 an upper bound for the weak error of approximations of spatio-temporal covariances of the solution process of a semilinear SPDE with multiplicative noise. In Section 3.4.2, we discuss a possible further application of our path-dependent mild Itô formula to linear SPDEs and the associated Kolmogorov equations. In Section 3.5 we collect some technical results mostly used in the proof of Theorem 3.2.2 below.

3.1 Preliminaries

For $T \in (0, \infty)$ and a separable real Hilbert space $\big(H, \|\cdot\|_H, \langle \cdot, \cdot \rangle_H\big)$, we denote by $D([0,T], H)$ the space of H-valued right-continuous with left limits (càdlàg) paths endowed with the uniform-norm $\|\cdot\|_{D([0,T],H)}$ given by $\|x\|_{D([0,T],H)} = \sup_{s\in[0,T]} \|x(s)\|_H$, $x \in D([0,T], H)$. If further $(E, \|\cdot\|_E)$ is a real Banach space, we denote by $C\big([0,T]\times D([0,T], H), E\big)$ the set of all continuous functions on $[0,T]\times D([0,T], H)$ with values in E. Given a function $f\colon [0,T]\times D([0,T], H) \to E$, then the mapping $\partial_1^+ f\colon [0,T) \times D([0,T], H) \to E$ defined by

$$\partial_1^+ f(t_0, x_0) = \lim_{h\to 0^+} \frac{f(t_0 + h, x_0) - f(t_0, x_0)}{h}, \qquad (t_0, x_0) \in [0,T] \times D([0,T], H),$$

is the right-sided derivative with respect to time parameter t, if the above limit exists. By $\partial_2 f(t_0, x_0)$ and $\partial_2^2 f(t_0, x_0)$ we denote, if they exist, the first and second Fréchet derivatives of the mapping $f(t_0, \cdot)\colon D([0,T],H) \to E$ at x_0. Note that $\partial_2 f(t_0,x_0) \in L\big(D([0,T],H),E\big)$ and $\partial_2^2 f(t_0,x_0) \in L\big(L(D([0,T],H),E),E\big)$. If the function f depends only on the variable x then we denote its first and second Fréchet derivatives with the conventional notation f' and f'', respectively.

Setting and assumptions

In the following we introduce the assumptions employed throughout this chapter.

Consider the stochastic process $X = (X(t))_{t\in[0,T]}$ given by

$$X(t) = S(t,0)X(0) + \int_0^t S(t,s)\Psi(s)\mathrm{d}s + \int_0^t S(t,s)\Phi(s)\mathrm{d}W(s), \quad t \in [0,T], \tag{3.1.1}$$

where we assume the following:

Assumption 3.1.1 (Process X)**.**

(i) Let $T \in (0,\infty)$ and let $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\big)$ be a filtered probability space such that the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ satisfies the usual conditions.

(ii) Let U and H be separable real Hilbert spaces and let $W = (W(t))_{t\in[0,T]}$ be an $(\mathcal{F}_t)$-adapted cylindrical Id_U-Wiener process.

(iii) Let $S = (S(t,s))_{0\le s\le t\le T}$ be a strongly continuous evolution family on H, i.e., for all $0 \le r \le s \le t \le T$ it holds that

$$S(t,s) \in L(H), \quad S(t,t) = \mathrm{Id}_H, \quad S(t,r) = S(t,s)S(s,r),$$

and for all $h \in H$ the mapping $(\tau,\sigma) \mapsto S(\tau,\sigma)h$ is continuous from $\{(\tau,\sigma) \in [0,T]^2 \colon \sigma \le \tau\}$ to H.

(iv) Let the stochastic processes $\Psi\colon [0,T]\times\Omega \to H$ and $\Phi\colon [0,T]\times\Omega \to \mathrm{HS}(U,H)$ be predictable mappings such that the following assumption is fulfilled:

$$\int_0^T \Big(\|\Psi(s)\|_H + \|\Phi(s)\|^2_{\mathrm{HS}(U,H)}\Big)\,\mathrm{d}s < \infty \qquad \mathbb{P}\text{-a.s.}$$

(v) Let $X(0)\colon \Omega \to H$ be $\mathcal{F}_0$-measurable.

(vi) Assume that the process X defined by (3.1.1) admits a modification with continuous sample paths $[0,T] \ni t \mapsto X(\omega,t) \in H$, $\omega \in \Omega$.

Note that, as a consequence of the uniform boundedness principle, we have that

$$\sup_{0\le s\le t\le T} \|S(t,s)\|_{L(H)} < \infty. \tag{3.1.2}$$

Therefore, the integrals in (3.1.1) are well-defined. Indeed, for all $t \in [0,T]$ we have that

$$\int_0^t \Big(\|S(t,s)\Psi(s)\|_H + \|S(t,s)\Phi(s)\|^2_{\mathrm{HS}(U,H)} \Big) \,\mathrm{d}s < \infty \qquad \mathbb{P}\text{-a.s.}$$

Moreover, note that, due to items (iii), (iv), and (vi) above, the H-valued processes $\big(\int_0^t S(t,s)\Psi(s)\mathrm{d}s\big)_{t\in[0,T]}$ and $\big(\int_0^t S(t,s)\Phi(s)\mathrm{d}W(s)\big)_{t\in[0,T]}$ admit continuous modifications. In the sequel, we will always consider these specific modifications without explicit mentioning it.

Example 3.1.2 (Mild solutions of SEEs)**.** Let Assumption 2.4.7 be fulfilled and assume further that $F\colon H \to H$ and $B\colon H \to \mathrm{HS}(U,H)$ are globally Lipschitz continuous. Consider the following stochastic evolution equation

$$\begin{cases} \mathrm{d}X(t) = \big[AX(t) + F\big(X(t)\big)\big]\mathrm{d}t + B\big(X(t)\big)\mathrm{d}W(t), & t \in [0,T] \\ X(0) = \xi \in L^2\big(\Omega, \mathcal{F}_0, \mathbb{P}; H\big). \end{cases} \tag{3.1.3}$$

In [42, proof of Theorem 6.2.3] it is shown that there exists a unique predictable mild solution of SEE (3.1.3), i.e.,

$$X(t) = S(t)\xi + \int_0^t S(t-s)F\big(X(s)\big)\mathrm{d}s + \int_0^t S(t-s)B\big(X(s)\big)\mathrm{d}W(s).$$

Moreover, due to [42, Lemma 6.2.9] and [42, Proposition 6.3.5], the process $(X(t))_{t\in[0,T]}$ has a continuous modification. Therefore the conditions in Assumption 3.1.1 are fulfilled with $S(t,s) = S(t-s)$, $\Psi(s) = F(X(s))$, and $\Phi(s) = B(X(s))$ for $0 \le s \le t \le T$ in the notation of Assumption 3.1.1.

Example 3.1.3. Assume the setting from Example 3.1.2 above. It is well-known; see, e.g., [15], that the linear implicit Euler approximations of the mild solution of the SEE (3.1.3), $\big(Y^N(t)\big)_{t\in\{0,\frac{T}{N},\dots,T\}}$, given by

$$\begin{aligned} Y^N\Big(\frac{(n+1)T}{N}\Big) = \Big(I - \frac{T}{N}A\Big)^{-1}\Bigg(& Y^N\big(\tfrac{nT}{N}\big) + \frac{T}{N}\cdot F\Big(Y^N\big(\tfrac{nT}{N}\big)\Big) \\ & + B\Big(Y^N\big(\tfrac{nT}{N}\big)\Big)\Big(W\Big(\frac{(n+1)T}{N}\Big) - W\Big(\frac{nT}{N}\Big)\Big)\Bigg), \end{aligned}$$

with $Y^N(0) = \xi$, $n = 0, 1, \ldots, N-1$, $N \in \mathbb{N}$, have continuous interpolations

$$\begin{aligned} Y^N(t) &= S^N(t,0)\xi + \int_0^t S^N(t, \lfloor s \rfloor_N) F\big(Y^N(\lfloor s \rfloor_N)\big)\mathrm{d}s + \int_0^t S^N(t, \lfloor s \rfloor_N) B\big(Y^N(\lfloor s \rfloor_N)\big)\mathrm{d}W(s) \\ &= S^N(t,0)\xi + \int_0^t S^N(t,s)\Big(I - A(s - \lfloor s \rfloor_N)\Big)^{-1} F\big(Y^N(\lfloor s \rfloor_N)\big)\mathrm{d}s \\ &\quad + \int_0^t S^N(t,s)\Big(I - A(s - \lfloor s \rfloor_N)\Big)^{-1} B\big(Y^N(\lfloor s \rfloor_N)\big)\mathrm{d}W(s), \end{aligned}$$

where $\big(S^N(t,s)\big)_{0 \le s \le t \le T} \subset L(H)$ is given by

$$S^N(t_2, t_1) = \Big(I - (t_1 - \lfloor t_1 \rfloor_N)A\Big)\Big(I - (t_2 - \lfloor t_2 \rfloor_N)A\Big)^{-1}\Big(I - \frac{T}{N}A\Big)^{(\lfloor t_1 \rfloor_N - \lfloor t_2 \rfloor_N)\frac{N}{T}},$$

for all t_1, $t_2 \in [0,T]$ with $t_1 < t_2$, all $N \in \mathbb{N}$, and

$$\lfloor t \rfloor_N := \max\Big\{ s \in \Big\{0, \frac{T}{N}, \frac{2T}{N}, \ldots, \frac{(N-1)T}{N}, T\Big\} \colon s \le t \Big\},$$

for all $t \in [0,T]$. Note that the semigroup $(S^N(t_2,t_1))_{0 \le t_1 \le t_2 \le T}$ depends explicitly on both variables t_1 and t_2 instead of the difference $t_2 - t_1$.

Example 3.1.4. Let Assumption 2.4.7 be fulfilled. Furthermore, let $(e_n)_{n \in \mathbb{N}}$ be an orthonormal basis of H, let $(\lambda_n)_{n\in\mathbb{N}} \subset (0,\infty)$ be an increasing sequence, and let the diagonal operator $A \colon D(A) \subset H \to H$ be such that $D(A) = \big\{ w \in H \colon \sum_{n \in \mathbb{N}} |\lambda_n|^2 |\left\langle e_n, w \right\rangle_H|^2 < \infty \big\}$ and for all $n \in \mathbb{N}$ it holds that $Ae_n = -\lambda_n e_n$. For $N \in \mathbb{N}$, let $(P_N)_{N\in\mathbb{N}} \subset L(H)$ satisfy $P_N(v) = \sum_{n=1}^N \left\langle e_n, v\right\rangle_H e_n$ for all $v \in H$, and let $X^N \colon [0,T] \times \Omega \to P_N(H)$ be a continuous analytically strong solution of

$$\begin{aligned} \mathrm{d}X^N(t) &= \Big[P_N A X^N(t) + P_N F(X^N(t))\Big]\mathrm{d}t + P_N B(X^N(t))\mathrm{d}W(t), \\ X^N(0) &= P_N \xi. \end{aligned}$$

The process $X^N = \big(X^N(t)\big)_{t\in[0,T]}$ is called spatial spectral Galerkin approximation [13], which satisfies Assumption 3.1.1 with

$$S(t,s) = e^{P_N A (t-s)},\ \Psi(s) = P_N F\big(X^N(s)\big),\ \Phi(s) = P_N B(X^N(s)),\ \text{and } X(0) = P_N \xi.$$

In the following we formulate our assumptions on the path-dependent functionals.

Assumption 3.1.5 (Functional f)**.**

(i) Let $(V, \|\cdot\|_V)$ be a separable real Hilbert space and let $f \colon [0,T] \times D([0,T],H) \to V$ be a mapping, where $T \in (0,\infty)$ and the separable real Hilbert space H are taken from Assumption 3.1.1.

(ii) Assume at each $(t, x) \in [0, T) \times D([0, T], H)$, the right-sided derivative $\partial_1^+ f(t, x) \in V$ and at each $(t, x) \in [0, T] \times D([0, T], H)$ the Fréchet derivatives $\partial_2 f(t, x) \in L(D([0, T], H), V)$ and $\partial_2^2 f(t, x) \in L^{(2)}(D([0, T], H), V)$ exist.

(iii) Let the mappings $f\colon [0, T] \times D([0, T], H) \to V$, $\partial_1^+ f\colon [0, T) \times D([0, T], H) \to V$, $\partial_2 f\colon [0, T] \times D([0, T], H) \to L(D([0, T], H), V)$ and $\partial_2^2 f\colon [0, T] \times D([0, T], H) \to L^{(2)}(D([0, T], H), V)$ be continuous and stay bounded on bounded subsets of $[0, T] \times D([0, T], H)$.

(iv) If $(h_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$ are bounded sequences in $D([0, T], H)$ and $h, g \in D([0, T], H)$ are such that for all $s \in [0, T]$ it holds that $\lim_{n\to\infty} h_n(s) = h(s)$ and $\lim_{n\to\infty} g_n(s) = g(s)$, let it hold for all $(t, x) \in [0, T] \times D([0, T], H)$ that

$$\partial_2 f(t, x)h_n \xrightarrow{n\to\infty} \partial_2 f(t, x)h \quad \text{(limit in } V\text{)},$$

and

$$\partial_2^2 f(t, x)(g_n, h_n) \xrightarrow{n\to\infty} \partial_2^2 f(t, x)(g, h) \quad \text{(limit in } V\text{)}.$$

Note that if a functional f satisfies the conditions in Assumption 3.1.5 above, then its second Fréchet derivative $\partial_2^2 f$ is a symmetric bounded bilinear operator; for details see Remark 2.3.3.

Example 3.1.6. (i) Let T, V and H be as in Assumption 3.1.5, and let $g\colon [0, T] \times H^m \to V$, $(t, (h_1, \ldots, h_m)) \mapsto g(t, (h_1, \ldots, h_m))$, be continuous and twice continuously Fréchet differentiable with respect to $(h_1, \ldots, h_m)$. Moreover, assume that the right-sided derivative $\partial_1^+ g(t, (h_1, \ldots, h_m))$ exists for all $t \in [0, T)$, $(h_1, \ldots, h_m) \in H^m$, that the mapping $\partial_1^+ g\colon [0, T) \times H^m \to V$ is continuous, and that $\partial_2 g$, $\partial_2^2 g$ and $\partial_1^+ g$ are bounded on bounded subsets of $[0, T] \times H^m$ and $[0, T) \times H^m$, respectively. Then the functional $f\colon [0, T] \times D([0, T], H) \to V$ defined by

$$f(t, x) = g(t, x(s_1), \ldots, x(s_n))$$

satisfies the conditions in Assumption 3.1.5. Indeed, note that

$$\partial_1^+ f(t, x) = \frac{\partial^+ g\big(t, \big(x(s_1), \ldots, x(s_m)\big)\big)}{\partial t},$$

for all $(t, x) \in [0, T) \times D([0, T], H)$ and that

$$\begin{aligned} \partial_2 f(t, x)(y) &= \partial_2 g\big(t, \big(x(s_1), \ldots, x(s_m)\big)\big)\big(y(s_1), \ldots, y(s_m)\big), \\ \partial_2^2 f(t, x)(y, z) &= \partial_2^2 g(t, x(s_1), \ldots, x(s_m))\big(y(s_1), \ldots, y(s_m), z(s_1), \ldots, z(s_m)\big), \end{aligned}$$

for all $t \in [0, T]$, x, y, and $z \in D([0, T], H)$.

(ii) Let T, V and H be as in Assumption 3.1.5, let μ be a finite Borel-measure on $[0, T]$, and let $g\colon [0, T] \times [0, T] \times H \to V$, $(t, s, h) \mapsto g(t, s, h)$, satisfy the following conditions:

(a) For all $(t,s) \in [0,T]^2$, the function $g(t,s,\cdot)\colon H \to V$ is two times continuously Fréchet differentiable.

(b) There exists a μ-measurable function $G\colon [0,T]\times[0,T] \to \mathbb{R}$ such that for all s, $t \in [0,T]$ and all h, h_1, $h_2 \in H$ with $\max\{\|h_1\|_H, \|h_2\|_H\} \leq 1$ it holds that $G(t,\cdot)$ is μ-integrable and that

$$\max\left\{ \|g(t,s,h)\|_V, \left\|\frac{\partial g\big((t,s,h\big)}{\partial x}(h_1)\right\|_V, \left\|\frac{\partial^2 g\big((t,s,h\big)}{\partial x^2}(h_1,h_2)\right\|_V \right\} \leq G(t,s).$$

Then the functional $f\colon [0,T]\times D([0,T],H) \to V$ defined by

$$f(t,x) = \int_{[0,T]} g(t,s,x(s))\mu(\mathrm{d}s)$$

satisfies the conditions in Assumption 3.1.5. Indeed, note that

$$\partial_1^+ f(t,x) = \int_{[0,T]} \frac{\partial^+ g(t,s,x(s))}{\partial t}\,\mu(\mathrm{d}s),$$

for all $(t,x) \in [0,T)\times D([0,T],H)$ and that

$$\begin{aligned} \partial_2 f(t,x)(y) &= \int_{[0,T]} \frac{\partial g\big(t,s,x(s)\big)}{\partial x}\big(y(s)\big)\,\mu(\mathrm{d}s), \quad \text{and} \\ \partial_2^2 f(t,x)(y,z) &= \int_{[0,T]} \frac{\partial^2 g\big((t,s,x(s)\big)}{\partial x^2}\big(y(s),z(s)\big)\,\mu(\mathrm{d}s), \end{aligned}$$

for all $t \in [0,T]$, x, y, and $z \in D([0,T],H)$ as a consequence of the assumptions on g and the dominated convergence theorem. See Lemma 3.5.4 below for details.

3.2 A path-dependent mild Itô formula

In this section, we introduce a path-dependent mild Itô formula in infinite-dimensional Hilbert spaces. In order to prove our path-dependent mild Itô formula, Theorem 3.2.2 below, we introduce a perturbation of the trajectories of $X = (X(t))_{t\in[0,T]}$, in (3.2.1) below.

Let Assumption 3.1.1 be given, let $X = (X(t))_{t\in[0,T]}$ be the H-valued mild Itô process given by (3.1.1), and set for all t, $r \in [0,T]$

$$X_t^S(r) = \begin{cases} X(r) & \text{if } r \in [0,t), \\ S(r,t)X(t) & \text{if } r \in [t,T]. \end{cases} \tag{3.2.1}$$

Then we have that

$$X_t^S(r) = S(r,0)X(0) + \int_0^{r\wedge t} S(r,s)\Psi(s)\mathrm{d}s + \int_0^{r\wedge t} S(r,s)\Phi(s)\mathrm{d}W(s). \tag{3.2.2}$$

It is clear that, under Assumption 3.1.1, for all $t \in [0,T]$ and $\mathbb{P}$-a.e. $\omega \in \Omega$ the trajectories $[0,T] \ni r \mapsto X_t^S(r,\omega) \in H$ of the H-valued process $(X_t^S(r))_{r\in[0,T]}$ are continuous, and therefore $X^S = (X_t^S)_{t\in[0,T]}$ is a stochastic process with values in $C([0,T],H) \subset D([0,T],H)$. Moreover, note that $X_T^S(r) = X(r)$ for all $r \in [0,T]$.

Lemma 3.2.1. Let Assumption 3.1.1 be fulfilled, let the H-valued process $(X(t))_{t\in[0,T]}$ be given by (3.1.1), and let the $C([0,T],H)$-valued process $(X_t^S)_{t\in[0,T]}$ be given by (3.2.2). Then the trajectories $[0,T] \ni t \mapsto X_t^S(\cdot,\omega) \in C([0,T],H)$, $\omega \in \Omega$, are continuous. Here and below we denote for every $(t,\omega) \in [0,T]\times\Omega$ by $X_t^S(\cdot,\omega) \in C([0,T],H)$ the mapping $[0,T] \ni s \mapsto X_t^S(s,\omega) \in H$.

Proof. We show the continuity at a fixed point $t_0 \in [0,T]$. To this end, we use the fact that for all $x \in C([0,T],H)$ the mapping

$$\left\{(t,s) \in [0,T]^2 \colon s \le t\right\} \ni (t,s) \mapsto S(t,s)x(s) \in H$$

is uniformly continuous, see Lemma 3.5.1 below for details. As a consequence, for $t > t_0$ it holds that

$$\begin{aligned}\left\|X_t^S - X_{t_0}^S\right\|_{C([0,T],H)} &= \sup_{r\in[0,T]} \left\|X_t^S(r) - X_{t_0}^S(r)\right\|_H \\ &\le \sup_{r\in[t_0,t]} \left\|S(r,r)X(r) - S(r,t_0)X(t_0)\right\|_H \\ &\quad + \sup_{r\in[t,T]} \left\|S(r,t)X(t) - S(r,t_0)X(t_0)\right\|_H\end{aligned}$$

goes to zero as t decreases to t_0. Therefore, the mapping $t \mapsto X_t^S(\cdot,\omega)$ is right-continuous at t_0 for all $\omega \in \Omega$. The left-continuity can be shown in the similar way: For $t < t_0$, observe that

$$\begin{aligned}\left\|X_t^S - X_{t_0}^S\right\|_{C([0,T],H)} &\le \sup_{r\in[t,t_0]} \left\|S(r,t)X(t) - S(r,r)X(r)\right\|_H \\ &\quad + \sup_{r\in[t_0,T]} \left\|S(r,t)X(t) - S(r,t_0)X(t_0)\right\|_H\end{aligned}$$

goes to zero as t increases to t_0. This implies the assertion. ∎

We are now able to present our path-dependent mild Itô formula:

Theorem 3.2.2. Let Assumption 3.1.1 be fulfilled, let $(X(t))_{t\in[0,T]}$ be given by (3.1.1), and let $f\colon [0,T]\times D([0,T],H) \to V$ satisfy Assumption 3.1.5. Moreover, let the $C([0,T],H)$-valued process $(X_t^S)_{t\in[0,T]}$ be given by (3.2.2) and let $\mathcal{U}$ be an orthonormal basis of U. Then it holds $\mathbb{P}$-a.s. that

$$\begin{aligned}\int_0^T &\Big\{\big\|\partial_1^+ f(s, X_s^S)\big\|_V + \big\|\partial_2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Psi(s)\big)\big\|_V \\ &+ \big\|\partial_2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)\big)\big\|^2_{\mathrm{HS}(U,V)} \\ &+ \sum_{u\in\mathcal{U}} \big\|\partial_2^2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u, \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u\big)\big\|_V\Big\}\mathrm{d}s \\ &< \infty\end{aligned} \tag{3.2.3}$$

and for all $t \in [0,T]$ we have $\mathbb{P}$-a.s. that

$$\begin{aligned} f(t, X_t^S) = f(0, X_0^S) &+ \int_0^t \partial_1^+ f(s, X_s^S)\,\mathrm{d}s \\ &+ \int_0^t \partial_2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Psi(s)\big)\,\mathrm{d}s \\ &+ \int_0^t \partial_2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)\,\mathrm{d}W(s)\big) \\ &+ \frac{1}{2}\int_0^t \sum_{u\in\mathcal{U}} \partial_2^2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u, \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u\big)\,\mathrm{d}s. \end{aligned} \tag{3.2.4}$$

Remark 3.2.3. Note that all integrals in (3.2.4) exist due to the integrability property (3.2.3) and the fact that the mappings

$$\begin{aligned} &[0,T]\times\Omega \ni (s,\omega) \mapsto \partial_1^+ f\big(s, X_s^S(\cdot,\omega)\big) \in V, \\ &[0,T]\times\Omega \ni (s,\omega) \mapsto \partial_2 f\big(s, X_s^S(\cdot,\omega)\big)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Psi(s,\omega)\big) \in V, \\ &[0,T]\times\Omega \ni (s,\omega) \mapsto \Big[U \ni u \mapsto \partial_2 f\big(s, X_s^S(\cdot,\omega)\big)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s,\omega)u\big) \in V\Big] \in \mathrm{HS}(U,V), \\ &[0,T]\times\Omega \ni (s,\omega) \mapsto \sum_{u\in\mathcal{U}} \partial_2^2 f\big(s, X_s^S(\cdot,\omega)\big)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s,\omega)u, \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s,\omega)u\big) \in V \end{aligned}$$

are $\mathcal{P}_T$ measurable. In particular, observe that for all $(s,\omega) \in [0,T]\times\Omega$, $u \in U$ we have $\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Psi(s,\omega) \in D([0,T],H)$ and $\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s,\omega)u \in D([0,T],H)$. Moreover, if the second mapping listed above is denoted by $\Xi\colon [0,T]\times\Omega \to \mathrm{HS}(U,V)$, then the stochastic integral appearing in (3.2.4) can be rewritten as $\int_0^t \Xi(s)\,\mathrm{d}W(s)$. See Section 3.5 below for details.

Remark 3.2.4 (Comparison with related results from literature)**.** (i) In order to compare the Itô formula (3.2.4) above with the functional Itô formula introduced in [12, 20], one can set the spaces U, H, and V equal to the space of real numbers $\mathbb{R}$ and the evolutionary family $S(t,s) = \mathrm{Id}_{\mathbb{R}}$, for $t,s \in [0,T]$. Then the path-valued random variable X_t^S defined in (3.2.1) is simply the stopped Itô process $X_t = X(\cdot \wedge t)$ and the formula in (3.2.4) can be rewritten as

$$\begin{aligned} f(t, X_t) = f(0, X_0) &+ \int_0^t \partial_1^+ f(s, X_s)\, \mathrm{d}s \\ &+ \int_0^t \partial_2 f(s, X_s)\big(\mathbb{1}_{[s,T]}(\cdot)\Psi(s)\big)\, \mathrm{d}s \\ &+ \int_0^t \partial_2 f(s, X_s)\big(\mathbb{1}_{[s,T]}(\cdot)\Phi(s)\big)\, \mathrm{d}W(s) \\ &+ \frac{1}{2}\int_0^t \partial_2^2 f(s, X_s)\big(\mathbb{1}_{[s,T]}(\cdot)\Phi(s), \mathbb{1}_{[s,T]}(\cdot)\Phi(s)\big)\, \mathrm{d}s. \end{aligned} \tag{3.2.5}$$

Considering the "non-anticipativity" assumption on the functional f in [12], observe that the right-sided time-derivative $\partial_1^+ f(s, X_s)$ and Fréchet derivatives $\partial_2 f(s, X_s)\big(\mathbb{1}_{[s,T]}(\cdot)h(s)\big)$, $h \in H$ appearing in (3.2.5) correspond to the horizontal and vertical derivatives in [12], respectively. Therefore, the formula (3.2.4) corresponds to the functional Itô formula introduced in [12, 20].

(ii) The infinite-dimensional functional Itô formula, presented in [52], is similar to our Itô formula in (3.2.4) if we set $S(t,s) = \mathrm{Id}_H$ for all $t, s \in [0,T]$. Note that the notation and assumptions in [52] are slightly different than in our setting. For instance, in [52] a left-sided time-derivative is used instead of right-sided time-derivative.

(iii) Let us finally compare Theorem 3.2.2 to the mild Itô formula in [15]. To this end let in Theorem 3.2.2 above $t \in [0,T]$ be fixed. If for all $s \in [0,T]$ and $x \in D([0,T], H)$ it holds that $f(s,x) = \tilde{f}(s, x(t))$ with an $\tilde{f} \in C^{1,2}([0,T] \times H, V)$, then $f \in C^{1,2}([0,T] \times D([0,T], H), V)$ and the formula (3.2.4) implies that

$$\begin{aligned} \tilde{f}(t, X(t)) = \tilde{f}(0, S(t,0)X(0)) &+ \int_0^t \partial_1^+ \tilde{f}\big(s, S(t,s)X(s)\big)\, \mathrm{d}s \\ &+ \int_0^t \partial_2 \tilde{f}\big(s, S(t,s)X(s)\big)\big(S(t,s)\Psi(s)\big)\, \mathrm{d}s \\ &+ \int_0^t \partial_2 \tilde{f}\big(s, S(t,s)X(s)\big)\big(S(t,s)\Phi(s)\big)\, \mathrm{d}W(s) \\ &+ \frac{1}{2}\int_0^t \sum_{u\in\mathcal{U}} \partial_2^2 \tilde{f}\big(s, S(t,s)X(s)\big)\big(S(t,s)\Phi(s)u, S(t,s)\Phi(s)u\big)\, \mathrm{d}s, \end{aligned} \tag{3.2.6}$$

which coincides, for the fixed t, with the mild Itô formula introduced in [15].

Remark 3.2.5 (alternative version of the path-dependent mild Itô formula)**.** Here we present an alternative, slightly modified version of the path-dependent mild Itô formula (3.2.4) involving a suitable evolution family $(S_{t,s})_{0\le s\le t\le T}$ of bounded linear operators on the path-space $D([0,T], H)$.

To this end, let Assumption 3.1.1 be fulfilled and for every $t \in [0,T]$ let $X_t = \big(X(r\wedge t)\big)_{r\in[0,T]}$ be the process stopped at t, which we identify with the corresponding $C([0,T], H)$-valued random variable. Therefore $(X_t)_{t\in[0,T]}$ is a stochastic process with values in $C([0,T],H) \subset D([0,T],H)$. For all $s,t \in [0,T]$ with $s \leq t$ let $S_{t,s} \in L(D([0,T],H))$ be defined by

$$\begin{aligned}\big(S_{t,s}x\big)(r) &= \mathbb{1}_{[0,s)}(r)x(r) + \mathbb{1}_{[s,t)}(r)S(r,s)x(r) + \mathbb{1}_{[t,T]}(r)S(t,s)x(r)\\ &= \mathbb{1}_{[0,s)}(r)x(r) + \mathbb{1}_{[s,T]}(r)S(r\wedge t,s)x(r),\end{aligned} \tag{3.2.7}$$

$x \in D([0,T],H)$, $r \in [0,T]$. Note that $(S_{t,s})_{0\leq s\leq t\leq T} \subset L(D([0,T],H))$ is a strongly continuous evolution family on $D([0,T],H)$. To see this, let $r,s,t \in [0,T]$ with $r \leq s \leq t$ and let $x \in D([0,T],H)$. Observe that $S_{t,t} = \mathrm{Id}_{D([0,T],H)}$ and that for all $u \in [0,T]$ it holds that

$$\begin{aligned}\big(S_{t,s}S_{s,r}x\big)(u) &= \mathbb{1}_{[0,s)}(u)\big(S_{s,r}x\big)(u) + \mathbb{1}_{[s,T]}(u)S(u\wedge t,s)\big(S_{s,r}x\big)(u)\\ &= \mathbb{1}_{[0,s)}(u)\Big[\mathbb{1}_{[0,r)}(u)x(u) + \mathbb{1}_{[r,T]}(u)S(u\wedge s,r)x(u)\Big]\\ &\quad + \mathbb{1}_{[s,T]}(u)S(u\wedge t,s)\Big[\mathbb{1}_{[0,r)}(u)x(u) + \mathbb{1}_{[r,T]}(u)S(u\wedge s,r)x(u)\Big]\\ &= \mathbb{1}_{[0,r)}(u)x(u) + \mathbb{1}_{[r,s)}(u)S(u\wedge s,r)x(u) + \mathbb{1}_{[s,T]}(u)S(u\wedge t,s)S(u\wedge s,r)x(r)\\ &= \mathbb{1}_{[0,r)}(u)x(u) + \mathbb{1}_{[r,s)}(u)S(u\wedge t,r)x(u) + \mathbb{1}_{[s,T]}(u)S(u\wedge t,s)S(s,r)x(r)\\ &= \big(S_{t,r}x\big)(u).\end{aligned} \tag{3.2.8}$$

The strong continuity of $(S_{t,s})_{0\leq s\leq t\leq T}$ can be shown in a similar way as in the proof of Lemma 3.2.1. Next suppose that the mapping $g\colon [0,T]\times D([0,T],H) \to V$ fulfills the conditions formulated in Assumption 3.1.5 with g in place of f, and let $\mathcal{U}$ be an orthonormal basis of U. Then, by reasoning along the lines of the proof of Theorem 3.2.2 one obtains that for all $t \in [0,T]$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned}&\int_0^T \Big\{\big\|\partial_1^+ g(s,S_{t,s}X_s)\big\|_V + \big\|\partial_2 g(s,S_{t,s}X_s)S_{t,s}\big(\mathbb{1}_{[s,T]}(\cdot)\Psi(s)\big)\big\|_V\\ &\quad + \big\|\partial_2 g(s,S_{t,s}X_s)S_{t,s}\big(\mathbb{1}_{[s,T]}(\cdot)\Phi(s)\big)\big\|^2_{\mathrm{HS}(U,V)}\\ &\quad + \sum_{u\in\mathcal{U}}\Big\|\partial_2^2 g(s,S_{t,s}X_s)\Big(S_{t,s}\big(\mathbb{1}_{[s,T]}(\cdot)\Phi(s)u\big),S_{t,s}\big(\mathbb{1}_{[s,T]}(\cdot)\Phi(s)u\big)\Big)\Big\|_V\Big\}\mathrm{d}s\\ &< \infty\end{aligned} \tag{3.2.9}$$

and

$$\begin{aligned}g(t,X_t) &= g(0,S_{t,0}X_0) + \int_0^t \partial_1 g(s,S_{t,s}X_s)\mathrm{d}s\\ &\quad + \int_0^t \partial_2 g(s,S_{t,s}X_s)S_{t,s}\big(\mathbb{1}_{[s,T]}(\cdot)\Psi(s)\big)\mathrm{d}s\\ &\quad + \int_0^t \partial_2 g(s,S_{t,s}X_s)S_{t,s}\big(\mathbb{1}_{[s,T]}(\cdot)\Phi(s)\big)\mathrm{d}W(s)\\ &\quad + \frac{1}{2}\int_0^t \sum_{u\in\mathcal{U}}\partial_2^2 g(s,S_{t,s}X_s)\Big(S_{t,s}\big(\mathbb{1}_{[s,T]}(\cdot)\Phi(s)u\big),S_{t,s}\big(\mathbb{1}_{[s,T]}(\cdot)\Phi(s)u\big)\Big)\mathrm{d}s.\end{aligned} \tag{3.2.10}$$

In particular, analogous measurability assertions to the ones in Remark 3.2.3 hold true.

To elucidate the relation of (3.2.10) and (3.2.4), let $t \in [0,T]$ be fixed, let the mapping $\mathcal{J}_t \colon D([0,T],H) \to D([0,T],H)$ be defined by $\mathcal{J}_t(x) = \mathbb{1}_{[0,t)}(\cdot)x(\cdot) + \mathbb{1}_{[t,T]}(\cdot)S(\cdot,t)x(\cdot)$, $x \in D([0,T],H)$, and observe that for all $s \in [0,t]$ it holds that

$$X_s^S = \mathcal{J}_t(S_{t,s}X_s), \tag{3.2.11}$$

where we use the notation introduced in (3.2.1). Thus, if : $[0,T] \times D([0,T],H) \to V$ is a further mapping satisfying the conditions formulated in Assumption 3.1.5, the identity (3.2.4) coincides with (3.2.10) for the specific choice

$$g(s,x) = f(s,\mathcal{J}_t(x)),$$

$s \in [0,T]$, $x \in D([0,T],H)$. Indeed, it holds for all $h \in H$ that $\mathcal{J}_t\big(S_{t,s}\big(\mathbb{1}_{[s,T]}(\cdot)h\big)\big) = \mathbb{1}_{[s,T]}(\cdot)S(\cdot,t)h$ and by Definition 2.3.2 that

$$\partial_2 \tilde{f}(s,x)(y) = \partial_2 f\big(s,\mathcal{J}_t(x)\big)\big(\mathcal{J}_t(y)\big), \tag{3.2.12}$$

$x, y \in D([0,T],H)$.

Remark 3.2.6 (simplified setting for the path-dependent mild Itô formula)**.** The proof of our mild Itô formula is significantly simplified if we work in the much more restrictive setting obtained by replacing the Banach spaces $C([0,T],H)$ and $D([0,T],H)$ in the setting of Remark 3.2.5 above by the Hilbert space $L^2(\mu;H) = L^2([0,T],\mathcal{B}([0,T]),\mu;H)$, where μ is some given finite Borel-measure on $[0,T]$. To discuss this in some more detail, let Assumption 3.1.1 be fulfilled and, as in Remark 3.2.5 above, for every $t \in [0,T]$ let $X_t = \big(X(r \wedge t)\big)_{r\in[0,T]}$ be the process stopped at t. We now interpret each X_t as an $L^2(\mu;H)$-valued random variable, so that $(X_t)_{t\in[0,T]}$ is an $L^2(\mu;H)$-valued stochastic process. Similarly, the operators $S_{t,s}$, $0 \le s \le t \le T$, defined in (3.2.7) are now considered as operators on $L^2(\mu;H)$, i.e., $S_{t,s} \in L(L^2(\mu;H))$. It then follows that $(S_{t,s})_{0\le s\le t\le T}$ is a strongly continuous evolution family on $L^2(\mu;H)$ and that the $L^2(\mu;H)$-valued process $(X_t)_{t\in[0,T]}$ is in fact an $L^2(\mu;H)$-valued mild Itô process which satisfies for all $t \in [0,T]$ that

$$X_t = S_{t,0}X_0 + \int_0^t S_{t,s}\big(\mathbb{1}_{[s,T]}(\cdot)\Psi(s)\big)\mathrm{d}s + \int_0^t S_{t,s}\big(\mathbb{1}_{[s,T]}(\cdot)\Phi(s)\big)\mathrm{d}W(s), \tag{3.2.13}$$

$\mathbb{P}$-a.s. as an equality in $L^2(\mu;H)$. Here the Bochner integral and the stochastic integral on the right hand side are meant to be $L^2(\mu;H)$-valued integrals; compare Lemma 4.4.1 in Section 4.4 below for related considerations. As the mild Itô process (3.2.13) essentially fits into the framework in [15], we are now able to directly apply the "standard" mild Itô formula in [15] to obtain that the identity (3.2.10) in Remark 3.2.5 above holds true for every $g \in C^{1,2}\big([0,T] \times$

$L^2(\mu; H), V\big)$. Note, however, that the simplified setting considered here is much more restrictive than the setting considered in Remark 3.2.5 above. An example of a mapping $g\colon [0,T] \times D([0,T], H) \to V$ that can not be extended to a mapping from $[0,T] \times L^2(\mu; H)$ to V is given by

$$g(t,x) = \mathbb{E}\Big[\phi\big(\mathbb{1}_{[0,t)}(\cdot)x(\cdot) + \mathbb{1}_{[t,T]}(\cdot)X^{t,x(t)}(\cdot)\big)\Big], \tag{3.2.14}$$

$t \in [0,T]$, $x \in D([0,T], H)$, where $\phi\colon D([0,T], H) \to V$ is sufficiently regular and $\big(X^{t,\xi}(r)\big)_{r\in[t,T]}$, $t \in [0,T]$, $\xi \in H$, is a family of mild solutions (in a probabilistically strong sense) of a given semilinear SPDE with respective starting time t and starting position ξ. Such functionals are of interest in the context of path-dependent Kolmogorov equations; compare, e.g., [5, 35].

3.3 Proof of the path-dependent mild Itô formula

To present the proof of Theorem 3.2.2 we need some auxiliary results and one new notation. For $t_j^n = \frac{j}{2^n}T$, $j \in \{0, 1, \dots, 2^n\}$, $n \in \mathbb{N}$, and $r \in [0,T]$ we define the H-valued random variables $Y_j^n(r)$ by

$$Y_j^n(r) = \sum_{k=1}^{2^n} \mathbb{1}_{[t_{k-1}^n, t_k^n)}(r) X_{t_j^n}^S(t_k^n) + \mathbb{1}_{\{T\}}(r) X_{t_j^n}^S(T).$$

Note that $Y_j^n = Y_j^n(\cdot)$ is thus a $D([0,T], H)$-valued random variable,

$$Y_j^n(\cdot) = \sum_{k=1}^{2^n} \mathbb{1}_{[t_{k-1}^n, t_k^n)}(\cdot) X_{t_j^n}^S(t_k^n) + \mathbb{1}_{\{T\}}(\cdot) X_{t_j^n}^S(T). \tag{3.3.1}$$

The idea of the definitions of X_t^S and Y_j^n are illustrated in Figure 3.1 below.

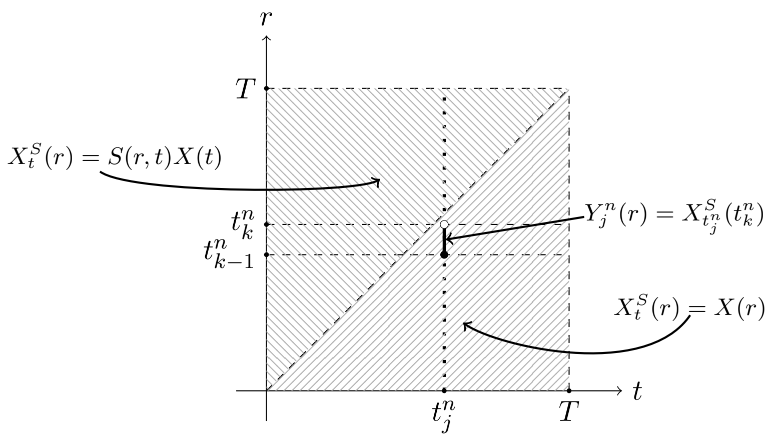

FIGURE 3.1: Perturbation of sample paths of X and its approximation via Y_j^n.

Lemma 3.3.1. The $D([0,T],H)$-valued processes $(Y_j^n)_{j=0,\dots,2^n}$, $n \in \mathbb{N}$, defined in (3.3.1), are uniform approximations of $(X_t^S)_{t\in[0,T]}$ in the following sense:

$$\max_{j=0,\dots,2^n} \sup_{t\in[t_{j-1}^n,t_j^n)} \left\|Y_j^n(\cdot,\omega) - X_t^S(\cdot,\omega)\right\|_{D([0,T],H)} \xrightarrow{n\to\infty} 0 \quad \text{for all } \omega \in \Omega.$$

Proof. The assertion follows from the fact that the mapping

$$[0,T]^2 \ni (t,r) \mapsto X_t^S(r,\omega) \in H$$

is continuous, by Lemma 3.2.1, and therefore uniformly continuous. Indeed, it holds that

$$\max_{k,j=0,\dots,2^n} \sup_{t,r\in[0,T]} \mathbb{1}_{[t_{j-1}^n,t_j^n)}(t)\, \mathbb{1}_{[t_{k-1}^n,t_k^n)}(r) \left\|X_{t_j^n}^S(t_k^n) - X_t^S(r)\right\|_H \xrightarrow{n\to\infty} 0.$$

□

The following auxiliary result is closely related to the so-called Λ-Lemma in [20]. It is crucial that the value δ appearing in (3.3.2) below does not depend on t.

Lemma 3.3.2. Let $T \in (0,\infty)$, let B be a real Banach space, and let H be a real Hilbert space. If the mappings $f\colon [0,T]\times D([0,T],H) \to B$ and $x_\cdot\colon [0,T] \to D([0,T],H)$, $[0,T] \ni t \mapsto x_t \in D([0,T],H)$, are continuous, then it holds that

$$\begin{gathered} \forall \varepsilon > 0\ \exists \delta > 0\colon\ \forall y \in D([0,T],H),\ \forall s,t \in [0,T] \\ \Big(|t-s| + \|x_t - y\|_{D([0,T],H)} < \delta \Rightarrow \|f(t,x_t) - f(s,y)\|_B < \varepsilon\Big). \end{gathered} \tag{3.3.2}$$

Proof. Assume that the assertion does not hold. Then there exist $\varepsilon > 0$, sequences $(t_n)_{n\in\mathbb{N}}$, $(s_n)_{n\in\mathbb{N}}$ in $[0,T]$, and $(y_n)_{n\in\mathbb{N}}$ in $D([0,T],H)$ such that $|t_n - s_n| + \|x_{t_n} - y_n\|_{D([0,T],H)} < \frac{1}{n}$ and $\|f(t_n,x_{t_n}) - f(s_n,y_n)\|_B > \varepsilon$ for all $n \in \mathbb{N}$. Let $(t_{n_k})_{k\in\mathbb{N}}$ be a subsequence of $(t_n)_{n\in\mathbb{N}}$ which converges to some $t^* \in [0,T]$ and write

$$\begin{aligned} \left\|f(t_{n_k},x_{t_{n_k}}) - f(s_{n_k},y_{n_k})\right\|_B &\le \left\|f(t_{n_k},x_{t_{n_k}}) - f(t^*,x_{t^*})\right\|_B \\ &\quad + \left\|f(t^*,x_{t^*}) - f(s_{n_k},y_{n_k})\right\|_B. \end{aligned} \tag{3.3.3}$$

Notice that

$$\begin{aligned} \left|t^* - s_{n_k}\right| + \left\|x_{t^*} - y_{n_k}\right\|_{D([0,T],H)} &< \left|t^* - t_{n_k}\right| + \left\|x_{t^*} - x_{t_{n_k}}\right\|_{D([0,T],H)} \\ &\quad + \left|t_{n_k} - s_{n_k}\right| + \left\|x_{t_{n_k}} - y_{n_k}\right\|_{D([0,T],H)} \\ &\le \frac{1}{n_k} + \left|t_{n_k} - t^*\right| + \left\|x_{t_{n_k}} - x_{t^*}\right\|_{D([0,T],H)} \xrightarrow{n\to\infty} 0, \end{aligned}$$

by continuity of $s \mapsto x_s$. Now the continuity of f at (t^*, x_{t^*}) implies that the right-hand-side of the inequality in (3.3.3) goes to zero. This contradicts the assumption that $\|f(t_n, x_{t_n}) - f(s_n, y_n)\|_B > \varepsilon$ for all $n \in \mathbb{N}$. □

Proof of Theorem 3.2.2. For the sake of notational simplicity, we prove formula (3.2.4) for the case $t = T$:

$$\begin{aligned} f(T, X_T^S) = f(0, X_0^S) &+ \int_0^T \partial_1^+ f(s, X_s^S)\, \mathrm{d}s \\ &+ \int_0^T \partial_2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot) S(\cdot, s)\Psi(s)\big)\, \mathrm{d}s \\ &+ \int_0^T \partial_2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot) S(\cdot, s)\Phi(s)\, \mathrm{d}W(s)\big) \\ &+ \frac{1}{2}\int_0^T \sum_{u\in\mathcal{U}} \partial_2^2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot) S(\cdot, s)\Phi(s)u, \mathbb{1}_{[s,T]}(\cdot) S(\cdot, s)\Phi(s)u\big)\, \mathrm{d}s. \end{aligned} \tag{3.3.4}$$

The proof for an arbitrary $t \in [0, T]$ is more complicated from a notational point of view but not from a mathematical point of view. To simplify the exposition, we first describe the strategy of our proof. The single arguments are carried out in Steps 1-3 below. We begin by showing the integrability assertion (3.2.3) in Step 1. Further, a stopping argument in Step 2 shows that without loss of generality we can assume that

$$\begin{aligned} \sup_{t\in[0,T]} \max\Bigg(\|X(0)\|_H, \int_0^T \|\Psi(s)\|_H\, \mathrm{d}s, \int_0^T \|\Phi(s)\|^2_{\mathrm{HS}(U,H)}\, \mathrm{d}s,& \\ \left\| \int_0^t S(t,s)\Phi(s)\mathrm{d}W(s) \right\|_H \Bigg) < N \quad \mathbb{P}\text{-a.s.}& \end{aligned} \tag{3.3.5}$$

for some $N \in (0, \infty)$. The main idea of our proof is to use the $D([0, T], H)$-valued processes $(Y_j^n)_{j\in\{0,\ldots,2^n\}}$, $n \in \mathbb{N}$, defined in (3.3.1), in order to write

$$\begin{aligned} f(T, X_T^S) - f(0, X_0^S) &= \Big(f(T, X_T^S) - f(T, Y_n^n)\Big) + \sum_{j=1}^{2^n} \Big(f(t_j^n, Y_j^n) - f(t_{j-1}^n, Y_{j-1}^n)\Big) \\ &\quad + \Big(f(0, Y_0^n) - f(0, X_0^S)\Big) \\ &= \Big(f(T, X_T^S) - f(T, Y_n^n)\Big) + \sum_{j=1}^{2^n} \Big(f(t_j^n, Y_{j-1}^n) - f(t_{j-1}^n, Y_{j-1}^n)\Big) \\ &\quad + \sum_{j=1}^{2^n} \Big(f(t_j^n, Y_j^n) - f(t_j^n, Y_{j-1}^n)\Big) + \Big(f(0, Y_0^n) - f(0, X_0^S)\Big) \\ &= A^n + B^n + C^n + D^n, \end{aligned} \tag{3.3.6}$$

where

$$\begin{aligned} A^n &= f(T, X^S_T) - f(T, Y^n_n), \\ B^n &= \sum_{j=1}^{2^n} \Big(f(t^n_j, Y^n_{j-1}) - f(t^n_{j-1}, Y^n_{j-1}) \Big), \\ C^n &= \sum_{j=1}^{2^n} \Big(f(t^n_j, Y^n_j) - f(t^n_j, Y^n_{j-1}) \Big), \\ D^n &= f(0, Y^n_0) - f(0, X^S_0). \end{aligned} \tag{3.3.7}$$

A straightforward argumentation in Step 3 shows that for $\mathbb{P}$-a.s. $\omega \in \Omega$ we have

$$\begin{aligned} A^n(\omega) &\xrightarrow{n\to\infty} 0, \\ B^n(\omega) &\xrightarrow{n\to\infty} \int_0^T \partial_1^+ f(s, X^S_s(\omega)) \mathrm{d}s, \\ D^n(\omega) &\xrightarrow{n\to\infty} 0. \end{aligned}$$

In order to handle the more complicated term C^n in (3.3.6), we set

$$\Delta Y^n_j = Y^n_{j+1} - Y^n_j, \quad j = 0, \dots, 2^n - 1, \tag{3.3.8}$$

and observe that by (3.2.2) and (3.3.1), we have

$$\begin{aligned} \Delta Y^n_{j-1} &= \sum_{k=j}^{2^n} \mathbb{1}_{[t^n_{k-1}, t^n_k)}(\cdot) \Big(X^S_{t^n_j}(t^n_k) - X^S_{t^n_{j-1}}(t^n_k) \Big) + \mathbb{1}_{\{T\}}(\cdot) \Big(X^S_{t^n_j}(T) - X^S_{t^n_{j-1}}(T) \Big) \\ &= \sum_{k=j}^{2^n} \mathbb{1}_{[t^n_{k-1}, t^n_k)}(\cdot) \bigg(\int_{t^n_k \wedge t^n_{j-1}}^{t^n_k \wedge t^n_j} S(t^n_k, s)\Psi(s) \mathrm{d}s + \int_{t^n_k \wedge t^n_{j-1}}^{t^n_k \wedge t^n_j} S(t^n_k, s)\Phi(s) \mathrm{d}W(s) \bigg) \\ &\quad + \mathbb{1}_{\{T\}}(\cdot) \bigg(\int_{t^n_{j-1}}^{t^n_j} S(T, s)\Psi(s) \mathrm{d}s + \int_{t^n_{j-1}}^{t^n_j} S(T, s)\Phi(s) \mathrm{d}W(s) \bigg) \\ &= \bigg(\sum_{k=j}^{2^n} \mathbb{1}_{[t^n_{k-1}, t^n_k)}(\cdot) S(t^n_k, t^n_j) + \mathbb{1}_{\{T\}}(\cdot) S(T, t^n_j) \bigg) \\ &\quad \times \bigg(\int_{t^n_{j-1}}^{t^n_j} S(t^n_j, s)\Psi(s) \mathrm{d}s + \int_{t^n_{j-1}}^{t^n_j} S(t^n_j, s)\Phi(s) \mathrm{d}W(s) \bigg). \end{aligned} \tag{3.3.9}$$

For the sake of better readability, we define for $j \in \{0, 1, \dots, 2^n - 1\}$, $n \in \mathbb{N}$, the linear operators $\Gamma^n_j \in L(H, D([0,T], H))$ by

$$\Gamma^n_j h = \sum_{k=j}^{2^n} \mathbb{1}_{[t^n_{k-1}, t^n_k)}(\cdot) S(t^n_k, t^n_j) h + \mathbb{1}_{\{T\}}(\cdot) S(T, t^n_j) h \qquad h \in H. \tag{3.3.10}$$

Therefore, we can rewrite (3.3.9) in the form

$$\Delta Y_{j-1}^n = \Gamma_j^n \bigg(\int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Psi(s)\mathrm{d}s + \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s) \bigg). \tag{3.3.11}$$

Looking back at the term C^n in (3.3.7), we see that, by the Taylor expansion presented in Lemma 2.3.4, the following equalities hold:

$$\begin{aligned} C^n &= \sum_{j=1}^{2^n} \Big(f(t_j^n, Y_{j-1}^n + \Delta Y_{j-1}^n) - f(t_j^n, Y_{j-1}^n) \Big) \\ &= \sum_{j=1}^{2^n} \bigg\{ \partial_2 f(t_j^n, Y_{j-1}^n)(\Delta Y_{j-1}^n) + \frac{1}{2}\partial_2^2 f(t_j^n, Y_{j-1}^n)(\Delta Y_{j-1}^n, \Delta Y_{j-1}^n) \\ &\quad + \int_0^1 (1-\theta)\Big[\partial_2^2 f(t_j^n, Y_{j-1}^n + \theta \Delta Y_{j-1}^n) - \partial_2^2 f(t_j^n, Y_{j-1}^n)\Big] \big(\Delta Y_{j-1}^n, \Delta Y_{j-1}^n\big)\mathrm{d}\theta \bigg\} \\ &= C_1^n + C_2^n + C_3^n + C_4^n + C_5^n + C_6^n, \end{aligned} \tag{3.3.12}$$

where

$$\begin{aligned} C_1^n &= \sum_{j=1}^{2^n} \partial_2 f(t_j^n, Y_{j-1}^n)\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Psi(s)\mathrm{d}s, \\ C_2^n &= \sum_{j=1}^{2^n} \partial_2 f(t_j^n, Y_{j-1}^n)\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s), \\ C_3^n &= \sum_{j=1}^{2^n} \frac{1}{2}\partial_2^2 f(t_j^n, Y_{j-1}^n)\bigg(\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Psi(s)\mathrm{d}s, \Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Psi(s)\mathrm{d}s\bigg), \\ C_4^n &= \sum_{j=1}^{2^n} \partial_2^2 f(t_j^n, Y_{j-1}^n)\bigg(\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Psi(s)\mathrm{d}s, \Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s)\bigg), \\ C_5^n &= \sum_{j=1}^{2^n} \frac{1}{2}\partial_2^2 f(t_j^n, Y_{j-1}^n)\bigg(\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s), \Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s)\bigg), \\ C_6^n &= \sum_{j=1}^{2^n} \int_0^1 (1-\theta)\Big[\partial_2^2 f(t_j^n, Y_{j-1}^n + \theta \Delta Y_{j-1}^n) - \partial_2^2 f(t_j^n, Y_{j-1}^n)\Big] \big(\Delta Y_{j-1}^n, \Delta Y_{j-1}^n\big)\mathrm{d}\theta. \end{aligned} \tag{3.3.13}$$

We will later in Step 4 investigate the terms above and show that

$$\begin{aligned} &C_1^n \xrightarrow{n\to\infty} \int_0^T \partial_2 f(s, X_s^S)\Big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Psi(s)\Big)\mathrm{d}s \quad \mathbb{P}\text{-a.s.} \\ &C_2^n \xrightarrow{n\to\infty} \int_0^T \partial_2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)\big)\,\mathrm{d}W(s) \quad \text{in probability,} \\ &C_5^n \xrightarrow{n\to\infty} \frac{1}{2}\int_0^T \sum_{u\in\mathcal{U}} \partial_2^2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)u, \mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)u\big)\mathrm{d}s \quad \text{in probability,} \\ &C_i^n \xrightarrow{n\to\infty} 0, \quad i = 3, 4, 6. \end{aligned}$$

Step 1: We begin by showing the integrability assertion (3.2.3). Due to Lemma 3.2.1, the mapping $[0,T] \ni t \mapsto X_t^S(\cdot,\omega) \in C([0,T],H)$ is continuous for all $\omega \in \Omega$. Therefore

$$\sup_{s\in[0,T]} \left\|X_s^S(\cdot,\omega)\right\|_{C([0,T],H)} < \infty,$$

and since $\partial_1^+ f$, $\partial_2 f$ and $\partial_2^2 f$ are bounded on bounded subsets of $[0,T]\times D([0,T],H)$, it follows that

$$K(\omega) = \sup_{s\in[0,T]} \Big\{ \left\|\partial_1^+ f\big(s, X_s^S(\cdot,\omega)\big)\right\|_V, \left\|\partial_2 f\big(s, X_s^S(\cdot,\omega)\big)\right\|_{L(D([0,T],H),V)}, \\ \left\|\partial_2^2 f\big(s, X_s^S(\cdot,\omega)\big)\right\|_{L^{(2)}(D([0,T],H),V)} \Big\} < \infty.$$

Now observe that it holds $\mathbb{P}$-a.s. that

$$\begin{aligned}
&\int_0^T \bigg\{ \left\|\partial_1^+ f(s, X_s^S)\right\|_V + \left\|\partial_2 f(s, X_s^S)\Big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Psi(s)\Big)\right\|_V \\
&\qquad + \left\|\partial_2 f(s, X_s^S)\Big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)\Big)\right\|_{\mathrm{HS}(U,V)}^2 \\
&\qquad + \sum_{u\in\mathcal{U}} \left\|\partial_2^2 f(s, X_s^S)\Big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u, \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u\Big)\right\|_V \bigg\} \mathrm{d}s \\
&\quad \le K \int_0^T \bigg\{ 1 + \sup_{0\le u\le v\le T} \|S(v,u)\|_{L(H)} \|\Psi(s)\|_H \\
&\qquad + K \sup_{0\le r\le t\le T} \|S(t,r)\|_{L(H)}^2 \|\Phi(s)\|_{\mathrm{HS}(U,H)}^2 \\
&\qquad + \sup_{0\le r\le t\le T} \|S(t,r)\|_{L(H)}^2 \|\Phi(s)\|_{\mathrm{HS}(U,H)}^2 \bigg\} \mathrm{d}s < \infty,
\end{aligned}$$

which shows the integrability property (3.2.3).

Step 2: As mentioned before, we would like to reduce the assertion (3.3.4) to the bounded case (3.3.5). For each $N \in \mathbb{N}$, define the stopping times

$$\tau_N = \inf\bigg\{ t\in[0,T]\colon \max\bigg(\|X(0)\|_H, \int_0^t \|\Psi(s)\|_H \,\mathrm{d}s, \int_0^t \|\Phi(s)\|_{\mathrm{HS}(U,H)}^2 \,\mathrm{d}s, \\ \left\| \int_0^t S(t,s)\Phi(s)\mathrm{d}W(s)\right\|_H \bigg) > N \bigg\} \wedge T, \tag{3.3.14}$$

with $\inf\emptyset := \infty$, and also the stochastic intervals

$$(\!(0,\tau_N]\!] = \big\{(t,\omega)\in(0,T]\times\Omega\colon t\le\tau_N(\omega)\big\} \in \mathcal{P}_T,$$

where the predictability follows from, e.g., [51, proof of Proposition 2.3.8]. Let $X^N(0) = \mathbb{1}_{\{\|X(0)\|_H \leq N\}} X(0)$, $\Psi^N = \mathbb{1}_{(\!(0,\tau_N]\!]}\Psi$, $\Phi^N = \mathbb{1}_{(\!(0,\tau_N]\!]}\Phi$. Then $\big(X^N(t)\big)_{t\in[0,T]}$ given by

$$X^N(t) = S(t,0)X^N(0) + \int_0^t S(t,s)\Psi^N(s)\mathrm{d}s + \int_0^t S(t,s)\Phi^N(s)\mathrm{d}W(s)$$

is well-defined. Due to Assumption 3.1.1, the processes $\big(I(t)\big)_{t\in[0,T]} = \Big(\int_0^t S(t,s)\Psi(s)\mathrm{d}s\Big)_{t\in[0,T]}$ and $\big(J(t)\big)_{t\in[0,T]} = \Big(\int_0^t S(t,s)\Phi(s)\mathrm{d}W(s)\Big)_{t\in[0,T]}$ have continuous modifications, which implies that the processes $\big(I^N(t)\big)_{t\in[0,T]} = \Big(\int_0^t S(t,s)\Psi^N(s)\mathrm{d}s\Big)_{t\in[0,T]}$ and $\big(J^N(t)\big)_{t\in[0,T]} = \Big(\int_0^t S(t,s)\Phi^N(s)\mathrm{d}W(s)\Big)_{t\in[0,T]}$ have continuous modifications as well. Indeed, for all $t\in[0,T]$ we have

$$I^N(t) = \int_0^t \mathbb{1}_{(\!(0,\tau_N]\!]}(s)S(t,s)\Psi(s)\mathrm{d}s = \int_0^{t\wedge\tau_N} S(t,s)\Psi(s)\mathrm{d}s = S(t,t\wedge\tau_N)I(t\wedge\tau_N) \quad \mathbb{P}\text{-a.s.} \tag{3.3.15}$$

and

$$J^N(t) = \int_0^t \mathbb{1}_{(\!(0,\tau_N]\!]}(s)S(t,s)\Phi(s)\mathrm{d}W(s) = S(t,t\wedge\tau_N)J(t\wedge\tau_N) \quad \mathbb{P}\text{-a.s.,} \tag{3.3.16}$$

Note that, while the second equality in (3.3.15) is a consequence of [51, Lemma 2.3.9] and the third equality in (3.3.15) follows from standard properties of the Bochner integral, the third equality in (3.3.16) can be justified by approximating τ_N by a series of simple stopping times and verifying the equality for this simple stopping times first; see Lemma 3.5.10 below for details. The latter verification is necessary since the operator $S(t,t\wedge\tau_N)$ involves a stopping time and is thus a random operator. The processes $\big(S(t,t\wedge\tau_N)I(t\wedge\tau_N)\big)_{t\in[0,T]}$ and $\big(S(t,t\wedge\tau_N)J(t\wedge\tau_N)\big)_{t\in[0,T]}$ have continuous modifications by Lemma 3.5.1 in Section 3.5 below. We fix continuous modifications of the H-valued processes I^N and J^N, $N\in\mathbb{N}$, once and for all.
Moreover, for $N\in\mathbb{N}$ define the $C([0,T],H)$-valued process $X^{S,N} = (X_t^{S,N})_{t\in[0,T]}$ by

$$X_t^{S,N}(r) = \begin{cases} X^N(r) & \text{if } r\in[0,t), \\ S(r,t)X^N(t) & \text{if } r\in[t,T]. \end{cases}$$

Then, for all $N\in\mathbb{N}$ the assumptions of Theorem 3.2.2 are fulfilled by X^N, Ψ^N, Φ^N and $X^{S,N}$ in place of X, Ψ, Φ and X^S, respectively. Next note that

$$\mathbb{P}\Big(\mathbb{1}_{\{\|X(0)\|_H\leq N\}}X_t^{S,N} = \mathbb{1}_{\{\|X(0)\|_H\leq N\}}X_{t\wedge\tau_N}^S \quad \text{for all } t\in[0,T], N\in\mathbb{N}\Big) = 1. \tag{3.3.17}$$

Indeed, by [51, Lemma 2.3.9] it holds $\mathbb{P}$-a.s. for all t, $r \in [0,T]$ that

$$\begin{aligned} X_t^{S,N}(r) &= S(r,0)X^N(0) + \int_0^{r\wedge t} S(r,s)\Psi^N(s)\mathrm{d}s + \int_0^{r\wedge t} S(r,s)\Phi^N(s)\mathrm{d}W(s) \\ &= S(r,0)X^N(0) + \int_0^{r\wedge(t\wedge\tau_N)} S(r,s)\Psi(s)\mathrm{d}s + \int_0^{r\wedge(t\wedge\tau_N)} S(r,s)\Phi(s)\mathrm{d}W(s), \end{aligned}$$

Since all considered processes are continuous we obtain that, with probability one,

$$\mathbb{1}_{\{\|X(0)\|_H\le N\}} X_t^{S,N}(r) = \mathbb{1}_{\{\|X(0)\|_H\le N\}} X_{t\wedge\tau_N}^{S}(r) \quad \text{for all } t,r\in[0,T], N\in\mathbb{N},$$

which implies (3.3.17). Assume that we have shown statement (3.3.4) for the process $\big(X^N(t)\big)_{t\in[0,T]}$, i.e., for all $t\in[0,T]$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned} f(T,X_T^{S,N}) &= f(0,X_0^{S,N}) + \int_0^T \partial_1^+ f(s,X_s^{S,N})\mathrm{d}s \\ &\quad + \int_0^T \partial_2 f(s,X_s^{S,N})\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Psi^N(s)\big)\mathrm{d}s \\ &\quad + \int_0^T \partial_2 f(s,X_s^{S,N})\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi^N(s)\big)\mathrm{d}W(s) \\ &\quad + \frac{1}{2}\int_0^T \sum_{u\in\mathcal{U}} \partial_2^2 f(s,X_s^{S,N})\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi^N(s)u, \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi^N(s)u\big)\mathrm{d}s. \end{aligned} \tag{3.3.18}$$

We would like to prove the statement for the process $\big(X(t)\big)_{t\in[0,T]}$. Observe that

$$\begin{aligned} f(T,X_T^{S,N}) - f(0,X_0^{S,N}) &= \mathbb{1}_{\{\|X(0)\|_H\le N\}}\Big(f(T,X_{T\wedge\tau_N}^S) - f(0,X_0^S)\Big) \\ &\quad + \mathbb{1}_{\{\|X(0)\|_H> N\}}\Big(f(T,X_T^{S,N}) - f(0,0)\Big), \\ &\xrightarrow{N\to\infty} f(T,X_T^S) - f(0,X_0^S) \quad \mathbb{P}\text{-a.s.} \end{aligned} \tag{3.3.19}$$

due to (3.3.17), the fact that $\mathbb{P}\big(\{\|X(0)\|_H\le N\}\big) \xrightarrow{N\to\infty} 1$, $\tau_N \xrightarrow{N\to\infty} T$ $\mathbb{P}$-a.s. and that $f(T,\cdot)$ is continuous from $C([0,T],H)$ to V. For the same reason, we have

$$\begin{aligned} \int_0^T \partial_1^+ f(s,X_s^{S,N})\mathrm{d}s &= \mathbb{1}_{\{\|X(0)\|_H\le N\}}\int_0^T \partial_1^+ f(s,X_{s\wedge\tau_N}^S)\mathrm{d}s \\ &\quad + \mathbb{1}_{\{\|X(0)\|_H> N\}}\int_0^T \partial_1^+ f(s,X_s^{S,N})\mathrm{d}s \\ &\xrightarrow{N\to\infty} \int_0^T \partial_1^+ f(s,X_s^S)\mathrm{d}s \quad \mathbb{P}\text{-a.s.} \end{aligned} \tag{3.3.20}$$

where we also use the fact that $\sup_{s\in[0,T]}\big\|\partial_1^+ f(s,X_{s\wedge\tau_N}^S)\big\|_V < \infty$ $\mathbb{P}$-a.s. due to the boundedness of $\Big(X_{s\wedge\tau_N}^S(\cdot,\omega)\Big)_{s\in[0,T]}$, $N\in\mathbb{N}$, in $C([0,T],H)$ and the assumption that $\partial_1^+ f$ stays bounded on bounded subsets of $[0,T] \times C([0,T],H)$. Next, we use (3.3.17), the definition of $\Psi^N = \mathbb{1}_{(\!(0,\tau_N]\!]}\Psi$, the fact that $\mathbb{P}\big(\{\|X(0)\|_H \le N\}\big) \xrightarrow{N\to\infty} 1$, and $\tau_N \xrightarrow{N\to\infty} T$ $\mathbb{P}$-a.s. and the continuity of the Bochner-integral as a function of the upper limit, to obtain

$$\begin{aligned}
&\int_0^T \partial_2 f(s, X_s^{S,N})\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Psi^N(s)\big)\mathrm{d}s \\
&\qquad = \mathbb{1}_{\{\|X(0)\|_H \le N\}} \int_0^{T\wedge\tau_N} \partial_2 f(s, X_s^{S})\Big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Psi(s)\Big)\mathrm{d}s \\
&\qquad\quad + \mathbb{1}_{\{\|X(0)\|_H > N\}} \int_0^{T} \partial_2 f(s, X_s^{S,N})\Big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Psi^N(s)\Big)\mathrm{d}s \\
&\qquad \xrightarrow{N\to\infty} \int_0^T \partial_2 f(s, X_s^{S})\Big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Psi(s)\Big)\mathrm{d}s \quad \mathbb{P}\text{-a.s.}
\end{aligned} \tag{3.3.21}$$

Similarly, using also [51, Lemma 2.3.9] we obtain $\mathbb{P}$-a.s for all $t \in [0,T]$ that

$$\begin{aligned}
&\int_0^T \partial_2 f(s, X_s^{S,N})\Big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi^N(s)\Big)\mathrm{d}W(s) \\
&\qquad = \mathbb{1}_{\{\|X(0)\|_H \le N\}} \int_0^{T\wedge\tau_N} \partial_2 f(s, X_s^{S})\Big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)\Big)\mathrm{d}W(s) \\
&\qquad\quad + \mathbb{1}_{\{\|X(0)\|_H > N\}} \int_0^{T} \partial_2 f(s, X_s^{S,N})\Big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi^N(s)\Big)\mathrm{d}W(s)
\end{aligned}$$

Now, by choosing a continuous modification of the first stochastic integral on the right-side of the equation above and the continuity of the stochastic integral as a function of the upper limit, we have

$$\begin{aligned}
&\int_0^T \partial_2 f(s, X_s^{S,N})\Big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi^N(s)\Big)\mathrm{d}W(s) \\
&\qquad \xrightarrow{N\to\infty} \int_0^T \partial_2 f(s, X_s^{S})\Big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)\Big)\mathrm{d}W(s) \quad \mathbb{P}\text{-a.s.}
\end{aligned} \tag{3.3.22}$$

With the same argument as for (3.3.21) we finally obtain

$$\begin{aligned}
&\frac{1}{2}\int_0^T \sum_{u\in\mathcal{U}} \partial_2^2 f(s, X_s^{S,N})\Big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi^N(s)u, \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi^N(s)u\Big)\mathrm{d}s \\
&\xrightarrow{N\to\infty} \frac{1}{2}\int_0^T \sum_{u\in\mathcal{U}} \partial_2^2 f(s, X_s^{S})\Big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u, \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u\Big)\mathrm{d}s \quad \mathbb{P}\text{-a.s.}
\end{aligned} \tag{3.3.23}$$

The combination of (3.3.18), (3.3.19), (3.3.20), (3.3.21), (3.3.22), and (3.3.23) implies (3.3.4).

Step 3: From the previous step, we know that without loss of generality we can assume that there exists an $N \in \mathbb{N}$ such that inequality (3.3.5) holds. Under this assumption we obtain that

$$M = \sup_{s\in[0,T]} \max\Big(\|X_s^S\|_{C([0,T],H)}, \|\partial_1^+ f(s,X_s^S)\|_V, \|\partial_2 f(s,X_s^S)\|_{L(D([0,T],H),V)}, \|\partial_2^2 f(s,X_s^S)\|_{L^{(2)}(D([0,T],H),V)}, \sup_{0\le s\le t\le T} \|S(t,s)\|_{L(H)} \Big) < \infty. \tag{3.3.24}$$

We are going to verify formula (3.3.4), by making use of (3.3.5) and (3.3.24) and analyzing the terms appearing in (3.3.7) step by step. Without loss of generality we also assume that the Hilbert spaces U and H are infinite-dimensional and $(u_i)_{i\in\mathbb{N}}$ and $(h_i)_{i\in\mathbb{N}}$ are orthonormal bases of U and H, respectively. For each $n \in \mathbb{N}$ and $j \in \{0,1,\dots,2^n-1\}$, let $t_j^n = j\frac{T}{2^n}$, and Y_j^n and Γ_j^n be given as in (3.3.1) and (3.3.10), respectively. Furthermore, considering the boundedness property (3.3.24), we assume without loss of generality that

$$\sup_{\substack{j=1,\dots,2^n\\ n\in\mathbb{N}}} \sup_{s\in[t_{j-1}^n,t_j^n)} \max\Big(\|Y_j^n\|_{D([0,T],H)}, \|\partial_1^+ f(s,Y_j^n)\|_V, \|\partial_2 f(s,Y_j^n)\|_{L(D([0,T],H),V)}, \|\partial_2^2 f(s,Y_j^n)\|_{L^{(2)}(D([0,T],H),V)} \Big) < M. \tag{3.3.25}$$

Note that due to continuity of the functional f and by Lemma 3.3.1 it holds $\mathbb{P}$-a.s. that

$$A^n = f(T,X_T^S) - f(T,Y_n^n) \xrightarrow{n\to\infty} 0,$$
$$D^n = f(0,Y_0^n) - f(0,X_0^S) \xrightarrow{n\to\infty} 0.$$

For the term B^n in (3.3.7) note that it holds due to Lemma 3.5.3 below that

$$B^n = \sum_{j=1}^{2^n} \int_{t_{j-1}^n}^{t_j^n} \partial_1^+ f(s,Y_{j-1}^n)\mathrm{d}s = \int_0^T \sum_{j=1}^{2^n} \mathbb{1}_{[t_{j-1}^n,t_j^n)} \partial_1^+ f(s,Y_{j-1}^n)\mathrm{d}s.$$

Now observe that due to the continuity of $\partial_1^+ f$ and Lemma 3.3.1 it holds $\mathbb{P}$-a.s. for all $s\in[0,T]$ that

$$\sum_{j=1}^{2^n} \mathbb{1}_{[t_{j-1}^n,t_j^n)}(s)\partial_1^+ f(s,Y_{j-1}^n) \xrightarrow{n\to\infty} \partial_1^+ f(s,X_s^S).$$

By the boundedness property (3.3.25) and the dominated convergence theorem, we obtain $\mathbb{P}$-a.s. that

$$B^n \xrightarrow{n\to\infty} \int_0^T \partial_1^+ f\big(s,X_s^S\big)\mathrm{d}s.$$

Step 4: Next, we analyze the third term in (3.3.7), C^n. For that, we will analyze the terms $C_1^n,\dots,C_6^n$ appearing in (3.3.13) above in Steps 4.1, …, 4.6 below, respectively.

Step 4.1: Using the fact that for fixed $\omega \in \Omega$, $\partial_2 f(t_j^n, Y_{j-1}^n(\cdot,\omega))\Gamma_j^n$ is an element of $L(H,V)$ and $\int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Psi(s,\omega)\mathrm{d}s$ is an H-valued Bochner-integral, we have that

$$\begin{aligned} C_1^n &= \sum_{j=1}^{2^n} \partial_2 f(t_j^n, Y_{j-1}^n)\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Psi(s)\mathrm{d}s \\ &= \sum_{j=1}^{2^n} \int_{t_{j-1}^n}^{t_j^n} \partial_2 f(t_j^n, Y_{j-1}^n)\Gamma_j^n S(t_j^n, s)\Psi(s)\mathrm{d}s \\ &= \int_0^T \sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n, t_j^n]}(s)\partial_2 f(t_j^n, Y_{j-1}^n)\Big(\Gamma_j^n S(t_j^n, s)\Psi(s)\Big)\mathrm{d}s. \end{aligned}$$

An application of the triangular inequality implies that

$$\begin{aligned} &\int_0^T \Big\| \sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n, t_j^n]}(s)\partial_2 f\big(t_j^n, Y_{j-1}^n\big)\big(\Gamma_j^n S(t_j^n, s)\Psi(s)\big) - \partial_2 f\big(s, X_s^S\big)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Psi(s)\big)\Big\|_V \mathrm{d}s \\ &\le \int_0^T \sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n, t_j^n]}(s)\Big\{\Big\|\Big(\partial_2 f\big(t_j^n, Y_{j-1}^n\big) - \partial_2 f\big(s, X_s^S\big)\Big)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Psi(s)\big)\Big\|_V \\ &\quad + \Big\|\partial_2 f\big(t_j^n, Y_{j-1}^n\big)\Big(\Gamma_j^n S(t_j^n, s)\Psi(s) - \mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Psi(s)\Big)\Big\|_V\Big\}\mathrm{d}s. \end{aligned} \tag{3.3.26}$$

Now observe that (3.3.24) and (3.3.25) imply for all $j \in \{1, \dots, 2^n\}$, $n \in \mathbb{N}$ that

$$\Big\|\partial_2 f\big(t_j^n, Y_{j-1}^n\big)\Big(\Gamma_j^n S(t_j^n, s)\Psi(s) - \mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Psi(s)\Big)\Big\|_V \le 2M^2 \|\Psi(s)\|_H$$

and by Lemma 3.5.1 below, it holds $\mathbb{P}$-a.s. that

$$\big\|\Gamma_j^n S(t_j^n, s)\Psi(s) - \mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Psi(s)\big\|_{D([0,T],H)} \xrightarrow{n\to\infty} 0.$$

The dominated convergence theorem and the fact that $\int_0^T \|\Psi(s)\|_H \mathrm{d}s < \infty$ $\mathbb{P}$-a.s. thus yield $\mathbb{P}$-a.s. that

$$\int_0^T \sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n, t_j^n]}(s)\Big\|\partial_2 f\big(t_j^n, Y_{j-1}^n\big)\Big(\Gamma_j^n S(t_j^n, s)\Psi(s) - \mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Psi(s)\Big)\Big\|_V \mathrm{d}s \xrightarrow{n\to\infty} 0. \tag{3.3.27}$$

Moreover, Lemma 3.3.1 and Lemma 3.3.2 and a further application of the dominated convergence theorem imply that

$$\int_0^T \sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n, t_j^n]}(s)\Big\|\Big(\partial_2 f\big(t_j^n, Y_{j-1}^n\big) - \partial_2 f\big(s, X_s^S\big)\Big)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Psi(s)\big)\Big\|_V \mathrm{d}s \xrightarrow{n\to\infty} 0. \tag{3.3.28}$$

Combining (3.3.24)–(3.3.28) gives

$$C_1^n \xrightarrow{n\to\infty} \int_0^T \partial_2 f(s, X_s^S)\Big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Psi(s)\Big)\mathrm{d}s \quad \mathbb{P}\text{-a.s.}$$

Step 4.2: Next we consider the term

$$C_2^n = \sum_{j=1}^{2^n} \partial_2 f(t_j^n, Y_{j-1}^n)\left(\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s)\right).$$

Analogously to Lemma 3.5.8, one can show for each $j \in \{1,\dots,2^n\}$ that the mapping $[0,T]\times\Omega \ni (s,\omega) \mapsto \partial_2 f(t_j^n, Y_{j-1}^n(\cdot,\omega))\big(\Gamma_j^n S(t_j^n,s)\Phi(s,\omega)\big) \in \mathrm{HS}(U,V)$ is $\mathcal{P}_T$-$\mathcal{B}(\mathrm{HS}(U,V))$-measurable, and due to (3.3.5) and (3.3.25) we have that

$$\sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n,t_j^n]}(s)\partial_2 f(t_j^n, Y_{j-1}^n)\big(\Gamma_j^n S(t_j^n,s)\Phi(s)\big) \in L^2(\mathbb{P}\otimes \mathrm{d}s, \mathcal{P}_T; \mathrm{HS}(U,V)).$$

Moreover, due to the standard properties of the stochastic integral, we have that

$$C_2^n = \int_0^T \sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n,t_j^n]}(s)\partial_2 f(t_j^n, Y_{j-1}^n)\big(\Gamma_j^n S(t_j^n,s)\Phi(s)\big)\mathrm{d}W(s).$$

Basically by following a similar idea as in Step 4.1 we show that

$$\sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n,t_j^n]}(s)\partial_2 f(t_j^n, Y_{j-1}^n)\big(\Gamma_j^n S(t_j^n,s)\Phi(s)\big) \xrightarrow{n\to\infty} \partial_2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)\big), \quad (3.3.29)$$

in $L^2(\mathbb{P}\otimes \mathrm{d}s; \mathrm{HS}(U,V))$. For this we use the following triangular inequality to obtain pointwise convergence and then apply the dominated convergence theorem. It holds by the triangular inequality that

$$\begin{aligned}
&\Big\| \sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n,t_j^n]}(s)\partial_2 f(t_j^n, Y_{j-1}^n)\big(\Gamma_j^n S(t_j^n,s)\Phi(s)\big) - \partial_2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)\big)\Big\|_{\mathrm{HS}(U,V)} \\
&\quad \le \sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n,t_j^n]}(s) \left\|\big(\partial_2 f(t_j^n, Y_{j-1}^n) - \partial_2 f(s, X_s^S)\big)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)\big)\right\|_{\mathrm{HS}(U,V)} \\
&\quad\quad + \sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n,t_j^n]}(s) \left\|\partial_2 f(t_j^n, Y_{j-1}^n)\big(\Gamma_j^n S(t_j^n,s)\Phi(s) - \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)\big)\right\|_{\mathrm{HS}(U,V)}.
\end{aligned} \quad (3.3.30)$$

Next observe that (3.3.24) and (3.3.25) imply for all $j \in \{1,\dots,2^n\}$, $n \in \mathbb{N}$, that

$$\left\|\partial_2 f(t_j^n, Y_{j-1}^n)\big(\Gamma_j^n S(t_j^n,s)\Phi(s) - \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)\big)\right\|_{\mathrm{HS}(U,V)}^2 \le 4M^4 \|\Phi(s)\|_{\mathrm{HS}(U,H)}^2$$

and by Lemma 3.5.1 below it holds $\mathbb{P}$-a.s. for all $u \in \mathcal{U}$ that

$$\left\|\Gamma_j^n S(t_j^n, s)\Phi(s)u - \mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)u\right\|_{D([0,T],H)}^2 \xrightarrow{n\to\infty} 0.$$

The dominated convergence theorem and the fact that $\int_0^T \|\Phi(s)\|_{\mathrm{HS}(U,H)}^2 \,\mathrm{d}s < \infty$ thus imply that

$$\sum_{u\in\mathcal{U}} \left\|\Gamma_j^n S(t_j^n, s)\Phi(s)u - \mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)u\right\|_{D([0,T],H)}^2 \xrightarrow{n\to\infty} 0$$

and consequently that

$$\mathbb{E}\int_0^T \sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n,t_j^n]}(s) \left\|\partial_2 f(t_j^n, Y_{j-1}^n)\big(\Gamma_j^n S(t_j^n, s)\Phi(s) - \mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)\big)\right\|_{\mathrm{HS}(U,V)}^2 \mathrm{d}s \xrightarrow{n\to\infty} 0. \tag{3.3.31}$$

Moreover, combining Lemma 3.3.1, Lemma 3.3.2, the boundedness assumption (3.3.25), and the fact that

$$\sum_{u\in\mathcal{U}} \left\|\mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)u\right\|_{D([0,T],H)}^2 < M^2 \|\Phi(s)\|_{\mathrm{HS}(U,H)}^2 \in L^1(\mathbb{P}\otimes \mathrm{d}s),$$

allows us to conclude that

$$\mathbb{E}\int_0^T \sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n,t_j^n]}(s) \left\|\big(\partial_2 f(t_j^n, Y_{j-1}^n) - \partial_2 f(s, X_s^S)\big)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)\big)\right\|_{\mathrm{HS}(U,V)}^2 \mathrm{d}s \xrightarrow{n\to\infty} 0. \tag{3.3.32}$$

Thus, as a consequence of (3.3.30), (3.3.31) and (3.3.32) we have that

$$\mathbb{E}\int_0^T \left\|\sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n,t_j^n]}(s)\partial_2 f(t_j^n, Y_{j-1}^n)\big(\Gamma_j^n S(t_j^n, s)\Phi(s)\big) - \partial_2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)\big)\right\|_{\mathrm{HS}(U,V)}^2 \mathrm{d}s \xrightarrow{n\to\infty} 0.$$

Consequently we obtain the following convergence that

$$C_2^n \xrightarrow{n\to\infty} \int_0^T \partial_2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)\big)\,\mathrm{d}W(s) \quad \text{in probability.}$$

Step 4.3: The next term in (3.3.13) that we analyze is

$$C_3^n = \sum_{j=1}^{2^n} \frac{1}{2}\partial_2^2 f(t_j^n, Y_{j-1}^n)\left(\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Psi(s)\mathrm{d}s, \Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Psi(s)\mathrm{d}s\right).$$

A straightforward computation shows that

$$\begin{aligned}\|C_3^n\|_V &\le \frac{1}{2} \max_{\substack{j=1,\dots,2^n\\ n\in\mathbb{N}}} \left\|\partial_2^2 f(t_j^n, Y_{j-1}^n)\right\|_{L^{(2)}(D([0,T],H),V)} \sup_{0\le r\le\tau\le T} \|S(\tau,r)\|_{L(H)}^2 \\ &\quad\times \sum_{j=1}^{2^n} \left\| \int_{t_{j-1}^n}^{t_j^n} S(t_j^n,s)\Psi(s)\mathrm{d}s \right\|_H^2 \\ &\le M_1 \int_0^T \|\Psi(s)\|_H \,\mathrm{d}s \max_{j=1,\dots,2^n} \left\| \int_{t_{j-1}^n}^{t_j^n} S(t_j^n,s)\Psi(s)\mathrm{d}s \right\|_H \\ &\le M_1 \cdot N \max_{j=1,\dots,2^n} \int_{t_{j-1}^n}^{t_j^n} \sup_{0\le r\le\tau\le T} \|S(\tau,r)\|_{L(H)} \|\Psi(s)\|_H \,\mathrm{d}s \\ &\le M_1 \cdot N \cdot M \max_{j=1,\dots,2^n} \left\| \int_0^{t_j^n} \Psi(s)\mathrm{d}s - \int_0^{t_{j-1}^n} \Psi(s)\mathrm{d}s \right\|_H \xrightarrow{n\to\infty} 0 \quad \mathbb{P}\text{-a.s.}\end{aligned}$$

where M_1 is some positive constant independent of n, and N is from (3.3.5).

Step 4.4: The following equalities hold due to standard properties of the Bochner-integral:

$$\begin{aligned} C_4^n &= \sum_{j=1}^{2^n} \partial_2^2 f(t_j^n, Y_{j-1}^n)\left(\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n,s)\Psi(s)\mathrm{d}s, \Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n,s)\Phi(s)\mathrm{d}W(s)\right) \\ &= \sum_{j=1}^{2^n} \int_{t_{j-1}^n}^{t_j^n} \partial_2^2 f(t_j^n, Y_{j-1}^n)\left(\Gamma_j^n S(t_j^n,s)\Psi(s), \Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n,r)\Phi(r)\mathrm{d}W(r)\right)\mathrm{d}s \\ &= \int_0^T \sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n,t_j^n]}(s)\partial_2^2 f(t_j^n, Y_{j-1}^n)\left(\Gamma_j^n S(t_j^n,s)\Psi(s), \Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n,r)\Phi(r)\mathrm{d}W(r)\right)\mathrm{d}s. \end{aligned}$$

Therefore we have that

$$\|C_4^n\|_V \le M_1 \int_0^T \|\Psi(s)\|_H \,\mathrm{d}s \max_{j=1,\dots,2^n} \left\| \int_{t_{j-1}^n}^{t_j^n} S(t_j^n,r)\Phi(r)\mathrm{d}W(r) \right\|_H .$$

Here and below, M_1 denotes a finite constant which does not depend on n and may change its value with every appearance. Next note that

$$\begin{aligned} &\max_{j=1,\dots,2^n} \left\| \int_{t_{j-1}^n}^{t_j^n} S(t_j^n,r)\Phi(r)\mathrm{d}W(r) \right\|_H \\ &= \max_{j=1,\dots,2^n} \left\| \int_0^{t_j^n} S(t_j^n,r)\Phi(r)\mathrm{d}W(r) - S(t_j^n,t_{j-1}^n) \int_0^{t_{j-1}^n} S(t_{j-1}^n,r)\Phi(r)\mathrm{d}W(r) \right\|_H \xrightarrow{n\to\infty} 0 \quad \mathbb{P}\text{-a.s.} \end{aligned}$$

where the convergence holds due to the continuity of the stochastic integral as a function of the upper limit (see Assumption 3.1.1) and Lemma 3.5.1 below. Hence, it holds $\mathbb{P}$-a.s. that

$$C_4^n \xrightarrow{n\to\infty} 0.$$

Step 4.5: In this step we will use the conditional expectation of Hilbert space-valued random variables. Analogously to the real-valued case, one can show the existence of the conditional expectation; see Section 2.4. Remember that

$$C_5^n = \frac{1}{2}\sum_{j=1}^{2^n} \partial_2^2 f(t_j^n, Y_{j-1}^n)\Big(\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s), \Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s)\Big),$$

and that $(u_i)_{i\in\mathbb{N}}$ and $(h_i)_{i\in\mathbb{N}}$ are orthonormal bases of U and H, respectively. First we show that, for $j \in \{1, \ldots, 2^n\}$, it holds $\mathbb{P}$-a.s. that

$$\begin{aligned}
&\mathbb{E}\bigg(\partial_2^2 f(t_j^n, Y_{j-1}^n)\Big(\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s), \Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s)\Big)\bigg|\mathcal{F}_{t_{j-1}^n}\bigg) \\
&\quad = \mathbb{E}\bigg(\int_{t_{j-1}^n}^{t_j^n} \sum_{i=1}^{\infty} \partial_2^2 f(t_j^n, Y_{j-1}^n)\Big(\Gamma_j^n S(t_j^n, s)\Phi(s)u_i, \Gamma_j^n S(t_j^n, s)\Phi(s)u_i\Big)\mathrm{d}s\bigg|\mathcal{F}_{t_{j-1}^n}\bigg),
\end{aligned} \tag{3.3.33}$$

To this end, let first $j \in \{1, \ldots, 2^n\}$ be fixed. We know, by (2.4.4), that

$$\int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\,\mathrm{d}W(s) = \sum_{i=1}^{\infty} \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)u_i\,\mathrm{d}\beta_i(s), \tag{3.3.34}$$

where $(\beta_i(t))_{t\in[0,T]}$, $i \in \mathbb{N}$, are a family of independent real-valued Brownian motions and the infinite sum above converges in $L^2(\mathbb{P}; H)$. Consequently, we have that

$$\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s) = \lim_{N\to\infty} \Gamma_j^n \sum_{i=1}^{N} \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)u_i\mathrm{d}\beta_i(s) \quad \text{in } L^2\big(\mathbb{P}; D([0,T], H)\big). \tag{3.3.35}$$

The boundedness and bilinearity of $\partial_2^2 f(t_j^n, Y_{j-1}^n)$ thus imply that

$$\begin{aligned}
&\partial_2^2 f(t_j^n, Y_{j-1}^n)\Big(\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s), \Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s)\Big) \\
&\quad = \lim_{N\to\infty}\lim_{N'\to\infty} \sum_{i=1}^{N}\sum_{k=1}^{N'} \partial_2^2 f(t_j^n, Y_{j-1}^n)\Big(\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)u_i\,\mathrm{d}\beta_i(s), \\
&\qquad \Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)u_k\,\mathrm{d}\beta_k(s)\Big),
\end{aligned} \tag{3.3.36}$$

with convergence in $L^1(\mathbb{P}; V)$. Now the continuity of the mapping $\mathbb{E}\big(\cdot \big| \mathcal{F}_{t_{j-1}^n}\big) : L^1(\Omega, \mathcal{F}, \mathbb{P}; V) \to L^1(\Omega, \mathcal{F}_{t_{j-1}^n}, \mathbb{P}; V)$ hence implies that

$$\begin{aligned}
&\mathbb{E}\bigg(\partial_2^2 f(t_j^n, Y_{j-1}^n)\Big(\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s), \Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s)\Big)\Big|\mathcal{F}_{t_{j-1}^n}\bigg) \\
&\quad = \lim_{N\to\infty} \lim_{N'\to\infty} \sum_{i=1}^{N} \sum_{k=1}^{N'} \mathbb{E}\bigg(\partial_2^2 f(t_j^n, Y_{j-1}^n)\Big(\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)u_i \mathrm{d}\beta_i(s), \\
&\qquad \Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)u_k \mathrm{d}\beta_k(s)\Big)\Big|\mathcal{F}_{t_{j-1}^n}\bigg),
\end{aligned} \tag{3.3.37}$$

with convergence in $L^1(\mathbb{P}; V)$. To continue with our computation, note that for all $i, k \in \mathbb{N}$ it holds that

$$\begin{aligned}
&\partial_2^2 f(t_j^n, Y_{j-1}^n)\Big(\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)u_i \mathrm{d}\beta_i(s), \Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)u_k \mathrm{d}\beta_k(s)\Big) \\
&\quad = \partial_2^2 f(t_j^n, Y_{j-1}^n)\Big(\Gamma_j^n \sum_{\mu=1}^{\infty} \Big\langle \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)u_i \mathrm{d}\beta_i(s), h_\mu \Big\rangle_H h_\mu, \\
&\qquad \Gamma_j^n \sum_{\nu=1}^{\infty} \Big\langle \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)u_k \mathrm{d}\beta_k(s), h_\nu \Big\rangle_H h_\nu\Big).
\end{aligned} \tag{3.3.38}$$

With a similar argument as used in (3.3.37), one can show that

$$\begin{aligned}
&\mathbb{E}\bigg(\partial_2^2 f(t_j^n, Y_{j-1}^n)\Big(\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)u_i \mathrm{d}\beta_i(s), \Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)u_k \mathrm{d}\beta_k(s)\Big)\Big|\mathcal{F}_{t_{j-1}^n}\bigg) \\
&\quad = \lim_{M\to\infty} \lim_{M'\to\infty} \sum_{\mu=1}^{M} \sum_{\nu=1}^{M'} \mathbb{E}\bigg(\partial_2^2 f(t_j^n, Y_{j-1}^n)\Big(\Gamma_j^n \Big\langle \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)u_i \mathrm{d}\beta_i(s), h_\mu \Big\rangle_H h_\mu, \\
&\qquad \Gamma_j^n \Big\langle \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)u_k \mathrm{d}\beta_k(s), h_\nu \Big\rangle_H h_\nu\Big)\Big|\mathcal{F}_{t_{j-1}^n}\bigg),
\end{aligned} \tag{3.3.39}$$

with convergence in $L^1(\mathbb{P};V)$. Further, due to the $\mathcal{F}_{t^n_{j-1}}$-measurability of the V-valued random variables $\partial_2^2 f(t^n_j, Y^n_{j-1})(\Gamma^n_j h_\mu, \Gamma^n_j h_\nu)$, $\mu,\nu \in \mathbb{N}$, we obtain that

$$\begin{aligned}
&\mathbb{E}\bigg(\partial_2^2 f(t^n_j, Y^n_{j-1})\Big(\Gamma^n_j \left\langle \int_{t^n_{j-1}}^{t^n_j} S(t^n_j,s)\Phi(s)u_i \mathrm{d}\beta_i(s), h_\mu \right\rangle_H h_\mu, \\
&\qquad \Gamma^n_j \left\langle \int_{t^n_{j-1}}^{t^n_j} S(t^n_j,s)\Phi(s)u_k \mathrm{d}\beta_k(s), h_\nu \right\rangle_H h_\nu\Big)\Big|\mathcal{F}_{t^n_{j-1}}\bigg) \\
&= \partial_2^2 f(t^n_j, Y^n_{j-1})(\Gamma^n_j h_\mu, \Gamma^n_j h_\nu)\mathbb{E}\bigg(\left\langle \int_{t^n_{j-1}}^{t^n_j} S(t^n_j,s)\Phi(s)u_i \mathrm{d}\beta_i(s), h_\mu \right\rangle_H \\
&\qquad \cdot \left\langle \int_{t^n_{j-1}}^{t^n_j} S(t^n_j,s)\Phi(s)u_k \mathrm{d}\beta_k(s), h_\nu \right\rangle_H \Big|\mathcal{F}_{t^n_{j-1}}\bigg) \\
&= \partial_2^2 f(t^n_j, Y^n_{j-1})(\Gamma^n_j h_\mu, \Gamma^n_j h_\nu)\mathbb{E}\bigg(\int_{t^n_{j-1}}^{t^n_j} \big\langle S(t^n_j,s)\Phi(s)u_i, h_\mu\big\rangle_H \mathrm{d}\beta_i(s) \\
&\qquad \cdot \int_{t^n_{j-1}}^{t^n_j} \big\langle S(t^n_j,s)\Phi(s)u_k, h_\nu\big\rangle_H \mathrm{d}\beta_k(s)\Big|\mathcal{F}_{t^n_{j-1}}\bigg) \\
&= \partial_2^2 f(t^n_j, Y^n_{j-1})(\Gamma^n_j h_\mu, \Gamma^n_j h_\nu) \\
&\qquad \cdot \mathbb{E}\bigg(\int_{t^n_{j-1}}^{t^n_j} \big\langle S(t^n_j,s)\Phi(s)u_i, h_\mu\big\rangle_H \cdot \big\langle S(t^n_j,s)\Phi(s)u_k, h_\nu\big\rangle_H \mathrm{d}s\, \delta_{ik}\Big|\mathcal{F}_{t^n_{j-1}}\bigg),
\end{aligned} \tag{3.3.40}$$

where the last equality holds by standard properties of stochastic integrals; see, e.g., [33, Propostion 2.17], and δ_{ik} is the Kronecker delta function. Combining (3.3.37), (3.3.39), and (3.3.40), we obtain the following identities in $L^1(\mathbb{P};V)$,

$$\begin{aligned}
&\mathbb{E}\bigg(\partial_2^2 f(t^n_j, Y^n_{j-1})\Big(\Gamma^n_j \int_{t^n_{j-1}}^{t^n_j} S(t^n_j,s)\Phi(s)\mathrm{d}W(s), \Gamma^n_j \int_{t^n_{j-1}}^{t^n_j} S(t^n_j,s)\Phi(s)\mathrm{d}W(s)\Big)\Big|\mathcal{F}_{t^n_{j-1}}\bigg) \\
&= \lim_{N\to\infty}\lim_{N'\to\infty} \sum_{i=1}^{N}\sum_{k=1}^{N'}\sum_{\mu=1}^{\infty}\sum_{\nu=1}^{\infty} \partial_2^2 f(t^n_j, Y^n_{j-1})(\Gamma^n_j h_\mu, \Gamma^n_j h_\nu) \\
&\qquad \cdot \mathbb{E}\bigg(\int_{t^n_{j-1}}^{t^n_j} \big\langle S(t^n_j,s)\Phi(s)u_i, h_\mu\big\rangle_H \cdot \big\langle S(t^n_j,s)\Phi(s)u_k, h_\nu\big\rangle_H \mathrm{d}s\, \delta_{ik}\Big|\mathcal{F}_{t^n_{j-1}}\bigg) \\
&= \lim_{N\to\infty}\lim_{N'\to\infty} \sum_{i=1}^{\min(N,N')}\sum_{\mu=1}^{\infty}\sum_{\nu=1}^{\infty} \partial_2^2 f(t^n_j, Y^n_{j-1})(\Gamma^n_j h_\mu, \Gamma^n_j h_\nu) \\
&\qquad \cdot \mathbb{E}\bigg(\int_{t^n_{j-1}}^{t^n_j} \big\langle S(t^n_j,s)\Phi(s)u_i, h_\mu\big\rangle_H \cdot \big\langle S(t^n_j,s)\Phi(s)u_i, h_\nu\big\rangle_H \mathrm{d}s\,\Big|\mathcal{F}_{t^n_{j-1}}\bigg) \\
&= \lim_{N\to\infty} \sum_{i=1}^{N}\sum_{\mu=1}^{\infty}\sum_{\nu=1}^{\infty} \partial_2^2 f(t^n_j, Y^n_{j-1})(\Gamma^n_j h_\mu, \Gamma^n_j h_\nu) \\
&\qquad \cdot \mathbb{E}\bigg(\int_{t^n_{j-1}}^{t^n_j} \big\langle S(t^n_j,s)\Phi(s)u_i, h_\mu\big\rangle_H \cdot \big\langle S(t^n_j,s)\Phi(s)u_i, h_\nu\big\rangle_H \mathrm{d}s\,\Big|\mathcal{F}_{t^n_{j-1}}\bigg).
\end{aligned} \tag{3.3.41}$$

We continue with the computation above by using the linearity of $\partial_2^2 f(t^n_j, Y^n_{j-1})$, linearity and continuity of the mapping $\mathbb{E}(\,\cdot\,|\mathcal{F}_{t^n_{j-1}})\colon L^1(\Omega,\mathcal{F},\mathbb{P};V)\to L^1(\Omega,\mathcal{F}_{t^n_{j-1}},\mathbb{P};V)$, $\mathcal{F}_{t^n_{j-1}}$-measurability of $\partial_2^2 f(t^n_j, Y^n_{j-1})$, and the boundedness properties and obtain for every $i\in\mathbb{N}$ that

$$\begin{aligned}
&\sum_{\mu=1}^{\infty}\sum_{\nu=1}^{\infty}\partial_2^2 f(t^n_j,Y^n_{j-1})(\Gamma^n_j h_\mu,\Gamma^n_j h_\nu)\mathbb{E}\Big(\int_{t^n_{j-1}}^{t^n_j}\left\langle S(t^n_j,s)\Phi(s)u_i,h_\mu\right\rangle_H\cdot\left\langle S(t^n_j,s)\Phi(s)u_i,h_\nu\right\rangle_H\mathrm{d}s\,\Big|\mathcal{F}_{t^n_{j-1}}\Big)\\
&\quad=\sum_{\mu=1}^{\infty}\mathbb{E}\Big(\partial_2^2 f(t^n_j,Y^n_{j-1})\Big(\Gamma^n_j h_\mu,\Gamma^n_j\int_{t^n_{j-1}}^{t^n_j}\left\langle S(t^n_j,s)\Phi(s)u_i,h_\mu\right\rangle_H\\
&\qquad\cdot\sum_{\nu=1}^{\infty}\left\langle S(t^n_j,s)\Phi(s)u_i,h_\nu\right\rangle_H h_\nu\mathrm{d}s\Big)\,\Big|\mathcal{F}_{t^n_{j-1}}\Big)\\
&\quad=\sum_{\mu=1}^{\infty}\mathbb{E}\Big(\int_{t^n_{j-1}}^{t^n_j}\partial_2^2 f(t^n_j,Y^n_{j-1})\Big(\Gamma^n_j\left\langle S(t^n_j,s)\Phi(s)u_i,h_\mu\right\rangle_H h_\mu,\\
&\qquad\Gamma^n_j\sum_{\nu=1}^{\infty}\left\langle S(t^n_j,s)\Phi(s)u_i,h_\nu\right\rangle_H h_\nu\Big)\mathrm{d}s\,\Big|\mathcal{F}_{t^n_{j-1}}\Big)\\
&\quad=\mathbb{E}\Big(\int_{t^n_{j-1}}^{t^n_j}\partial_2^2 f(t^n_j,Y^n_{j-1})\Big(\Gamma^n_j\sum_{\mu=1}^{\infty}\left\langle S(t^n_j,s)\Phi(s)u_i,h_\mu\right\rangle_H h_\mu,\Gamma^n_j S(t^n_j,s)\Phi(s)u_i\Big)\mathrm{d}s\,\Big|\mathcal{F}_{t^n_{j-1}}\Big)\\
&\quad=\mathbb{E}\Big(\int_{t^n_{j-1}}^{t^n_j}\partial_2^2 f(t^n_j,Y^n_{j-1})\Big(\Gamma^n_j S(t^n_j,s)\Phi(s)u_i,\Gamma^n_j S(t^n_j,s)\Phi(s)u_i\Big)\mathrm{d}s\,\Big|\mathcal{F}_{t^n_{j-1}}\Big).
\end{aligned}\tag{3.3.42}$$

Combining (3.3.41) and (3.3.42) and using again the linearity and continuity of the mapping $\mathbb{E}(\,\cdot\,|\mathcal{F}_{t^n_{j-1}})\colon L^1(\Omega,\mathcal{F},\mathbb{P};V)\to L^1(\Omega,\mathcal{F}_{t^n_{j-1}},\mathbb{P};V)$ we get that

$$\begin{aligned}
&\mathbb{E}\Big(\partial_2^2 f(t^n_j,Y^n_{j-1})\Big(\Gamma^n_j\int_{t^n_{j-1}}^{t^n_j}S(t^n_j,s)\Phi(s)\mathrm{d}W(s),\Gamma^n_j\int_{t^n_{j-1}}^{t^n_j}S(t^n_j,s)\Phi(s)\mathrm{d}W(s)\Big)\Big|\mathcal{F}_{t^n_{j-1}}\Big)\\
&\quad=\lim_{N\to\infty}\sum_{i=1}^{N}\mathbb{E}\Big(\int_{t^n_{j-1}}^{t^n_j}\partial_2^2 f(t^n_j,Y^n_{j-1})\Big(\Gamma^n_j S(t^n_j,s)\Phi(s)u_i,\Gamma^n_j S(t^n_j,s)\Phi(s)u_i\Big)\mathrm{d}s\,\Big|\mathcal{F}_{t^n_{j-1}}\Big)\\
&\quad=\mathbb{E}\Big(\int_{t^n_{j-1}}^{t^n_j}\sum_{i=1}^{\infty}\partial_2^2 f(t^n_j,Y^n_{j-1})\Big(\Gamma^n_j S(t^n_j,s)\Phi(s)u_i,\Gamma^n_j S(t^n_j,s)\Phi(s)u_i\Big)\mathrm{d}s\,\Big|\mathcal{F}_{t^n_{j-1}}\Big).
\end{aligned}\tag{3.3.43}$$

This proves the equality in (3.3.33). On the other hand, we show that

$$\begin{aligned}
&\mathbb{E}\Big(\Big\|\sum_{j=1}^{2^n}\partial_2^2 f(t^n_j,Y^n_{j-1})\Big(\Gamma^n_j\int_{t^n_{j-1}}^{t^n_j}S(t^n_j,s)\Phi(s)\mathrm{d}W(s),\Gamma^n_j\int_{t^n_{j-1}}^{t^n_j}S(t^n_j,s)\Phi(s)\mathrm{d}W(s)\Big)\\
&\quad-\sum_{j=1}^{2^n}\mathbb{E}\Big(\partial_2^2 f(t^n_j,Y^n_{j-1})\Big(\Gamma^n_j\int_{t^n_{j-1}}^{t^n_j}S(t^n_j,s)\Phi(s)\mathrm{d}W(s),\Gamma^n_j\int_{t^n_{j-1}}^{t^n_j}S(t^n_j,s)\Phi(s)\mathrm{d}W(s)\Big)\Big|\mathcal{F}_{t^n_{j-1}}\Big)\Big\|_V^2\Big)\\
&\qquad\xrightarrow{n\to\infty}0.
\end{aligned}\tag{3.3.44}$$

To simplify notation, let $G_j^n = \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s)$, $j \in \{1, \ldots, 2^n\}$, $n \in \mathbb{N}$ and note that

$$\begin{aligned}
&\mathbb{E}\Bigg(\Bigg\| \sum_{j=1}^{2^n} \partial_2^2 f(t_j^n, Y_{j-1}^n)\Big(\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s), \Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s)\Big) \\
&\qquad - \sum_{j=1}^{2^n} \mathbb{E}\Big(\partial_2^2 f(t_j^n, Y_{j-1}^n)\Big(\Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s), \Gamma_j^n \int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s)\Big)\Big|\mathcal{F}_{t_{j-1}^n}\Big)\Bigg\|_V^2\Bigg) \\
&\quad = \sum_{j=1}^{2^n} \mathbb{E}\left(\left\|\partial_2^2 f(t_j^n, Y_{j-1}^n)\big(\Gamma_j^n G_j^n, \Gamma_j^n G_j^n\big) - \mathbb{E}\left(\partial_2^2 f(t_j^n, Y_{j-1}^n)\big(\Gamma_j^n G_j^n, \Gamma_j^n G_j^n\big)\Big|\mathcal{F}_{t_{j-1}^n}\right)\right\|_V^2\right) \\
&\qquad + \sum_{j \neq k} \mathbb{E}\Bigg(\Big\langle \partial_2^2 f(t_j^n, Y_{j-1}^n)\big(\Gamma_j^n G_j^n, \Gamma_j^n G_j^n\big) - \mathbb{E}\left(\partial_2^2 f(t_j^n, Y_{j-1}^n)\big(\Gamma_j^n G_j^n, \Gamma_j^n G_j^n\big)\Big|\mathcal{F}_{t_{j-1}^n}\right), \\
&\qquad\qquad \partial_2^2 f(t_k^n, Y_{k-1}^n)\big(\Gamma_k^n G_k^n, \Gamma_k^n G_k^n\big) - \mathbb{E}\left(\partial_2^2 f(t_k^n, Y_{k-1}^n)\big(\Gamma_k^n G_k^n, \Gamma_k^n G_k^n\big)\Big|\mathcal{F}_{t_{k-1}^n}\right)\Big\rangle_V\Bigg).
\end{aligned} \tag{3.3.45}$$

The second sum after the equality sign above is equal to zero due to simple properties of the conditional expectation. Indeed, let $k, j \in \mathbb{N}$, $k < j$ and observe that

$$\begin{aligned}
&\mathbb{E}\Big[\Big\langle \partial_2^2 f(t_j^n, Y_{j-1}^n)\Big(\Gamma_j^n G_j^n, \Gamma_j^n G_j^n\Big), \partial_2^2 f(t_k^n, Y_{k-1}^n)\Big(\Gamma_k^n G_k^n, \Gamma_k^n G_k^n\Big)\Big\rangle_V\Big] \\
&\quad = \mathbb{E}\Big[\mathbb{E}\big[\Big\langle \partial_2^2 f(t_j^n, Y_{j-1}^n)\Big(\Gamma_j^n G_j^n, \Gamma_j^n G_j^n\Big), \partial_2^2 f(t_k^n, Y_{k-1}^n)\Big(\Gamma_k^n G_k^n, \Gamma_k^n G_k^n\Big)\Big\rangle_V |\mathcal{F}_{t_{j-1}^n}\big]\Big] \\
&\quad = \mathbb{E}\Big[\Big\langle \mathbb{E}\big[\partial_2^2 f(t_j^n, Y_{j-1}^n)\Big(\Gamma_j^n G_j^n, \Gamma_j^n G_j^n\Big)|\mathcal{F}_{t_{j-1}^n}\big], \partial_2^2 f(t_k^n, Y_{k-1}^n)\Big(\Gamma_k^n G_k^n, \Gamma_k^n G_k^n\Big)\Big\rangle_V\Big],
\end{aligned}$$

and that

$$\begin{aligned}
&\mathbb{E}\Big[\Big\langle \mathbb{E}\big[\partial_2^2 f(t_k^n, Y_{k-1}^n)\Big(\Gamma_k^n G_k^n, \Gamma_k^n G_k^n\Big)|\mathcal{F}_{t_{k-1}^n}\big], \partial_2^2 f(t_j^n, Y_{j-1}^n)\Big(\Gamma_j^n G_j^n, \Gamma_j^n G_j^n\Big)\Big\rangle_V\Big] \\
&\quad = \mathbb{E}\Big[\Big\langle \mathbb{E}\big[\partial_2^2 f(t_j^n, Y_{j-1}^n)\Big(\Gamma_j^n G_j^n, \Gamma_j^n G_j^n\Big)|\mathcal{F}_{t_{j-1}^n}\big], \mathbb{E}\big[\partial_2^2 f(t_k^n, Y_{k-1}^n)\Big(\Gamma_k^n G_k^n, \Gamma_k^n G_k^n\Big)|\mathcal{F}_{t_{k-1}^n}\big]\Big\rangle_V\Big].
\end{aligned}$$

Moreover, note that for each $j \in \{1, \dots, 2^n\}$ it holds that

$$
\begin{aligned}
&\mathbb{E}\left(\left\|\partial_2^2 f(t_j^n, Y_{j-1}^n)(\Gamma_j^n G_j^n, \Gamma_j^n G_j^n) - \mathbb{E}\left(\partial_2^2 f(t_j^n, Y_{j-1}^n)(\Gamma_j^n G_j^n, \Gamma_j^n G_j^n)\middle|\mathcal{F}_{t_{j-1}^n}\right)\right\|_V^2\right) \\
&\quad \leq 4\,\mathbb{E}\left(\left\|\partial_2^2 f(t_j^n, Y_{j-1}^n)(\Gamma_j^n G_j^n, \Gamma_j^n G_j^n)\right\|_V^2\right) + 4\,\mathbb{E}\left(\left\|\mathbb{E}\left(\partial_2^2 f(t_j^n, Y_{j-1}^n)(\Gamma_j^n G_j^n, \Gamma_j^n G_j^n)\middle|\mathcal{F}_{t_{j-1}^n}\right)\right\|_V^2\right) \\
&\quad \leq 4 \max_{\substack{i=1,\dots,2^m \\ m\in\mathbb{N}}} \left\|\partial_2^2 f(t_i^m, Y_{i-1}^m)\right\|_{L^{(2)}(D([0,T],H),V)}^2 \sup_{0\leq r\leq\tau\leq T} \|S(\tau,r)\|_{L(H)}^4\,\mathbb{E}\left(\|G_j^n\|_H^4\right) \\
&\qquad + 4\,\mathbb{E}\left(\mathbb{E}\left(\left\|\partial_2^2 f(t_j^n, Y_{j-1}^n)(\Gamma_j^n G_j^n, \Gamma_j^n G_j^n)\right\|_V \middle|\mathcal{F}_{t_{j-1}^n}\right)^2\right) \\
&\quad \leq 8 \max_{\substack{i=1,\dots,2^m \\ m\in\mathbb{N}}} \left\|\partial_2^2 f(t_i^m, Y_{i-1}^m)\right\|_{L^{(2)}(D([0,T],H),V)}^2 \sup_{0\leq r\leq\tau\leq T} \|S(\tau,r)\|_{L(H)}^4\,\mathbb{E}\left(\|G_j^n\|_H^4\right) \\
&\quad \leq 8M^6\,\mathbb{E}\left(\|G_j^n\|_H^4\right),
\end{aligned}
\tag{3.3.46}
$$

where $M \in (0,\infty)$ is the constant defined in (3.3.24). By Lemma 3.1 in [23] there exists a $c \in (0,\infty)$ such that for all $j \in \{1,\dots,2^n\}$, $n \in \mathbb{N}$, it holds that

$$
\mathbb{E}\left[\left\|\int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s)\right\|_H^4\right] \leq c\,\mathbb{E}\left[\left(\int_{t_{j-1}^n}^{t_j^n} \|S(t_j^n, s)\Phi(s)\|_{\mathrm{HS}(U,H)}^2\,\mathrm{d}s\right)^2\right],
$$

which implies that

$$
\begin{aligned}
&\sum_{j=1}^{2^n}\mathbb{E}\left(\left\|\partial_2^2 f(t_j^n, Y_{j-1}^n)(\Gamma_j^n G_j^n, \Gamma_j^n G_j^n) - \mathbb{E}\left(\partial_2^2 f(t_j^n, Y_{j-1}^n)(\Gamma_j^n G_j^n, \Gamma_j^n G_j^n)\middle|\mathcal{F}_{t_{j-1}^n}\right)\right\|_V^2\right) \\
&\quad \leq 8M^6 \cdot c \cdot \sup_{0\leq r\leq\tau\leq T} \|S(\tau,r)\|_{L(H)}^4 \\
&\qquad \cdot\,\mathbb{E}\left[\max_{j=1,\dots,2^n}\int_{t_{j-1}^n}^{t_j^n} \|\Phi(s)\|_{\mathrm{HS}(U,H)}^2\mathrm{d}s \cdot \int_0^T \|\Phi(s)\|_{\mathrm{HS}(U,H)}^2\mathrm{d}s\right] \xrightarrow{n\to\infty} 0,
\end{aligned}
\tag{3.3.47}
$$

by the boundedness assumption (3.3.5) on $\int_0^T \|\Phi(s)\|_{\mathrm{HS}(U,H)}^2\mathrm{d}s$, the fact that $\max_{j=1,\dots,2^n}\int_{t_{j-1}^n}^{t_j^n} \|\Phi(s)\|_{\mathrm{HS}(U,H)}^2\mathrm{d}s \xrightarrow{n\to\infty} 0$, $\mathbb{P}$-a.s., and the dominated convergence theorem. Indeed, the uniform continuity of the mapping $[0,T] \ni t \mapsto \int_0^t \|\Phi(s)\|_{\mathrm{HS}(U,H)}^2\,\mathrm{d}s \in \mathbb{R}$ is a result of the continuity of the Lebesgue integral of an integrable function. Combining (3.3.33) and (3.3.44) we obtain that

$$
\begin{aligned}
&\mathbb{E}\Bigg(\Bigg\|\sum_{j=1}^{2^n}\partial_2^2 f(t_j^n, Y_{j-1}^n)\left(\Gamma_j^n\int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s), \Gamma_j^n\int_{t_{j-1}^n}^{t_j^n} S(t_j^n, s)\Phi(s)\mathrm{d}W(s)\right) \\
&\quad -\sum_{j=1}^{2^n}\mathbb{E}\Bigg(\int_{t_{j-1}^n}^{t_j^n}\sum_{i=1}^{\infty}\partial_2^2 f(t_j^n, Y_{j-1}^n)\left(\Gamma_j^n S(t_j^n, s)\Phi(s)u_i, \Gamma_j^n S(t_j^n, s)\Phi(s)u_i\right)\mathrm{d}s\Bigg|\mathcal{F}_{t_{j-1}^n}\Bigg)\Bigg\|_V^2\Bigg) \xrightarrow{n\to\infty} 0.
\end{aligned}
\tag{3.3.48}
$$

Let $s \in [0,T]$ and $(t^n_{j_n})_{n\geqslant 1}$ be such that $s \in (t^n_{j_n-1}, t^n_{j_n}]$, $n \in \mathbb{N}$. By using the triangular inequality and the bilinearity and symmetry of $\partial_2^2 f(t^n_{j_n}, Y^n_{j_n-1})(\cdot,\cdot)$, one can show for each $i \in \mathbb{N}$ that

$$\begin{aligned}
&\Big\| \partial_2^2 f(t^n_{j_n}, Y^n_{j_n-1})\big(\Gamma^n_{j_n} S(t^n_{j_n}, s)\Phi(s)u_i, \Gamma^n_{j_n} S(t^n_{j_n}, s)\Phi(s)u_i\big) \\
&\qquad - \partial_2^2 f(s, X^S_s)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u_i, \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u_i\big)\Big\|_V \\
&\quad \leq \Big\| \Big(\partial_2^2 f(t^n_{j_n}, Y^n_{j_n-1}) - \partial_2^2 f(s, X^S_s)\Big)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u_i, \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u_i\big)\Big\|_V \\
&\quad\quad + \Big\| \partial_2^2 f(t^n_{j_n}, Y^n_{j_n-1})\Big(\Gamma^n_{j_n} S(t^n_{j_n}, s)\Phi(s)u_i - \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u_i, \\
&\qquad\quad \Gamma^n_{j_n} S(t^n_{j_n}, s)\Phi(s)u_i + \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u_i\Big)\Big\|_V.
\end{aligned}$$

Now, by following a similar idea as in the proof of (3.3.29) we obtain that

$$\begin{aligned}
&\mathbb{E}\Bigg(\Big\| \int_0^T \sum_{j=1}^{2^n} \mathbb{1}_{(t^n_{j-1}, t^n_j]}(s) \sum_{i=1}^{\infty} \partial_2^2 f(t^n_j, Y^n_{j-1})\big(\Gamma^n_j S(t^n_j, s)\Phi(s)u_i, \Gamma^n_j S(t^n_j, s)\Phi(s)u_i\big)\mathrm{d}s \\
&\quad - \int_0^T \sum_{i=1}^{\infty} \partial_2^2 f(s, X^S_s)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u_i, \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u_i\big)\mathrm{d}s\Big\|_V\Bigg) \xrightarrow{n\to\infty} 0.
\end{aligned} \tag{3.3.49}$$

We use the above convergence to show that the following holds:

$$\begin{aligned}
&\mathbb{E}\Bigg(\Big\| \sum_{j=1}^{2^n} \mathbb{E}\Big(\int_{t^n_{j-1}}^{t^n_j} \sum_{i=1}^{\infty} \partial_2^2 f(t^n_j, Y^n_{j-1})\big(\Gamma^n_j S(t^n_j, s)\Phi(s)u_i, \Gamma^n_j S(t^n_j, s)\Phi(s)u_i\big)\mathrm{d}s \Big| \mathcal{F}_{t^n_{j-1}}\Big) \\
&\quad - \int_0^T \sum_{i=1}^{\infty} \partial_2^2 f(s, X^S_s)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u_i, \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u_i\big)\mathrm{d}s\Big\|_V\Bigg) \\
&\leq \mathbb{E}\Bigg(\Big\| \sum_{j=1}^{2^n} \mathbb{E}\Big(\int_{t^n_{j-1}}^{t^n_j} \sum_{i=1}^{\infty} \partial_2^2 f(t^n_j, Y^n_{j-1})\big(\Gamma^n_j S(t^n_j, s)\Phi(s)u_i, \Gamma^n_j S(t^n_j, s)\Phi(s)u_i\big)\mathrm{d}s \\
&\quad - \int_{t^n_{j-1}}^{t^n_j} \sum_{i=1}^{\infty} \partial_2^2 f(s, X^S_s)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u_i, \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u_i\big)\mathrm{d}s \Big| \mathcal{F}_{t^n_{j-1}}\Big)\Big\|_V\Bigg) \\
&\quad + \mathbb{E}\Bigg(\Big\| \int_0^T \sum_{i=1}^{\infty} \partial_2^2 f(s, X^S_s)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u_i, \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u_i\big)\mathrm{d}s \\
&\quad - \sum_{j=1}^{2^n} \mathbb{E}\Big(\int_{t^n_{j-1}}^{t^n_j} \sum_{i=1}^{\infty} \partial_2^2 f(s, X^S_s)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u_i, \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s)u_i\big)\mathrm{d}s \Big| \mathcal{F}_{t^n_{j-1}}\Big)\Big\|_V\Bigg).
\end{aligned} \tag{3.3.50}$$

An application of the Jensen inequality on the first summand after the inequality sign above and using (3.3.49) imply that

$$\begin{aligned}
&\mathbb{E}\Bigg(\bigg\| \sum_{j=1}^{2^n} \mathbb{E}\bigg(\int_{t_{j-1}^n}^{t_j^n} \sum_{i=1}^{\infty} \partial_2^2 f(t_j^n, Y_{j-1}^n)\big(\Gamma_j^n S(t_j^n, s)\Phi(s)u_i, \Gamma_j^n S(t_j^n, s)\Phi(s)u_i\big)\mathrm{d}s \\
&\qquad - \int_{t_{j-1}^n}^{t_j^n} \sum_{i=1}^{\infty} \partial_2^2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)u_i, \mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)u_i\big)\mathrm{d}s \bigg| \mathcal{F}_{t_{j-1}^n}\bigg)\bigg\|_V\Bigg) \\
&\leq \mathbb{E}\Bigg(\bigg\| \int_0^T \sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n, t_j^n]}(s) \sum_{i=1}^{\infty} \partial_2^2 f(t_j^n, Y_{j-1}^n)\big(\Gamma_j^n S(t_j^n, s)\Phi(s)u_i, \Gamma_j^n S(t_j^n, s)\Phi(s)u_i\big)\mathrm{d}s \\
&\qquad - \int_0^T \sum_{i=1}^{\infty} \partial_2^2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)u_i, \mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)u_i\big)\mathrm{d}s\bigg\|_V\Bigg) \xrightarrow{n\to\infty} 0.
\end{aligned} \tag{3.3.51}$$

For the second summand in (3.3.50), we use the Jensen inequality, the martingale convergence theorem, and the dominated convergence theorem to obtain that

$$\begin{aligned}
&\mathbb{E}\Bigg(\bigg\| \int_0^T \sum_{i=1}^{\infty} \partial_2^2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)u_i, \mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)u_i\big)\mathrm{d}s \\
&\qquad - \sum_{j=1}^{2^n} \mathbb{E}\bigg(\int_{t_{j-1}^n}^{t_j^n} \sum_{i=1}^{\infty} \partial_2^2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)u_i, \mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)u_i\big)\mathrm{d}s \bigg| \mathcal{F}_{t_{j-1}^n}\bigg)\bigg\|_V\Bigg) \\
&\leq \int_0^T \sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n, t_j^n]}(s) \mathbb{E}\Bigg(\bigg\| \sum_{i=1}^{\infty} \partial_2^2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)u_i, \mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)u_i\big) \\
&\qquad - \mathbb{E}\bigg(\sum_{i=1}^{\infty} \partial_2^2 f(s, X_s^S)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)u_i, \mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)u_i\big) \bigg| \mathcal{F}_{t_{j-1}^n}\bigg)\bigg\|_V\Bigg)\mathrm{d}s \\
&\quad \xrightarrow{n\to\infty} 0.
\end{aligned} \tag{3.3.52}$$

More details are given in Lemma 3.5.11 below. The combination of (3.3.48) and (3.3.50)–(3.3.52) implies that

$$C_5^n \xrightarrow{n\to\infty} \frac{1}{2}\int_0^T \sum_{i=1}^{\infty} \partial_2^2 f(s, X_s^S)\Big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)u_i, \mathbb{1}_{[s,T]}(\cdot)S(\cdot, s)\Phi(s)u_i\Big)\mathrm{d}s \quad \text{in } L^1(\mathbb{P}; \mathbb{R}),$$

and particularly in probability.

Step 4.6: Finally observe that

$$
\begin{aligned}
\mathbb{E}\Big(\|C_6^n\|_V\Big) &= \mathbb{E}\Bigg(\Bigg\|\sum_{j=1}^{2^n}\int_0^1 (1-\theta)\Big[\partial_2^2 f(t_j^n, Y_{j-1}^n+\theta\Delta Y_{j-1}^n)-\partial_2^2 f(t_j^n, Y_{j-1}^n)\Big](\Delta Y_{j-1}^n \Delta Y_{j-1}^n)\mathrm{d}\theta\Bigg\|\Bigg)\\
&\le \sum_{j=1}^{2^n}\mathbb{E}\Bigg(\int_0^1\Big\|\Big[\partial_2^2 f(t_j^n, Y_{j-1}^n+\theta\Delta Y_{j-1}^n)-\partial_2^2 f(t_j^n, Y_{j-1}^n)\Big](\Delta Y_{j-1}^n, \Delta Y_{j-1}^n)\Big\|_V \mathrm{d}\theta\Bigg)\\
&\le \sum_{j=1}^{2^n}\mathbb{E}\Bigg(\int_0^1\Big\|\partial_2^2 f(t_j^n, Y_{j-1}^n+\theta\Delta Y_{j-1}^n)-\partial_2^2 f(t_j^n, Y_{j-1}^n)\Big\|_{L^{(2)}(D([0,T],H),V)}\cdot\Big\|\Delta Y_{j-1}^n\Big\|^2_{D([0,T],H)}\mathrm{d}\theta\Bigg)\\
&\le \mathbb{E}\Bigg(\sup_{\substack{i=1,\dots,2^n\\ \theta\in[0,1]}}\Big\|\partial_2^2 f(t_i^n, Y_{i-1}^n+\theta\Delta Y_{i-1}^n)-\partial_2^2 f(t_i^n, Y_{i-1}^n)\Big\|_{L^{(2)}(D([0,T],H),V)}\cdot\sum_{j=1}^{2^n}\Big\|\Delta Y_{j-1}^n\Big\|^2_{D([0,T],H)}\Bigg).
\end{aligned}
\tag{3.3.53}
$$

Moreover, by Lemma 3.3.1 and Lemma 3.3.2 we have that

$$
\begin{aligned}
&\Big\|\partial_2^2 f(t_j^n, Y_{j-1}^n+\theta\Delta Y_{j-1}^n)-\partial_2^2 f(t_j^n, Y_{j-1}^n)\Big\|_{L^{(2)}(D([0,T],H),V)}\\
&\quad\le \Big\|\partial_2^2 f(t_j^n, X_{t_j^n}^S)-\partial_2^2 f(t_j^n, Y_{j-1}^n)\Big\|_{L^{(2)}(D([0,T],H),V)}\\
&\qquad+\Big\|\partial_2^2 f(t_j^n, Y_{j-1}^n+\theta\Delta Y_{j-1}^n)-\partial_2^2 f(t_j^n, X_{t_j^n}^S)\Big\|_{L^{(2)}(D([0,T],H),V)}\xrightarrow{n\to\infty} 0,\quad \mathbb{P}\text{-a.s.}
\end{aligned}
\tag{3.3.54}
$$

uniformly in $\theta\in[0,1]$ and $j=1,\dots,2^n$. Indeed, by an application of Lemma 3.3.2 we set $B=L^{(2)}(D([0,T],H),V)$, $H=H$, $f=\partial_2^2 f$, and $x_t=X_t^S$ in the notation of the lemma. Furthermore, note that

$$
\begin{aligned}
\mathbb{E}\Big(\Big\|\Delta Y_{j-1}^n\Big\|^2_{D([0,T],H)}\Big) &= \mathbb{E}\Bigg(\Bigg\|\Gamma_j^n\Bigg(\int_{t_{j-1}^n}^{t_j^n} S(t_j^n,s)\Psi(s)\mathrm{d}s+\int_{t_{j-1}^n}^{t_j^n} S(t_j^n,s)\Phi(s)\mathrm{d}W(s)\Bigg)\Bigg\|^2_{D([0,T],H)}\Bigg)\\
&\le 2\sup_{0\le r\le\tau\le T}\|S(\tau,r)\|^2_{L(H)}\,\mathbb{E}\Bigg(\Bigg\|\int_{t_{j-1}^n}^{t_j^n} S(t_j^n,s)\Psi(s)\mathrm{d}s\Bigg\|^2_H\\
&\qquad+\Bigg\|\int_{t_{j-1}^n}^{t_j^n} S(t_j^n,s)\Phi(s)\mathrm{d}W(s)\Bigg\|^2_H\Bigg)\\
&\le 2M^4\mathbb{E}\Bigg[\Bigg(\int_{t_{j-1}^n}^{t_j^n}\|\Psi(s)\|_H\mathrm{d}s\Bigg)^2\Bigg]+2M^4\mathbb{E}\Bigg(\int_{t_{j-1}^n}^{t_j^n}\|\Phi(s)\|^2_{\mathrm{HS}(U,H)}\mathrm{d}s\Bigg),
\end{aligned}
\tag{3.3.55}
$$

where for the last inequality we used the Jensen inequality and the Itô isometry and M is the constant defined in (3.3.24). Therefore for all $n\in\mathbb{N}$ we have that

$$
\begin{aligned}
\mathbb{E}\Bigg(\sum_{j=1}^{2^n}\Big\|\Delta Y_{j-1}^n\Big\|^2_{D([0,T],H)}\Bigg) &\le 2M^4\mathbb{E}\Bigg(\int_0^T\|\Psi(s)\|_H\mathrm{d}s\Bigg)\cdot\max_{j=1,\dots,2^n}\int_{t_{j-1}^n}^{t_j^n}\|\Psi(s)\|_H\mathrm{d}s\\
&\quad+2M^4\mathbb{E}\Bigg(\int_0^T\|\Phi(s)\|^2_{\mathrm{HS}(U,H)}\mathrm{d}s\Bigg)<\infty.
\end{aligned}
\tag{3.3.56}
$$

Indeed, the term $\max_{j=1,\dots,2^n} \int_{t_{j-1}^n}^{t_j^n} \|\Psi(s)\|_H \, \mathrm{d}s$ is finite due to the assumption (3.3.5). Combining (3.3.53), (3.3.54) and (3.3.56), we obtain that

$$C_6^n \xrightarrow{n\to\infty} 0,$$

in $L^1(\mathbb{P};\mathbb{R})$ and therefore in probability. □

3.4 Applications to weak error analysis for SEEs

3.4.1 Approximation of spatio-temporal covariances

In this section we mainly assume the setting in [13, Proposition 2.1] and analyze the convergence rate for approximations of spatio-temporal covariances of the form $\operatorname{Cov}\big(\langle X(t_1), h_1\rangle_H, \langle X(t_2), h_2\rangle_H\big)$, where $t_1, t_2 \in [0,T]$ and $h_1, h_2 \in H$. The employed approximations are based on Galerkin projections onto finite-dimensional subspaces of H.

Assumption 3.4.1. Let items (i) and (ii) in Assumption 3.1.1 be fulfilled. Moreover, let $(e_i)_{i\in\mathbb{N}} \subset H$, be an orthonormal basis of H, let $A\colon D(A) \subset H \to H$ be a diagonal linear operator as in Lemma 2.2.14 with eigenvalues satisfying $\sup((\lambda_i)_{i\in\mathbb{N}}) < 0$, let $\big(H_r, \langle\cdot,\cdot\rangle_{H_r}, \|\cdot\|_{H_r}\big)$, $r \in \mathbb{R}$, be a family of interpolation spaces associated to $-A$ (see Definition 2.2.10), let $F \in \operatorname{Lip}\big(H, H\big)$, and let $B \in \operatorname{Lip}\big(H, \operatorname{HS}(U,H)\big)$. Suppose that $\xi \in H_\rho$ with $\rho \in [0,1)$, and let also $(P_N)_{N\in\mathbb{N}} \subset L(H)$ be defined as $P_N(v) = \sum_{n=1}^N \langle e_n, v\rangle_H e_n$ for all $v \in H$ and $N \in \mathbb{N}$. Note that for $x \in D([0,T], H)$ and $N \in \mathbb{N}$, by $P_N x$ we mean the path defined by

$$\big(P_N x\big)(s) = P_N(x(s)) \qquad s \in [0,T].$$

The following existence and uniqueness result is well-known; see, e.g., [31, Theorem 5.1].

Lemma 3.4.2. Let Assumption 3.4.1 be fulfilled. Then there exists a unique (up to modifications) predictable stochastic process $X\colon \Omega \times [0,T] \to H$ satisfying $\sup_{t\in[0,T]} \mathbb{E}\Big[\|X(t)\|_H^2\Big] < \infty$ and

$$X(t) = e^{tA}\xi + \int_0^t e^{(t-s)A} F\big(X(s)\big)\mathrm{d}s + \int_0^t e^{(t-s)A} B\big(X(s)\big)\mathrm{d}W(s), \tag{3.4.1}$$

$\mathbb{P}$-a.s. for all $t \in [0,T]$.

Remark 3.4.3. Note that under above assumption, the conditions in Assumption 3.1.1 are fulfilled with $S(t,s) = e^{(t-s)A}$, $\Psi(s) = F(X(s))$, $\Phi(s) = B(X(s))$ for $s,t \in [0,T]$ in the notation of Assumption 3.1.1. Indeed, the path-wise continuity of X follows from [42, Lemma 6.2.9] and [42, Proposition 6.3.5].

For $N \in \mathbb{N}$, observe that it holds $\mathbb{P}$-a.s. for all $t \in [0,T]$ that $P_N X(t) = e^{tA} P_N \xi + \int_0^t e^{(t-s)A} P_N F\big(X(s)\big)\mathrm{d}s + \int_0^t e^{(t-s)A} P_N B\big(X(s)\big)\mathrm{d}W(s)$. Note that for all $t \in [0,T]$ and $N \in \mathbb{N}$ the operators P_N and e^{tA} can commute due to the fact that A is a diagonal operator.

Lemma 3.4.4. Let Assumption 3.4.1 be fulfilled, let $X\colon \Omega \times [0,T] \to H$ be given by Lemma 3.4.2 with continuous sample paths, and let $N \in \mathbb{N}$. Then, for every t_1, $t_2 \in [0,T]$ it holds that

$$\Big|\mathbb{E}\Big[\big\langle X(t_1), X(t_2)\big\rangle_H - \big\langle P_N X(t_1), P_N X(t_2)\big\rangle_H\Big]\Big| \leq M \left\|\mathrm{Id} - P_N\right\|_{L(H,H_{-\rho})},$$

where

$$\begin{aligned} M &= 2\Big(\sup_{r\in[0,T]} \left\|e^{rA}\right\|_{L(H)}\Big)^2 \left\|\xi\right\|_{H_\rho}^2 \\ &\quad + \sup_{s\in[0,T]} \mathbb{E}\left[\left\|X(s)\right\|_H^2\right] \bigg(\big(2 \sup_{r\in[0,T]} \left\|e^{rA}\right\|_{L(H)} + 5\big)\frac{T^{1-\rho}}{1-\rho} \left\|F\right\|_{\mathrm{Lip}(H,H)} \\ &\quad + 2 \sup_{r\in[0,T]} \left\|e^{rA}\right\|_{L(H)} \frac{T^{1-\rho}}{1-\rho} \left\|B\right\|_{\mathrm{Lip}(H,\mathrm{HS}(U,H))}^2 \bigg). \end{aligned}$$

Note that the term $\left\|\mathrm{Id} - P_N\right\|_{L(H,H_{-\rho})}$ in the inequality above is finite. In particular, if $(\lambda_i)_{i\in\mathbb{N}}$ is a decreasing sequence then we have that

$$\left\|\mathrm{Id} - P_N\right\|_{L(H,H_{-\rho})} = (-\lambda_{N+1})^{-\rho}.$$

Proof of Lemma 3.4.4. Let t_1, $t_2 \in [0,T]$, with $t_1 < t_2$ be fixed and the mapping $f\colon D([0,T],H) \to \mathbb{R}$ be given as $f(x) = \langle x(t_1), x(t_2)\rangle_H$, for all $x \in D([0,T],H)$. Then f is two times Fréchet differentiable with continuous derivatives and therefore Assumption 3.1.5 is fulfilled. Indeed it holds for all $x, y, z \in D([0,T],H)$ that

$$\begin{aligned} f'(x)(y) &= \langle y(t_1), x(t_2)\rangle_H + \langle x(t_1), y(t_2)\rangle_H, \\ f''(x)(y,z) &= \langle y(t_1), z(t_2)\rangle_H + \langle z(t_1), y(t_2)\rangle_H. \end{aligned}$$

We denote by $\mathcal{U}$ the orthonormal basis of U. By Lemma 3.4.2 and Remark 3.4.3, the assumption of Theorem 3.2.2 is fulfilled and we can apply our path-dependent mild Itô formula and obtain that

$$\begin{aligned} &\mathbb{E}\left[f(X_T^S)\right] - \mathbb{E}\left[f(P_N X_T^S)\right] = \mathbb{E}\left[f(X_0^S)\right] - \mathbb{E}\left[f(P_N X_0^S)\right] \\ &\quad + \int_0^T \bigg(\mathbb{E}\Big[f'(X_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}F(X(t))\Big)\Big] - \mathbb{E}\Big[f'(P_N X_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}P_N F(X(t))\Big)\Big] \bigg)\mathrm{d}t \\ &\quad + \frac{1}{2}\int_0^T \mathbb{E}\Big[\sum_{u\in\mathcal{U}} f''(X_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}B\big(X(t)\big)u, \mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}B\big(X(t)\big)u\Big)\Big]\mathrm{d}t \\ &\quad - \frac{1}{2}\int_0^T \mathbb{E}\Big[\sum_{u\in\mathcal{U}} f''(P_N X_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}P_N B\big(X(t)\big)u, \mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}P_N B\big(X(t)\big)u\Big)\Big]\mathrm{d}t. \end{aligned} \tag{3.4.2}$$

Next observe that due to the definition of X_0^S it holds that

$$\begin{aligned}\left|f(X_0^S) - f(P_N X_0^S)\right| &= \left|\left\langle e^{t_1 A}\big(\operatorname{Id} - P_N\big)\xi, e^{t_2 A}\xi\right\rangle_H + \left\langle e^{t_1 A} P_N \xi, e^{t_2 A}\big(\operatorname{Id} - P_N\big)\xi\right\rangle_H\right| \\ &\leq \left\|e^{t_1 A}(-A)^{-\rho}\big(\operatorname{Id} - P_N\big)(-A)^{\rho}\xi\right\|_H \cdot \left\|e^{t_2 A}\xi\right\|_H \\ &\quad + \left\|e^{t_1 A} P_N \xi\right\|_H \cdot \left\|e^{t_2 A}(-A)^{-\rho}\big(\operatorname{Id} - P_N\big)(-A)^{\rho}\xi\right\|_H \\ &\leq 2\Big(\sup_{r\in[0,T]} \left\|e^{rA}\right\|_{L(H)}\Big)^2 \|\xi\|_{H_\rho}^2 \, \|\operatorname{Id} - P_N\|_{L(H,H_{-\rho})}.\end{aligned}$$

Therefore we have that

$$\mathbb{E}\left[\left|f(X_0^S) - f(P_N X_0^S)\right|\right] \leq 2\Big(\sup_{r\in[0,T]} \left\|e^{rA}\right\|_{L(H)}\Big)^2 \|\xi\|_{H_\rho}^2 \, \|\operatorname{Id} - P_N\|_{L(H,H_{-\rho})}. \tag{3.4.3}$$

Next observe that for $t \in [0,T]$ it holds that

$$\begin{aligned}&\left|f'(X_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}F(X(t))\Big) - f'(P_N X_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}P_N F(X(t))\Big)\right| \\ &\quad \leq \left|\Big(f'(X_t^S) - f'(P_N X_t^S)\Big)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}F(X(t))\Big)\right| \\ &\qquad + \left|f'(P_N X_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}F(X(t)) - \mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}P_N F(X(t))\Big)\right|.\end{aligned}$$

Before we continue with the above computation, we need to write the term $f'(X_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}F(X(t))\Big)$ explicitly. Observe that

$$\begin{aligned}&f'(X_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}F(X(t))\Big) \\ &\quad = \begin{cases}\Big\langle e^{(t_1-t)A}F(X(t)),\, e^{(t_2-t)A}X(t)\Big\rangle_H + \Big\langle e^{(t_2-t)A}F(X(t)),\, e^{(t_1-t)A}X(t)\Big\rangle_H & t\in[0,t_1] \\ \Big\langle e^{(t_2-t)A}F(X(t))\,,X(t_1)\Big\rangle_H & t\in(t_1,t_2] \\ 0 & t\in(t_2,T].\end{cases}\end{aligned}$$

Thus it holds that

$$\begin{aligned}&\Big(f'(X_t^S) - f'(P_N X_t^S)\Big)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}F(X(t))\Big) \\ &= \begin{cases}\Big\langle e^{(t_1-t)A}F(X(t)),\, e^{(t_2-t)A}\big(\operatorname{Id} - P_N\big)X(t)\Big\rangle_H + \Big\langle e^{(t_2-t)A}F(X(t)),\, e^{(t_1-t)A}\big(\operatorname{Id} - P_N\big)X(t)\Big\rangle_H & t\in[0,t_1] \\ \Big\langle e^{(t_2-t)A}F(X(t))\,,\big(\operatorname{Id} - P_N\big)X(t_1)\Big\rangle_H = \Big\langle e^{(\frac{t_2-t}{2})A}F(X(t)),\, e^{(\frac{t_2-t}{2})A}\big(\operatorname{Id} - P_N\big)X(t_1)\Big\rangle_H & t\in(t_1,t_2] \\ 0 & t\in(t_2,T],\end{cases}\end{aligned}$$

and that

$$\begin{aligned}&f'(P_N X_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}F(X(t)) - \mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}P_N F(X(t))\Big) \\ &= \begin{cases}\Big\langle e^{(t_1-t)A}\big(\operatorname{Id} - P_N\big)F(X(t)),\, e^{(t_2-t)A}P_N X(t)\Big\rangle_H + \Big\langle e^{(t_2-t)A}\big(\operatorname{Id} - P_N\big)F(X(t)), e^{(t_1-t)A}P_N X(t)\Big\rangle_H & t\in[0,t_1] \\ \Big\langle e^{(t_2-t)A}\big(\operatorname{Id} - P_N\big)F(X(t)), P_N X(t_1)\Big\rangle_H & t\in(t_1,t_2] \\ 0 & t\in(t_2,T].\end{cases}\end{aligned}$$

Consequently, we obtain by straightforward computations that

$$\begin{aligned}
&\left|\int_0^T \mathbb{E}\left[f'(X_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}F(X(t))\Big)\right] - \mathbb{E}\left[f'(P_N X_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}P_N F(X(t))\Big)\right]\mathrm{d}t\right| \\
&\quad\le \sup_{s\in[0,T]} \mathbb{E}\left[\|X(s)\|_H^2\right] \|F\|_{\mathrm{Lip}(H,H)} \|\mathrm{Id}-P_N\|_{L(H,H_{-\rho})} \\
&\qquad\cdot\left(2\big(1+\sup_{r\in[0,T]}\|e^{rA}\|_{L(H)}\big)\int_0^{t_1}(t_1-t)^{-\rho}\mathrm{d}t + \big(2^\rho+1\big)\int_{t_1}^{t_2}(t_2-t)^{-\rho}\mathrm{d}t\right) \\
&\quad\le \sup_{s\in[0,T]} \mathbb{E}\left[\|X(s)\|_H^2\right] \|F\|_{\mathrm{Lip}(H,H)} \|\mathrm{Id}-P_N\|_{L(H,H_{-\rho})} \\
&\qquad\cdot\left(2\big(1+\sup_{r\in[0,T]}\|e^{rA}\|_{L(H)}\big)\int_0^{T}(T-t)^{-\rho}\mathrm{d}t + \big(2+1\big)\int_{0}^{T}(T-t)^{-\rho}\mathrm{d}t\right) \\
&\quad\le \sup_{s\in[0,T]} \mathbb{E}\left[\|X(s)\|_H^2\right] \|F\|_{\mathrm{Lip}(H,H)} \frac{T^{1-\rho}}{1-\rho}\big(2\sup_{r\in[0,T]}\|e^{rA}\|_{L(H)}+5\big)\|\mathrm{Id}-P_N\|_{L(H,H_{-\rho})}.
\end{aligned} \tag{3.4.4}$$

Next note that for each $u\in\mathcal{U}$ it holds that

$$\begin{aligned}
&f''(X_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}B\big(X(t)\big)u, \mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}B\big(X(t)\big)u\Big) \\
&\quad= \begin{cases} 2\left\langle e^{(t_1-t)A}B\big(X(t)\big)u,\, e^{(t_2-t)A}B\big(X(t)\big)u\right\rangle_H & t\le t_1 \\ 0 & \text{otherwise.}\end{cases}
\end{aligned}$$

Therefore, we have that

$$\begin{aligned}
&\bigg|\frac{1}{2}\int_0^T \mathbb{E}\Big[\sum_{u\in\mathcal{U}} f''(X_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}B\big(X(t)\big)u, \mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}B\big(X(t)\big)u\Big)\Big]\mathrm{d}t \\
&\qquad - \frac{1}{2}\int_0^T \mathbb{E}\Big[\sum_{u\in\mathcal{U}} f''(P_N X_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}P_N B\big(X(t)\big)u, \mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}P_N B\big(X(t)\big)u\Big)\Big]\mathrm{d}t\bigg| \\
&\quad\le \int_0^{t_1}\mathbb{E}\Big[\sum_{u\in\mathcal{U}}\Big\langle e^{(t_1-t)A}\big(\mathrm{Id}-P_N\big)B\big(X(t)\big)u,\, e^{(t_2-t)A}B\big(X(t)\big)u\Big\rangle_H\Big] \\
&\qquad + \int_0^{t_1}\mathbb{E}\Big[\sum_{u\in\mathcal{U}}\Big\langle e^{(t_1-t)A}P_N B\big(X(t)\big)u,\, e^{(t_2-t)A}\big(\mathrm{Id}-P_N\big)B\big(X(t)\big)u\Big\rangle_H\Big] \\
&\quad\le \sup_{s\in[0,T]}\mathbb{E}\left[\|X(s)\|_H^2\right]\|B\|^2_{\mathrm{Lip}(H,\mathrm{HS}(U,H))}\|\mathrm{Id}-P_N\|_{L(H,H_{-\rho})}\int_0^{t_1}\Big((t_1-t)^{-\rho}+(t_2-t)^{-\rho}\Big)\mathrm{d}t \\
&\quad\le 2\sup_{s\in[0,T]}\mathbb{E}\left[\|X(s)\|_H^2\right]\|B\|^2_{\mathrm{Lip}(H,\mathrm{HS}(U,H))}\|\mathrm{Id}-P_N\|_{L(H,H_{-\rho})}\cdot\int_0^{t_1}(t_1-t)^{-\rho}\,\mathrm{d}t \\
&\quad\le 2\sup_{s\in[0,T]}\mathbb{E}\left[\|X(s)\|_H^2\right]\|B\|^2_{\mathrm{Lip}(H,\mathrm{HS}(U,H))}\|\mathrm{Id}-P_N\|_{L(H,H_{-\rho})}\cdot\frac{T^{1-\rho}}{1-\rho}.
\end{aligned} \tag{3.4.5}$$

The combination of (3.4.2), (3.4.3), (3.4.4) and (3.4.5) proves the assertion. □

In the following result we show the weak error rate for time approximations of covariances of mild solution of a linear SPDE, i.e., we assume the Assumtion 3.4.1 and set the drift term $F = 0$. Then $X\colon \Omega \times [0,T] \to H$ satisfies $\sup_{t\in[0,T]} \mathbb{E}\left[\|X(t)\|_H^2\right] < \infty$ and

$$X(t) = e^{tA}\xi + \int_0^t e^{(t-s)A} B\big(X(s)\big) \mathrm{d}W(s), \tag{3.4.6}$$

$\mathbb{P}$-a.s. for all $t \in [0,T]$.

Lemma 3.4.5. Let Assumption 3.4.1 be fulfilled with $F = 0$, let $X\colon \Omega \times [0,T] \to H$ with continuous sample paths be given by (3.4.6), and let $N \in \mathbb{N}$. Then, for every t_1, $t_2 \in [0,T]$ and $h_1, h_2 \in H$ it holds that

$$\begin{aligned}&\Big|\mathrm{Cov}\left(\big\langle X(t_1), h_1\big\rangle_H, \big\langle X(t_2), h_2\big\rangle_H\right) - \mathrm{Cov}\left(\big\langle P_N X(t_1), h_1\big\rangle_H, \big\langle P_N X(t_2), h_2\big\rangle_H\right)\Big| \\ &\quad \le M \left\|\mathrm{Id} - P_N\right\|_{L(H,H_{-\rho})},\end{aligned}$$

where

$$\begin{aligned}M = \|h_1\|_H \, \|h_2\|_H \Bigg(&\Big(\sup_{r\in[0,T]} \big\|e^{rA}\big\|_{L(H)}\Big)^2 \|\xi\|_{H_\rho}^2 \\ &+ 2 \sup_{s\in[0,T]} \mathbb{E}\left[\|X(s)\|_H^2\right] \frac{T^{1-\rho}}{1-\rho} \, \|B\|_{\mathrm{Lip}(H,\mathrm{HS}(U,H))}^2\Bigg) < \infty.\end{aligned}$$

Proof. Let t_1, $t_2 \in [0,T]$, with $t_1 < t_2$ and $h_1, h_2 \in H$ be fixed and let the mapping $f\colon D([0,T],H) \to \mathbb{R}$ be given as $f(x) = \langle x(t_1), h_1\rangle_H \cdot \langle x(t_2), h_2\rangle_H$, for all $x \in D([0,T],H)$. Then f is two times Fréchet differentiable with continuous derivatives and therefore Assumption 3.1.5 is fulfilled. Indeed it holds for all $x, y, z \in D([0,T],H)$ that

$$\begin{aligned} f'(x)(y) &= \langle y(t_1), h_1\rangle_H \cdot \langle x(t_2), h_2\rangle_H + \langle y(t_2), h_2\rangle_H \cdot \langle x(t_1), h_1\rangle_H\,, \quad \text{and} \\ f''(x)(y,z) &= \langle y(t_1), h_1\rangle_H \cdot \langle z(t_2), h_2\rangle_H + \langle y(t_2), h_2\rangle_H \cdot \langle z(t_1), h_1\rangle_H\,. \end{aligned}$$

We denote by $\mathcal{U}$ the orthonormal basis of U. By Theorem 3.4.2 and Remark 3.4.3, the assumption of Theorem 3.2.2 is fulfilled and we can apply the path-dependent mild Itô formula and obtain that

$$\begin{aligned}&\mathbb{E}\left[f(X_T^S)\right] - \mathbb{E}\left[f(P_N X_T^S)\right] = \mathbb{E}\left[f(X_0^S)\right] - \mathbb{E}\left[f(P_N X_0^S)\right] \\ &\quad + \frac{1}{2}\int_0^T \mathbb{E}\Big[\sum_{u\in\mathcal{U}} f''(X_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}B\big(X(t)\big)u, \mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}B\big(X(t)\big)u\Big)\Big]\mathrm{d}t \\ &\quad - \frac{1}{2}\int_0^T \mathbb{E}\Big[\sum_{u\in\mathcal{U}} f''(P_N X_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}P_N B\big(X(t)\big)u, \mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}P_N B\big(X(t)\big)u\Big)\Big]\mathrm{d}t.\end{aligned} \tag{3.4.7}$$

Next observe that due to the definition of X_0^S it holds that

$$\begin{aligned}\left|f(X_0^S)-f(P_NX_0^S)\right| &\le \left|\left\langle e^{t_1A}\big(\mathrm{Id}-P_N\big)\xi,h_1\right\rangle_H\cdot\left\langle e^{t_2A}\xi,h_2\right\rangle_H\right|\\ &\quad+\left|\left\langle e^{t_2A}\big(\mathrm{Id}-P_N\big)\xi,h_2\right\rangle_H\cdot\left\langle e^{t_1A}P_N\xi,h_1\right\rangle_H\right|\\ &\le \|h_1\|_H\ \|h_2\|_H\left(\sup_{r\in[0,T]}\left\|e^{rA}\right\|_{L(H)}\right)^2\|\xi\|_{H_\rho}^2\ \|\mathrm{Id}-P_N\|_{L(H,H_{-\rho})}.\end{aligned}$$

Therefore we have that

$$\mathbb{E}\left[\left|f(X_0^S)-f(P_NX_0^S)\right|\right]\le\|h_1\|_H\ \|h_2\|_H\left(\sup_{r\in[0,T]}\left\|e^{rA}\right\|_{L(H)}\right)^2\|\xi\|_{H_\rho}^2\ \|\mathrm{Id}-P_N\|_{L(H,H_{-\rho})}. \tag{3.4.8}$$

Next note that for each $u\in\mathcal{U}$ it holds that

$$\begin{aligned}&f''(X_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}B\big(X(t)\big)u,\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}B\big(X(t)\big)u\Big)\\ &=\begin{cases}2\left\langle e^{(t_1-t)A}B\big(X(t)\big)u,\ h_1\right\rangle_H\cdot\left\langle e^{(t_2-t)A}B\big(X(t)\big)u,\ h_2\right\rangle_H & t\le t_1\\ 0, & \text{otherwise.}\end{cases}\end{aligned}$$

Therefore we have that

$$\begin{aligned}&\Bigg|\frac{1}{2}\int_0^T\mathbb{E}\Big[\sum_{u\in\mathcal{U}}f''(X_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}B\big(X(t)\big)u,\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}B\big(X(t)\big)u\Big)\Big]\mathrm{d}t\\ &\quad-\frac{1}{2}\int_0^T\mathbb{E}\Big[\sum_{u\in\mathcal{U}}f''(P_NX_t^S)\Big(\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}P_NB\big(X(t)\big)u,\mathbb{1}_{[t,T]}(\cdot)e^{(\cdot-t)A}P_NB\big(X(t)\big)u\Big)\Big]\mathrm{d}t\Bigg|\\ &\le\int_0^{t_1}\mathbb{E}\Big[\sum_{u\in\mathcal{U}}\Big\langle e^{(t_1-t)A}\big(\mathrm{Id}-P_N\big)B\big(X(t)\big)u,\ h_1\Big\rangle_H\cdot\Big\langle e^{(t_2-t)A}B\big(X(t)\big)u,\ h_2\Big\rangle_H\Big]\mathrm{d}t\\ &\quad+\int_0^{t_1}\mathbb{E}\Big[\sum_{u\in\mathcal{U}}\Big\langle e^{(t_1-t)A}P_NB\big(X(t)\big)u,\ h_1\Big\rangle_H\cdot\Big\langle e^{(t_2-t)A}\big(\mathrm{Id}-P_N\big)B\big(X(t)\big)u,\ h_2\Big\rangle_H\Big]\mathrm{d}t\\ &\le\|h_1\|_H\ \|h_2\|_H\sup_{s\in[0,T]}\mathbb{E}\left[\|X(s)\|_H^2\right]\|B\|_{\mathrm{Lip}(H,\mathrm{HS}(U,H))}^2\ \|\mathrm{Id}-P_N\|_{L(H,H_{-\rho})}\\ &\quad\cdot\int_0^{t_1}\Big((t_1-t)^{-\rho}+(t_2-t)^{-\rho}\Big)\mathrm{d}t\\ &\le 2\,\|h_1\|_H\ \|h_2\|_H\sup_{s\in[0,T]}\mathbb{E}\left[\|X(s)\|_H^2\right]\|B\|_{\mathrm{Lip}(H,\mathrm{HS}(U,H))}^2\ \|\mathrm{Id}-P_N\|_{L(H,H_{-\rho})}\\ &\quad\cdot\int_0^{t_1}(t_1-t)^{-\rho}\mathrm{d}t\\ &\le 2\,\|h_1\|_H\ \|h_2\|_H\sup_{s\in[0,T]}\mathbb{E}\left[\|X(s)\|_H^2\right]\|B\|_{\mathrm{Lip}(H,\mathrm{HS}(U,H))}^2\ \|\mathrm{Id}-P_N\|_{L(H,H_{-\rho})}\cdot\frac{T^{1-\rho}}{1-\rho}.\end{aligned} \tag{3.4.9}$$

The combination of (3.4.7), (3.4.8) and (3.4.9) proves the assertion. □

3.4.2 Outlook on further applications

In this section, we discuss a possible further application of our path-dependent mild Itô formula (3.2.4) to linear SPDEs. For this, let μ be a finite Borel-measure, let $A\colon D(A) \subset H \to H$ be the generator of a strongly continuous semigroup $(e^{tA})_{t\geq 0} \subset L(H)$, and let $B \in \mathrm{HS}(U,H)$. Moreover, let Assumption 3.1.1 be fulfilled, where the strongly continuous evolution family S is given by $S(t,s) = e^{(t-s)A}$, $0 \leq s \leq t \leq T$, $\Psi(s) = 0$, and $\Phi(s) = B$, $s \in [0,T]$. Consider the mild solution

$$X(t) = e^{tA}\xi + \int_0^t e^{(t-s)A} B \,\mathrm{d}W(s) \qquad t \in [0,T],$$

of the linear SPDE

$$\begin{cases} \mathrm{d}X(t) = AX(t)\,\mathrm{d}t + B\,\mathrm{d}W(t), & t \in [0,T] \\ X(0) = \xi \in H. \end{cases}$$

The following approach to weak error analysis for path-dependent functionals is analogous to the approach in [38] in the non-path-dependent case. To simplify the exposition, we consider here a simplified $L^2(\mu;H)$-setting; as in Remark 3.2.6, however, we expect that the following weak error representation can be extended to the $D([0,T],H)$-setting.
As in Remark 3.2.5, we define the strongly continuous evolution family $(S_{t,s})_{0\leq s\leq t\leq T}$ on $L^2(\mu;H)$ and the $L^2(\mu;H)$-valued mild Itô process $(X_t)_{t\in[0,T]}$ by

$$\begin{aligned} S_{t,s}x &= \mathbb{1}_{[0,s)}(\cdot)x(\cdot) + \mathbb{1}_{[s,T]}(\cdot)e^{(\cdot\wedge t-s)A}x(\cdot) \qquad \forall x \in L^2(\mu;H), \\ X_t &= X(\cdot \wedge t), \end{aligned}$$

so that

$$X_t = S_{t,0}\xi + \int_0^t S_{t,s}\big(\mathbb{1}_{[s,T]}(\cdot)B\big)\,\mathrm{d}W(s), \tag{3.4.10}$$

where we interpret ξ as a constant H-valued path. Next let $f \in C_b^2\big(L^2(\mu;H),\mathbb{R}\big)$ and define the mapping $\phi\colon [0,T] \times L^2(\mu;H) \to \mathbb{R}$, $(t,x) \mapsto \phi(t,x)$, by

$$\phi(t,x) = \mathbb{E}\Big[f\Big(x + \int_t^T S_{T,s}\big(\mathbb{1}_{[s,T]}(\cdot)B\big)\,\mathrm{d}W(s)\Big)\Big]. \tag{3.4.11}$$

Then it is not difficult to check that for all $t \in [0,T]$, $x,y,z \in L^2(\mu;H)$ it holds that

$$\begin{aligned} \partial_2\phi(t,x)(y) &= \mathbb{E}\Big[f'\Big(x + \int_t^T S_{T,s}\big(\mathbb{1}_{[s,T]}(\cdot)B\big)\,\mathrm{d}W(s)\Big)(y)\Big], \\ \partial_2^2\phi(t,x)(y,z) &= \mathbb{E}\Big[f''\Big(x + \int_t^T S_{T,s}\big(\mathbb{1}_{[s,T]}(\cdot)B\big)\,\mathrm{d}W(s)\Big)(y,z)\Big]. \end{aligned} \tag{3.4.12}$$

Moreover, the following backward Kolmogorov equation holds for $(t,x) \in [0,T) \times L^2(\mu;H)$

$$\begin{cases} \partial_1^+ \phi(t,x) = -\frac{1}{2}\sum_{u\in\mathcal{U}} \partial_2^2\phi(t,x)\Big(S_{T,t}\big(\mathbb{1}_{[t,T]}(\cdot)Bu\big),\, S_{T,t}\big(\mathbb{1}_{[t,T]}(\cdot)Bu\big)\Big), \\ \phi(T,x) = f(x). \end{cases} \tag{3.4.13}$$

To see that (3.4.13) holds true, note that the $L^2(\mu;H)$-valued random variables $\int_t^T S_{T,s}\big(\mathbb{1}_{[s,T]}(\cdot)B\big)\,\mathrm{d}W(s)$ and $\int_0^{T-t} S_{T,T-s}\big(\mathbb{1}_{[T-s,T]}(\cdot)B\big)\,\mathrm{d}W(s)$ have the same distribution, so that

$$\phi(t,x) = \mathbb{E}\Big[f\Big(x + \int_0^{T-t} S_{T,T-s}\big(\mathbb{1}_{[T-s,T]}(\cdot)B\big)\,\mathrm{d}W(s)\Big)\Big]. \tag{3.4.14}$$

Now we can apply the Itô formula [16, Theorem 4.32] to the function $y \mapsto f(x+y)$, with fixed x, and the $L^2(\mu;H)$-valued process $\Big(\int_0^r S_{T,T-s}\big(\mathbb{1}_{[T-s,T]}(\cdot)B\big)\,\mathrm{d}W(s)\Big)_{r\in[0,T]}$, at time $T-t$, and take expectations on both sides of the Itô formula to obtain that

$$\begin{aligned} \mathbb{E}\Big[f\Big(x + \int_0^{T-t} S_{T,T-s}\big(\mathbb{1}_{[T-s,T]}(\cdot)B\big)\,\mathrm{d}W(s)\Big)\Big] &= f(x) \\ + \frac{1}{2}\int_0^{T-t} \sum_{u\in\mathcal{U}} \mathbb{E}\Big[f''\Big(x + \int_0^s S_{T,T-r}\big(\mathbb{1}_{[T-r,T]}(\cdot)B\big)\,\mathrm{d}W(r)\Big) \\ \Big(S_{T,T-s}\big(\mathbb{1}_{[T-s,T]}(\cdot)Bu\big),\, S_{T,T-s}\big(\mathbb{1}_{[T-s,T]}(\cdot)Bu\big)\Big)\Big]\,\mathrm{d}s. \end{aligned}$$

Now use (3.4.14) and (3.4.12) to get

$$\phi(t,x) = \phi(T,x) + \frac{1}{2}\int_0^{T-t} \sum_{u\in\mathcal{U}} \partial_2^2\phi(T-s,x)\Big(S_{T,T-s}\big(\mathbb{1}_{[T-s,T]}(\cdot)Bu\big),\, S_{T,T-s}\big(\mathbb{1}_{[T-s,T]}(\cdot)Bu\big)\Big)\mathrm{d}s. \tag{3.4.15}$$

Taking a right-sided time-derivative from both sides of (3.4.15) implies the backward Kolmogorov equation (3.4.13).

To derive a weak error representation, let $(\tilde{X}(t))_{t\in[0,T]}$ be the mild solution of another linear SPDE with initial value $\tilde{\xi}$ and diffusion operator $\tilde{B} \in \mathrm{HS}(U,H)$ and observe that

$$\begin{aligned} \mathbb{E}\big[f(\tilde{X}_T)\big] - E\big[f(X_T)\big] &= \mathbb{E}\big[\phi(T,\tilde{X}_T)\big] - \phi(0,S_{T,0}\xi) \\ &= \big(\phi(0,S_{T,0}\tilde{\xi}) - \phi(0,S_{T,0}\xi)\big) + \mathbb{E}\big[\phi(T,\tilde{X}_T) - \phi(0,S_{T,0}\tilde{\xi})\big], \end{aligned}$$

and assuming sufficient regularity of ϕ apply the mild Itô formula to the $L^2(\mu;H)$-valued mild Itô process $(\tilde{X}_t)_{t\in[0,T]}$ and the function $\phi(t,x)$. This leads to a weak error representation analogous to [38, Theorem 3.1] involving $\partial_2^2\phi(t,x)$, and the weak error can then be suitably estimated. In the $D([0,T],H)$-setting, we expect that our path-dependent mild Itô formula (3.2.4) can be applied in the procedure explained above, wherever an Itô formula was used, i.e., once to derive the backward Kolmogorov equation and once to estimate the weak error.

3.5 Technical proofs

In this section we collect some technical results mostly used in the proof of Theorem 3.2.2. Readers who are not interested in these technicalities are welcome to skip this section.

Lemma 3.5.1. Let $T \in (0,\infty)$, let $(H, \langle\cdot,\cdot\rangle_H, \|\cdot\|_H)$ be a separable real Hilbert space, let $S = (S(t,s))_{0\leq s\leq t\leq T}$ be a strongly continuous evolution family on H as in Assumption 3.1.1 (iii) and let $x \in C([0,T],H)$. Then it holds that

$$\big\{(t,s) \in [0,T]^2\colon s \leq t\big\} \ni (t,s) \mapsto S(t,s)x(s) \in H$$

is uniformly continuous.

Proof. Since $\big\{(t,s) \in [0,T]^2\colon s \leq t\big\}$ is a compact set, it suffices to show the continuity at an arbitrary fixed $(t_0,s_0) \in \big\{(t,s) \in [0,T]^2\colon s \leq t\big\}$. For this note that it holds for all $(\tau,\sigma) \in \big\{(t,s) \in [0,T]^2\colon s \leq t\big\}$ and $x \in C([0,T],H)$ that

$$\begin{aligned}
\|S(\tau,\sigma)x(s) - S(t_0,s_0)x(s_0)\|_H &\leq \|S(\tau,\sigma)x(s) - S(\tau,\sigma)x(s_0)\|_H \\
&\quad + \|S(\tau,\sigma)x(s_0) - S(t_0,s_0)x(s_0)\|_H \\
&\leq \sup_{0\leq u\leq v\leq T} \|S(v,u)\|_{L(H)} \|x(\sigma) - x(s_0)\|_H \\
&\quad + \big\|\big(S(\tau,\sigma) - S(t_0,s_0)\big)x(s_0)\big\|_H .
\end{aligned}$$

If $\sigma \to s_0$, the first term on the right-hand side goes to zero since $x \in C([0,T],H)$. Moreover, the second term goes to zero as $(\tau,\sigma) \to (t_0,s_0)$, by the continuity assumption on $S = (S(\tau,\sigma))_{0\leq s\leq t\leq T}$. □

Lemma 3.5.2. Let $(V, \|\cdot\|_V)$ be a real Banach space, let $a,b \in \mathbb{R}$ with $a < b$ and, let $g\colon [a,b] \to V$ be continuous and right-differentiable with $\partial^+ g|_{(a,b)} = 0$. Then g is constant on $[a,b]$.

Proof. Assume that the assertion does not hold. Then there exist $r,s \in (a,b)$ with $r < s$ and $g(r) \neq g(s)$. We define

$$\begin{aligned}
\varepsilon &= \frac{\|g(s) - g(r)\|_V}{2(s-r)} > 0, \quad \text{and} \\
c &= \inf\Big\{x \in (r,s]\colon \|g(x) - g(r)\|_V > \varepsilon(x-r)\Big\}.
\end{aligned}$$

The definitions above and the continuity of g imply that $\|g(c) - g(r)\|_V = \varepsilon(c-r)$ and $c < s$. Due to the assumption that $\partial^+ g(c) = 0$, it holds that

$$\exists d \in (c,s] \quad s.t. \quad \|g(x) - g(c)\|_V < \varepsilon(x-c) \qquad \forall x \in (c,d],$$

and consequently that

$$\|g(x) - g(r)\|_V \leq \|g(x) - g(c)\|_V + \|g(c) - g(r)\|_V \leq \varepsilon(x - r) \qquad \forall x \in (c, d],$$

which contradicts the definition of c. □

Lemma 3.5.3. Let $(V, \|\cdot\|_V)$ be a real Banach space, let $a, b \in \mathbb{R}$ with $a < b$ and let $g\colon [a, b] \to V$ be continuous and right-differentiable with continuous right-derivative $\partial^+ g$. Then it holds that

$$g(b) - g(a) = \int_a^b \partial^+ g(x) \mathrm{d}x.$$

Proof. Define $G(x) = \int_a^x \partial^+ g(s) \mathrm{d}s$, then by the fundamental theorem of calculus, [42, Proposition A.2.3], we have for all $x \in (a, b)$ that

$$G'(x) = \partial^+ g(x).$$

In particular $\partial^+(G - g)|_{(a,b)} = 0$. Now Lemma 3.5.2 completes the proof. □

Lemma 3.5.4. Assume the setting in Example 3.1.6 (ii). Then it holds that

$$\partial_2 f(t, x)(h) = \int_{[0,T]} \partial_2 g\big(t, s, x(s)\big)\big(h(s)\big)\mu(\mathrm{d}s).$$

Proof. By the first-order Taylor formula [58, Theorem 4.C] , we have that

$$g(t, s, r + h) = g(t, s, r) + \partial_2 g(t, s, r)(h) + \int_0^1 (1 - \theta)\Big(\partial_2 g(t, s, r + \theta h) - \partial_2 g(t, s, r)\Big)(h)\mathrm{d}\theta.$$

Then we obtain

$$\begin{aligned} f(t, x + h) &= \int_{[0,T]} g\big(t, s, x(s) + h(s)\big)\mu(\mathrm{d}s) \\ &= \int_{[0,T]} g\big(t, s, x(s)\big)\mu(\mathrm{d}s) + \int_0^T \partial_2 g\big(t, s, x(s)\big)\big(h(s)\big)\mu(\mathrm{d}s) \\ &\quad + \int_{[0,T]} \int_0^1 (1 - \theta)\Big[\partial_2 g\big(t, s, x(s) + \theta h(s)\big) - \partial_2 g\big(t, s, x(s) + h(s)\big)\Big] h(s)\mathrm{d}\theta\mu(\mathrm{d}s) \\ &= f(t, x) + \partial_2 f(t, x)(h) \\ &\quad + \int_{[0,T]} \int_0^1 (1 - \theta)\Big[\partial_2 g\big(t, s, x(s) + \theta h(s)\big) - \partial_2 g\big(t, s, x(s) + h(s)\big)\Big] h(s)\mathrm{d}\theta\mu(\mathrm{d}s), \end{aligned}$$

since

$$\begin{aligned}
&\Big\| \int_{[0,T]} \int_0^1 (1-\theta)\Big[\partial_2 g\big(t,s,x(s)+\theta h(s)\big) - \partial_2 g\big(t,s,x(s)+h(s)\big)\Big] h(s)\mathrm{d}\theta\mu(\mathrm{d}s)\Big\|_V \\
&\leq \|h\|_{D([0,T],H)} \int_{[0,T]} \int_0^1 \Big\|\partial_2 g\big(t,s,x(s)+\theta h(s)\big) - \partial_2 g\big(t,s,x(s)+h(s)\big)\Big\|_{L(H,V)} \mathrm{d}\theta\mu(\mathrm{d}s),
\end{aligned}$$

and therefore

$$\begin{aligned}
&\lim_{h\to 0} \frac{\Big\| \int_{[0,T]} \int_0^1 (1-\theta)\Big[\partial_2 g\big(t,s,x(s)+\theta h(s)\big) - \partial_2 g\big(t,s,x(s)+h(s)\big)\Big] h(s)\mathrm{d}\theta\mu(\mathrm{d}s)\Big\|_V}{\|h\|_{D([0,T],H)}} \\
&\leq \lim_{h\to 0} \int_{[0,T]} \int_0^1 \Big\|\partial_2 g\big(t,s,x(s)+\theta h(s)\big) - \partial_2 g\big(t,s,x(s)+h(s)\big)\Big\|_{L(H,V)} \mathrm{d}\theta\mu(\mathrm{d}s) \\
&= \int_{[0,T]} \int_0^1 \lim_{h\to 0} \Big\|\partial_2 g\big(t,s,x(s)+\theta h(s)\big) - \partial_2 g\big(t,s,x(s)+h(s)\big)\Big\|_{L(H,V)} \mathrm{d}\theta\mu(\mathrm{d}s) \\
&= 0
\end{aligned}$$

as $\partial_2 g$ is continuous in x and dominated by the μ-integrable function G. □

Measurability of integrands in the functional Itô formula

Lemma 3.5.5. Let $(E, \|\cdot\|_E)$ and $(K, \|\cdot\|_K)$ be real Banach spaces, let $(\Omega, \mathcal{A})$ be a measurable space, let $f\colon \Omega \to E$ be $\mathcal{A}$-$\mathcal{B}(E)$-measurable, let $F\colon \Omega \to L(E,K)$ be such that for all $e \in E$ the mapping $F(\cdot)e\colon \Omega \to K$ is $\mathcal{A}$-$\mathcal{B}(K)$-measurable, and assume further that at least one of the following conditions is fulfilled:

(i) $f(\Omega) \subset E$ is separable,

(ii) F is $\mathcal{A}$-$\mathcal{B}(L(E,K))$-measurable and $F(\Omega) \subset L(E,K)$ is separable.

Then the mapping $F(\cdot)f(\cdot)\colon \Omega \to K$, $\omega \mapsto F(\omega)f(\omega)$, is $\mathcal{A}$-$\mathcal{B}(K)$-measurable.

Proof. We first verify the assertion under assumption (i). In this case, let $f_n\colon \Omega \to E$, $n \in \mathbb{N}$, be a sequence of simple E-valued functions such that for every $\omega \in \Omega$ it holds that $\lim_{n\to\infty} f_n(\omega) = f(\omega)$, see [51, Lemma A.1.4]. Observe that for every $\omega \in \Omega$ we have

$$F(\omega)f(\omega) = \lim_{n\to\infty} F(\omega)f_n(\omega) \tag{3.5.1}$$

as an equality in K. Moreover, note that for every $n \in \mathbb{N}$ the mapping

$$F(\cdot)f_n(\cdot)\colon \Omega \to K,\ \omega \mapsto F(\omega)f_n(\omega) \tag{3.5.2}$$

is $\mathcal{A}$-$\mathcal{B}(K)$-measurable. Indeed, in order to verify it suffices to observe that for every $B \in \mathcal{B}(K)$ we have

$$\big(F(\cdot)f_n(\cdot)\big)^{-1}(B) = \bigcup_{e\in f_n(\Omega)} \Big[\big(F(\cdot)e\big)^{-1}(B) \cap f_n^{-1}(\{e\})\Big] \in \mathcal{A},$$

due to the measurability assumption on F. Combining (3.5.1), (3.5.2) with the fact that the set of $\mathcal{A}$-$\mathcal{B}(E)$-measurable functions is closed under the formation of pointwise limits; see, e.g., [51, Proposition A.1.3] yields the assertion under assumption (i).

Next we verify the assertion under assumption (ii). In this case, let $F_n\colon \Omega \to L(E,K)$, $n \in \mathbb{N}$ be a sequence of simple $L(E,K)$-valued functions such that for every $\omega \in \Omega$ it holds that

$$F(\omega) = \lim_{n\to\infty} F_n(\omega) \tag{3.5.3}$$

as an equality in $L(E,K)$; see [51, Lemma A.1.4]. Further, note that for every $n \in \mathbb{N}$ the mapping

$$F_n(\cdot)f(\cdot)\colon \Omega \to K,\ \omega \mapsto F_n(\omega)f(\omega) \text{ is } \mathcal{A}\text{-}\mathcal{B}(K)\text{-measurable.} \tag{3.5.4}$$

Indeed, in order to verify (3.5.4) it suffices to observe that for every $B \in \mathcal{B}(K)$ it holds that

$$\Big(F_n(\cdot)f(\cdot)\Big)^{-1}(B) = \bigcup_{L\in F_n(\Omega)} \Big[\big(L\circ f\big)^{-1}(B) \cap F_n^{-1}(\{L\})\Big] \in \mathcal{A}$$

due to the measurability assumption on f. Combining (3.5.3), (3.5.4) with the fact that the set of $\mathcal{A}$-$\mathcal{B}(E)$ measurable functions is closed under formation of pointwise limits; see, e.g., [51, Proposition A.1.3] yields the assertion under assumption (ii). □

Lemma 3.5.6. Assume the setting in Theorem 3.2.2 and let $h \in H$. Then the mapping

$$[0,T] \ni s \mapsto \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)h \in D([0,T],H)$$

is $\mathcal{B}\big([0,T]\big)$-$\mathcal{B}\big(D([0,T],H)\big)$-measurable. As a trivial consequence, it holds also that the mapping

$$[0,T]\times\Omega \ni (s,\omega) \mapsto \mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)h \in D([0,T],H)$$

is $\mathcal{P}_T$-$\mathcal{B}\big(D([0,T],H)\big)$-measurable.

Proof. We test the measurability by considering the sets of a suitable generator of $\mathcal{B}\big(D([0,T],H)\big)$. Let $n \in \mathbb{N}$, $B_1, B_2, \ldots, B_n \in \mathcal{B}(H)$, and $t_1, \ldots, t_n \in [0,T]$. Denote by

$\pi_{t_1,\dots,t_n}\colon D([0,T],H) \to H^n$ the projection $\pi_{t_1,\dots,t_n}(x) = \big(x(t_1),\dots,x(t_n)\big)$ and observe that

$$
\begin{aligned}
\big(\mathbb{1}_{[\cdot,T]}S\big)^{-1}\big(\pi^{-1}_{t_1,\dots,t_n}(B_1\times\ldots\times B_n)\big) &= \Big\{t\in[0,T]\colon \pi_{t_1,\dots,t_n}\big(\mathbb{1}_{[t,T]}(\cdot)S(\cdot,t)h\big)\in B_1\times\ldots\times B_n\Big\}\\
&= \Big\{t\in[0,T]\colon \mathbb{1}_{[t,T]}(t_1)S(t_1,t)h\in B_1,\dots,\mathbb{1}_{[t,T]}(t_n)S(t_n,t)h\in B_n\Big\}\\
&= \bigcap_{i=1}^{n}\underbrace{\Big\{t\in[0,T]\colon \mathbb{1}_{[t,T]}(t_i)S(t_i,t)h\in B_i\Big\}}_{\in\mathcal{B}([0,T])}\\
&\in \mathcal{B}\big([0,T]\big),
\end{aligned}
$$

which proves the assertion. □

Lemma 3.5.7. Assume the setting in Theorem 3.2.2. Then the mapping

$$[0,T]\times\Omega\ni(s,\omega)\mapsto \partial_2 f\big(s,X_s^S(\cdot,\omega)\big)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Psi(s,\omega)\big)\in V$$

is $\mathcal{P}_T$-$\mathcal{B}(V)$-measurable.

Proof. **Step 1:** The mapping $[0,T]\times\Omega\ni(s,\omega)\mapsto\big(s,X_s^S(\cdot,\omega)\big)\in[0,T]\times D([0,T],H)$ is $\mathcal{P}_T$-$\mathcal{B}\big(D([0,T],H)\big)$-measurable since the process $(X_s^S)_{s\in[0,T]}$ is $(\mathcal{F}_s)$-adapted and has continuous trajectories; see Lemma 3.2.1. The operator-valued mapping $\partial_2 f\colon[0,T]\times D([0,T],H)\to L\big(D([0,T],H),V\big)$ is continuous, therefore for all $y\in D([0,T],H)$ the composition $[0,T]\times\Omega\ni(s,\omega)\mapsto\partial_2 f\big(s,X_s^S(\cdot,\omega)\big)y\in V$ is $\mathcal{P}_T$-$\mathcal{B}\big(V\big)$-measurable.
Step 2: Note that the mapping

$$[0,T]\times\Omega\ni(s,\omega)\mapsto\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Psi(s,\omega)\in D([0,T],H)$$

is $\mathcal{P}_T$-$\mathcal{B}\big(D([0,T],H)\big)$-measurable, due to the predictability assumption on $(\Psi(t))_{t\in[0,T]}$, Lemma 3.5.6, and Lemma 3.5.5 with $E=H$ and $K=D([0,T],H)$, $f=\Psi$, $F=\mathbb{1}_{[\cdot,T]}S$, and $\Omega=[0,T]\times\Omega$ fulfilling assumption (i).
Step 3: Combining Step 1 and Step 2 and applying Lemma 3.5.5 once again with $E=D([0,T],H)$ fulfilling assumption (ii), we obtain that the mapping $(s,\omega)\mapsto\partial_2 f\big(s,X_s^S(\cdot,\omega)\big)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Psi(s,\omega)\big)\in V$ is $\mathcal{P}_T$-$\mathcal{B}(V)$-measurable. □

Lemma 3.5.8. Assume the setting in Theorem 3.2.2. Then the mapping

$$[0,T]\times\Omega\ni(s,\omega)\mapsto\partial_2 f\big(s,X_s^S(\cdot,\omega)\big)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s,\omega)\big)\in\mathrm{HS}(U,V)$$

is $\mathcal{P}_T$-$\mathcal{B}(\mathrm{HS}(U,V))$-measurable.

Proof. Let $\mathscr{H}\in\mathrm{HS}(U,H)$. With a similar argument as in the proof of Lemma 3.5.6, one can show that the mapping

$$[0,T]\times\Omega\ni(s,\omega)\mapsto\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\mathscr{H}\in D([0,T],\mathrm{HS}(U,H))$$

is $\mathcal{P}_T$-$\mathcal{B}\big(D([0,T],\mathrm{HS}(U,H))\big)$-measurable. Following an analogue idea as in Step 2 and Step 3 in the proof of Lemma 3.5.7 proves the assertion.

□

Lemma 3.5.9. Assume the setting in Theorem 3.2.2. Then the mapping

$$[0,T]\times\Omega\ni(s,\omega)\mapsto\sum_{i\in\mathbb{N}}\partial_2^2 f\big(s,X_s^S(\cdot,\omega)\big)\big(\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s,\omega)e_i,\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s,\omega)e_i\big)$$

is $\mathcal{P}_T$-$\mathcal{B}(V)$-measurable.

Proof. It suffices to show the measurability for only one summand. Let $i\in\mathbb{N}$ and $y\in D([0,T],H)$. As in the proof of Lemma 3.5.7, one can show that the mapping $[0,T]\times\Omega\ni(s,\omega)\mapsto\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s,\omega)e_i\in D([0,T],H)$ is $\mathcal{P}_T$-$\mathcal{B}\big(D([0,T],H)\big)$-measurable. Now, by Lemma 3.5.5 with $E=D([0,T],H)$ and $K=V$ fulfilling assumption (ii), we obtain that the mapping

$$[0,T]\times\Omega\ni(s,\omega)\mapsto\partial_2^2 f\big(s,X_s^S(\cdot,\omega)\big)\big(y\,,\mathbb{1}_{[s,T]}(\cdot)S(\cdot,s)\Phi(s,\omega)e_i\big)\in V$$

is $\mathcal{P}_T$-$\mathcal{B}(V)$-measurable. A second application of Lemma 3.5.5 thus finishes the proof.

□

Lemma 3.5.10. Assume the setting in Theorem 3.2.2, let $N\in\mathbb{N}$, and let $\tau_N\colon\Omega\to[0,T]$ be the stopping time defined in (3.3.14). Then it holds for all $t\in[0,T]$ that

$$\int_0^t \mathbb{1}_{(\!(0,\tau_N]\!]}(s)S(t,s)\Phi(s)\,\mathrm{d}W(s)=S(t,t\wedge\tau_N)\int_0^{t\wedge\tau_N}S(t\wedge\tau_N,s)\Phi(s)\,\mathrm{d}W(s),\quad\mathbb{P}\text{-a.s.}\tag{3.5.5}$$

Proof. In order to verify (3.5.5) above, we employ an approximation argument for stopping times similar to the argument in the proof of [51, Lemma 2.3.9]. To this end, we define the following sequence of simple stopping times

$$\tau_N^{(n)}=\sum_{k=0}^{2^n-1}\frac{T(k+1)}{2^n}\mathbb{1}_{(\frac{Tk}{2^n},\frac{T(k+1)}{2^n}]}\circ\tau_N,\quad n\in\mathbb{N},\tag{3.5.6}$$

with the property that $\tau_N^{(n)}\downarrow\tau_N$, $\mathbb{P}$-a.s. as $n\to\infty$. For each $n\in\mathbb{N}$, we can rewrite $\tau_N^{(n)}$ as

$$\tau_N^{(n)}=\sum_{k=0}^{2^n-1}a_k\mathbb{1}_{A_k}$$

where $a_k=\frac{T(k+1)}{2^n}$ and $A_k=\{\tau_N\in(\frac{Tk}{2^n},\frac{T(k+1)}{2^n}]\}=\{\tau_N^{(n)}=a_k\}$, $k=0,\cdots,2^n-1$. Thus, it holds that

$$\mathbb{1}_{(\!(0,\tau_N^{(n)}]\!]}(s)=\sum_{k=0}^{2^n-1}\mathbb{1}_{A_k}\mathbb{1}_{(0,a_k]}(s),\quad s\in[0,T].$$

Now note that for all $t \in [0,T]$ the following equalities hold in $L^2(\mathbb{P}\otimes\lambda; H)$

$$\begin{aligned}
\int_0^t \mathbb{1}_{(\!(0,\tau_N]\!]}(s)S(t,s)\Phi(s)\,\mathrm{d}W(s) &= \lim_{n\to\infty}\int_0^t \mathbb{1}_{(\!(0,\tau_N^{(n)}]\!]}(s)S(t,s)\Phi(s)\,\mathrm{d}W(s) \\
&= \lim_{n\to\infty}\int_0^t \sum_{k=0}^{2^n-1}\mathbb{1}_{A_k}\mathbb{1}_{(0,a_k]}(s)S(t,t\wedge\tau_N^{(n)})S(t\wedge\tau_N^{(n)},s)\Phi(s)\,\mathrm{d}W(s) \\
&= \lim_{n\to\infty}\int_0^t \sum_{k=0}^{2^n-1}\mathbb{1}_{A_k}\mathbb{1}_{(0,a_k]}(s)S(t,t\wedge a_k)S(t\wedge\tau_N^{(n)},s)\Phi(s)\,\mathrm{d}W(s) \\
&= \lim_{n\to\infty}\sum_{k=0}^{2^n-1} S(t,t\wedge a_k)\int_0^t \mathbb{1}_{A_k}\mathbb{1}_{(0,a_k]}(s)S(t\wedge\tau_N^{(n)},s)\Phi(s)\,\mathrm{d}W(s).
\end{aligned} \tag{3.5.7}$$

We continue with our computation by using [49, Proposition 8.11]

$$\begin{aligned}
&\lim_{n\to\infty}\sum_{k=0}^{2^n-1} S(t,t\wedge a_k)\int_0^t \mathbb{1}_{A_k}\mathbb{1}_{(0,a_k]}(s)S(t\wedge\tau_N^{(n)},s)\Phi(s)\,\mathrm{d}W(s) \\
&= \lim_{n\to\infty}\sum_{k=0}^{2^n-1} S(t,t\wedge\tau_N^{(n)})\int_0^t \mathbb{1}_{A_k}\mathbb{1}_{(0,a_k]}(s)S(t\wedge\tau_N^{(n)},s)\Phi(s)\,\mathrm{d}W(s) \\
&= \lim_{n\to\infty} S(t,t\wedge\tau_N^{(n)})\int_0^t \sum_{k=0}^{2^n-1}\mathbb{1}_{A_k}\mathbb{1}_{(0,a_k]}(s)S(t\wedge\tau_N^{(n)},s)\Phi(s)\,\mathrm{d}W(s) \\
&= \lim_{n\to\infty} S(t,t\wedge\tau_N^{(n)})\int_0^t \mathbb{1}_{(\!(0,\tau_N^{(n)}]\!]}(s)S(t\wedge\tau_N^{(n)},s)\Phi(s)\,\mathrm{d}W(s) \\
&= \lim_{n\to\infty} S(t,t\wedge\tau_N^{(n)})\int_0^{t\wedge\tau_N^{(n)}} S(t\wedge\tau_N^{(n)},s)\Phi(s)\,\mathrm{d}W(s).
\end{aligned} \tag{3.5.8}$$

Next, observe that due to continuity of the mapping $[0,T]\ni t\to\int_0^t S(t,s)\Phi(s)\mathrm{d}W(s)\in H$ and Lemma 3.5.1, we have $\mathbb{P}$-a.s. that

$$S(t,t\wedge\tau_N)\int_0^{t\wedge\tau_N} S(t\wedge\tau_N,s)\Phi(s)\,\mathrm{d}W(s) = \lim_{n\to\infty} S(t,t\wedge\tau_N^{(n)})\int_0^{t\wedge\tau_N^{(n)}} S(t\wedge\tau_N^{(n)},s)\Phi(s)\,\mathrm{d}W(s).$$

This together with (3.5.7) and (3.5.8) completes the proof. □

On time integrals of predictable processes

Lemma 3.5.11. Let $T\in(0,\infty)$, let $\big(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\in[0,T]},\mathbb{P}\big)$ be a filtered probability space such that the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ satisfies the usual conditions, and let V be a separable real Hilbert space. Moreover, let $Z=(Z(t))_{t\in[0,T]}$ be a bounded V-valued $\mathcal{P}_T$-measurable process and let $t_j^n=\frac{jT}{2^n}$, $j\in\{0,1,\ldots,2^n\}$, $n\in\mathbb{N}$. Then it holds that

$$\mathbb{E}\left(\left\|\int_0^T Z(t)\mathrm{d}t-\sum_{j=1}^{2^n}\mathbb{E}\left(\int_{t_{j-1}^n}^{t_j^n} Z(t)\mathrm{d}t\Big|\mathcal{F}_{t_{j-1}^n}\right)\right\|_V\right)\xrightarrow{n\to\infty}0.$$

Proof. Due to predictability of the process Z and by a similar argument as used in the proof of [51, Proposition 2.3.8], there exists a sequence $(Z^{(m)})_{m\in\mathbb{N}}$ of predictable simple processes of the form

$$Z^{(m)}(\omega,t)=\sum_{k=1}^{K^{(m)}} z_k^{(m)}\mathbb{1}_{F_k^{(m)}\times(t_{k-1}^{(m)},t_k^{(m)}]}(\omega,t), \tag{3.5.9}$$

where $z_k^{(m)}\in V$, $F_k^{(m)}\in\mathcal{F}_{t_{k-1}^{(m)}}$, $0\le t_0^{(m)}\le\ldots\le t_{K^{(m)}}^{(m)}\le T$, such that

$$\lim_{m\to\infty}\mathbb{E}\int_0^T\left\|Z(t)-Z^{(m)}(t)\right\|_V^2\,\mathrm{d}t=0.$$

In particular, $(Z^{(m)})_{m\in\mathbb{N}}$ is a Cauchy-sequence in $L^2(\lambda;L^2(\mathbb{P};V))$ and after passing to a subsequence it holds for λ-a.e. $t\in[0,T]$ that

$$\lim_{m\to\infty}Z^{(m)}(t)=Z(t)\quad\text{in } L^2(\mathbb{P};V). \tag{3.5.10}$$

For $j\in\{1,2,\ldots,2^n\}$, $n,m\in\mathbb{N}$, let the function $\eta_j^{n,m}\in\mathscr{L}^2\big(\Omega\times[0,T],\mathcal{F}_{t_{j-1}^n}\otimes\mathcal{B}([0,T]),\mathbb{P}\otimes\lambda;V\big)$ be given by

$$\eta_j^{n,m}(\omega,t)=\sum_{k=1}^{K^{(m)}} z_k^{(m)}\mathbb{E}\big(\mathbb{1}_{F_k^{(m)}}\big|\mathcal{F}_{t_{j-1}^n}\big)(\omega)\,\mathbb{1}_{(t_{k-1}^{(m)},t_k^{(m)}]}(t)\quad\forall(\omega,t)\in\Omega\times[0,T],$$

where by $\mathbb{E}\big(\mathbb{1}_{F_k^{(m)}}\big|\mathcal{F}_{t_{j-1}^n}\big)$ we mean a fixed version of the conditional expectation. Then it holds for all $t\in[0,T]$ that

$$\eta_j^{n,m}(\cdot,t)=\mathbb{E}\big(Z^{(m)}(t)\big|\mathcal{F}_{t_{j-1}^n}\big). \tag{3.5.11}$$

Remember that $(Z^{(m)})_{m\in\mathbb{N}}$ is a Cauchy-sequence in $L^2(\lambda;L^2(\mathbb{P};V))$ and note that the conditional expectation $\mathbb{E}\big(\,\cdot\,\big|\mathcal{F}_{t_{j-1}^n}\big)\colon L^2(\Omega,\mathcal{F},\mathbb{P};V)\to L^2(\Omega,\mathcal{F}_{t_{j-1}^n},\mathbb{P};V)$ is a bounded linear operator and therefore the sequence $(\eta_j^{n,m})_{m\in\mathbb{N}}$ is also a Cauchy-sequence in $L^2(\lambda;L^2(\mathbb{P};V))$. For $j\in\{1,2,\ldots,2^n\}$, $n\in\mathbb{N}$, let η_j^n be given by

$$\eta_j^n=\lim_{m\to\infty}\eta_j^{n,m}\quad\text{ in } L^2(\lambda;L^2(\Omega,\mathcal{F}_{t_{j-1}^n},\mathbb{P};V)).$$

Then, after a passing to a subsequence it holds for λ-a.e. $t\in[0,T]$ that

$$\eta_j^n(\cdot,t)=\lim_{m\to\infty}\eta_j^{n,m}(\cdot,t)\quad\text{ in } L^2(\Omega,\mathcal{F}_{t_{j-1}^n},\mathbb{P};V). \tag{3.5.12}$$

The combination of (3.5.10)–(3.5.12) implies for λ-a.e. $t\in[0,T]$ and $A\in\mathcal{F}_{t_{j-1}^n}$ that

$$\int_A Z(\omega,t)\mathbb{P}(\mathrm{d}\omega)=\int_A\eta_j^n(\omega,t)\mathbb{P}(\mathrm{d}\omega).$$

Therefore, for λ-a.e. $t \in (t_{j-1}^n, t_j^n]$ we get that

$$\mathbb{E}\left(Z(t)\middle|\mathcal{F}_{t_{j-1}^n}\right) = \eta_j^n(\cdot, t), \tag{3.5.13}$$

as an equality in $L^1(\Omega, \mathcal{F}_{t_{j-1}^n}, \mathbb{P}; V)$. Next we show that

$$\mathbb{E}\left(\int_{t_{j-1}^n}^{t_j^n} Z(t)\mathrm{d}t\middle|\mathcal{F}_{t_{j-1}^n}\right) = \int_{t_{j-1}^n}^{t_j^n} \eta_j^n(t)\mathrm{d}t \quad \mathbb{P}\text{-a.s.} \tag{3.5.14}$$

To this end, let $A \in \mathcal{F}_{t_{j-1}^n}$ and observe that

$$\int_A \int_{t_{j-1}^n}^{t_j^n} Z(t)\,\mathrm{d}t\,\mathrm{d}\mathbb{P} = \int_{t_{j-1}^n}^{t_j^n} \int_\Omega \mathbb{1}_A Z(t)\,\mathrm{d}\mathbb{P}\,\mathrm{d}t = \int_{t_{j-1}^n}^{t_j^n} \int_\Omega \mathbb{1}_A \eta_j^n(t)\,\mathrm{d}\mathbb{P}\,\mathrm{d}t = \int_A \int_{t_{j-1}^n}^{t_j^n} \eta_j^n(t)\,\mathrm{d}t\,\mathrm{d}\mathbb{P},$$

where for the second equality above we used (3.5.13). This proves (3.5.14). Therefore we obtain that

$$\begin{aligned}
\mathbb{E}\left(\left\|\int_0^T Z(t)\mathrm{d}t - \sum_{j=1}^{2^n} \mathbb{E}\left(\int_{t_{j-1}^n}^{t_j^n} Z(t)\mathrm{d}t\middle|\mathcal{F}_{t_{j-1}^n}\right)\right\|_V\right) &= \mathbb{E}\left(\left\|\int_0^T Z(t)\mathrm{d}t - \sum_{j=1}^{2^n}\int_{t_{j-1}^n}^{t_j^n} \eta_j^n(t)\mathrm{d}t\right\|_V\right) \\
= \mathbb{E}\left(\left\|\int_0^T Z(t)\mathrm{d}t - \int_0^T \sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n, t_j^n]}(t)\eta_j^n(t)\mathrm{d}t\right\|_V\right)& \\
\leq \int_0^T \sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n, t_j^n]}(t)\mathbb{E}\left(\left\|Z(t) - \eta_j^n(t)\right\|_V\right)\mathrm{d}t.&
\end{aligned}$$

Now let $t \in [0, T]$ be fixed and let $(j_n)_{n\in\mathbb{N}} \subset [0, T]$ be such that $\lim_{n\to\infty} t_{j_n-1}^n = t$, and $t \in (t_{j_n-1}^n, t_{j_n}^n]$, $n \in \mathbb{N}$. The martingale convergence theorem [50, Theorem 1.14] thus implies that

$$\mathbb{E}\left(Z(t)|\sigma(\bigcup_{s\in[0,t)} \mathcal{F}_s)\right) = \lim_{n\to\infty} \mathbb{E}\left(Z(t)\middle|\mathcal{F}_{t_{j_n-1}^n}\right) = \lim_{n\to\infty} \eta_{j_n}^n(\cdot, t) \quad \text{ in } L^1(P; V),$$

which, by considering the predictability of Z can be rewritten as follows

$$Z(t) = \lim_{n\to\infty} \eta_{j_n}^n(\cdot, t) \quad \text{ in } L^1(P; V). \tag{3.5.15}$$

Indeed, the definition (3.5.9) and the estimation (3.5.10) imply that

$$Z(t) = \mathbb{E}\left(Z(t)|\sigma(\bigcup_{s\in[0,t)} \mathcal{F}_s)\right) \quad \mathbb{P}\text{-a.s.}$$

The boundedness assumption on $(Z(t))_{t\in[0,T]}$ allows us to apply the dominated convergence theorem and obtain that

$$\int_0^T \sum_{j=1}^{2^n} \mathbb{1}_{(t_{j-1}^n, t_j^n]}(t)\mathbb{E}\Big(\big\|Z(t) - \eta_j^n(t)\big\|_V\Big)\mathrm{d}t \xrightarrow{n\to\infty} 0.$$

This concludes the assertion of the Lemma. □

Chapter 4

Weak error analysis of approximations of path-dependent functionals of mild solutions of SEEs via Malliavin calculus

In this chapter we develop a Malliavin calculus approach to analyze the weak error of spatial approximations of path-dependent functionals of mild solutions of stochastic evolution equations (SEEs). In order to do that we first introduce in Assumption 4.1.1 below the notation we use and the setting we assume throughout this chapter. After mentioning some regularity results from the literature, we state in Lemma 4.3.1 below a Malliavin regularity result for mild solutions of SEEs of the type

$$\begin{aligned} \mathrm{d}X(t) &= \big[AX(t) + F\big(X(t)\big)\big]\mathrm{d}t + B\big(X(t)\big)\mathrm{d}W(t), \qquad t \in [0,T] \\ X(0) &= \xi \in H, \end{aligned} \tag{4.0.1}$$

where H and U are separable real Hilbert spaces, $\big(W(t)\big)_{t\in[0,T]}$ is an Id_U-Wiener process defined on a filtered probability space $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\big)$, and $F\colon H \to H$ and $B\colon H \to L(U,H)$ are measurable mappings satisfying suitable regularity assumptions. Using results from [22], we show in Lemma 4.3.1 that the mild solution of (4.0.1) is Malliavin differentiable and the Malliavin derivative $\big(D_s(X(t))\big)_{s,t\in[0,T]}$ is continuous and satisfies suitable boundedness properties. For a finite Borel measure μ, we show in Lemma 4.3.4 that the $L^2(\mu;H)$-valued random variable $X(\cdot)$, i.e., the random variable whose values are the sample paths of the stochastic process $(X(t))_{t\in[0,T]}$, is also Malliavin differentiable. Moreover, we investigate the relation between Malliavin derivative of $(X(t))_{t\in[0,T]}$ and of $X(\cdot)$, using a proper isometry introduced in Lemma 4.3.3 below. In the last section of this chapter, the main result is presented in Theorem 4.4.3.

There we show an upper bound for the weak error $\mathbb{E}\big|f(X)-f(\tilde{X})\big|$, where $\tilde{X}$ is an approximation of X and f is a functional on the sample path space of X. Our approach to prove this upper bound is based on Malliavin calculus and the regularity results proved in this chapter.

4.1 Preliminaries

We start this section with introducing a setting which we assume throughout this chapter. Assumption 4.1.1 below is based on [22, Hypothesis 3.1]. Consider the following abstract SEE

$$\begin{aligned} \mathrm{d}X(t) &= \big[AX(t)+F\big(X(t)\big)\big]\mathrm{d}t + B\big(X(t)\big)\mathrm{d}W(t), \qquad t\in[0,T] \\ X(0) &= \xi \in H, \end{aligned} \tag{4.1.1}$$

where the following is assumed:

Assumption 4.1.1. Let Assumption 2.4.7 hold and additionally assume that

(i) the mappings $F\colon H\to H$ and $B\colon H\to L(U,H)$ satisfy for all $s>0$, $x\in H$, $u\in U$ that $F\in \mathrm{Lip}(H,H)$, $e^{sA}B(x)\in \mathrm{HS}(U,H)$ and that $B(\cdot)u\colon H\to H$ is measurable. Moreover, there exist constants $L>0$ and $\vartheta\in[0,1)$ such that for all $s>0$, $x,y\in H$ it holds that

$$\begin{aligned} \big\|e^{sA}B(x)\big\|_{\mathrm{HS}(U,H)} &\le Ls^{-\vartheta/2}\big(1+\|x\|_H\big), \\ \big\|e^{sA}B(x)-e^{sA}B(y)\big\|_{\mathrm{HS}(U,H)} &\le Ls^{-\vartheta/2}\,\|x-y\|_H \\ \|B(x)\|_{L(U,H)} &\le L(1+\|x\|_H). \end{aligned}$$

(ii) for every $s>0$ it holds that

$$F\in\mathscr{G}^1(H,H), \qquad e^{sA}B\in\mathscr{G}^1\big(H,\mathrm{HS}(U,H)\big).$$

Remember that by $\mathscr{G}^1$ we denote the space of Gâteaux differentiable mappings; see Definition 2.3.1.

Recall from Definition 2.4.8 that a mild solution of (4.1.1) is an H-valued predictable process such that for all $t\in[0,T]$ it holds $\mathbb{P}$-a.s. that

$$X(t) = e^{tA}\xi + \int_0^t e^{(t-s)A}F(X(s))\,\mathrm{d}s + \int_0^t e^{(t-s)A}B(X(s))\,\mathrm{d}W(s). \tag{4.1.2}$$

In particular, the integrals in (4.1.2) above have to be well-defined. According to [22, Proposition 3.2], Assumption 4.1.1 guarantees the existence of a unique mild solution $X=\big(X(t)\big)_{t\in[0,T]}$ of (4.1.1) with continuous sample paths such that (4.1.2) holds $\mathbb{P}$-a.s. for all $t\in[0,T]$. Moreover, for every $p\in[2,\infty)$ it holds that $X\in L^p\big(\mathbb{P};C([0,T],H)\big)$ with

$$\mathbb{E}\Big[\sup_{t\in[0,T]}\big\|X(t)\big\|_H^p\Big] < C\big(1+\|\xi\|_H\big)^p, \tag{4.1.3}$$

for some constant $C \in (0,\infty)$ depending only on p, ϑ, T, F, B and $M := \sup_{t\in[0,T]} \left\|e^{tA}\right\|_{L(H)}$.

Lemma 4.1.2. Let Assumption 4.1.1 be fulfilled, let $p \geq 2$ and let $X \in L^p(\mathbb{P}; C([0,T],H))$ be the mild solution of (4.1.1) with continuous paths. Then it holds that

$$X \in C\big([0,T], L^p(\mathbb{P};H)\big).$$

Proof. Let $t \in [0,T]$. Observe that for every $t_0 \in [0,T]$ it holds that $\|X(t) - X(t_0)\|_H^p \leq 2\sup_{s\in[0,T]} \|X(s)\|_H^p$ and that $\mathbb{E}(\sup_{s\in[0,T]} \|X(s)\|_H^p) < \infty$. As a consequence, the dominated convergence theorem implies that

$$\lim_{t\to t_0} \mathbb{E}\big(\|X(t) - X(t_0)\|_H^p \big) = 0.$$

□

4.2 Malliavin calculus in Hilbert spaces

In this section we review some definitions and auxiliary results from [39, Chapter 4] and [22, Section 3.3]. Let U and H be separable real Hilbert spaces, let $T \in (0,\infty)$, and let $\big(W(t)\big)_{t\in[0,T]}$ be an Id_U-Wiener process defined on a filtered probability space $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\big)$. We denote by $\mathscr{S}(H)$ the set of all smooth and cylindrical H-valued random variables F of the form

$$F = \sum_{j=1}^{n} f_j(W(\Phi_1), \ldots, W(\Phi_m))h_j,$$

where $m,n \in \mathbb{N}$, $\Phi_i \in L^2([0,T], \lambda; \mathrm{HS}(U,\mathbb{R}))$, $W(\Phi_i) = \int_0^T \Phi_i(t)\,\mathrm{d}W(t)$, $h_j \in H$, and $f_j\colon \mathbb{R}^m \to \mathbb{R}$ is infinitely differentiable function with the property that f_j and all its derivatives are at most polynomially growing, $i = 1,\ldots,m$, $j = 1,\ldots,n$. Then the Malliavin derivative DF of $F \in \mathscr{S}(H)$ is defined by

$$D_\tau F = \sum_{j=1}^{n}\sum_{i=1}^{m} \frac{\partial}{\partial x_i} f_j(W(\Phi_1),\ldots,W(\Phi_m))\, h_j \otimes \Phi_i(\tau) \quad \text{for } \tau \in [0,T],$$

where $h_j \otimes \Phi_i(\tau)$ denotes the tensor product of $h_j \in H$ and $\Phi_i(\tau) \in \mathrm{HS}(U,\mathbb{R})$ and indicates the operator $U \ni u \mapsto \big(\Phi_i(\tau)\big)[u] \cdot h_j \in H$.

In [39, Propostion 4.2], it is shown that the Malliavin derivative opertor $D\colon \mathscr{S}(H) \subset L^2(\mathbb{P};H) \to L^2(\Omega\times[0,T], \mathbb{P}\otimes\lambda; \mathrm{HS}(U,H))$ is well-defined. In particular, DF does not depend on the representation of $F \in \mathscr{S}(H)$. It is well-known; see, e.g., [39, Proposition 4.4], that the Malliavin derivative operator $D\colon \mathscr{S}(H) \subset L^2(\mathbb{P};H) \to L^2(\Omega\times[0,T], \mathbb{P}\otimes\lambda; \mathrm{HS}(U,H))$ is closable, i.e., if

$(F_n)_{n\in\mathbb{N}} \subset \mathscr{S}(H)$ is a sequence with

$$\lim_{n\to\infty} F_n = 0, \quad \text{in } L^2(\mathbb{P};H),$$
$$\lim_{n\to\infty} DF_n = \Upsilon, \quad \text{in } L^2(\mathbb{P}\otimes\lambda;\mathrm{HS}(U,H)),$$

then $\Upsilon = 0$. We write $\mathbb{D}^{1,2}(H)$ for the closure of $\mathscr{S}(H)$ in $L^2(\mathbb{P};H)$ with respect to the seminorm

$$\|F\|_{\mathbb{D}^{1,2}(H)} = \big(\|F\|^2_{L^2(\mathbb{P};H)} + \|DF\|^2_{L^2(\mathbb{P}\otimes\lambda;\mathrm{HS}(U,H))}\big)^{1/2}$$

and obtain a well-defined extension of the Malliavin derivative operator

$$D\colon \mathbb{D}^{1,2}(H) \subset L^2(\mathbb{P};H) \to L^2(\mathbb{P}\otimes\lambda;\mathrm{HS}(U,H)).$$

We also remark that, for $F \in \mathbb{D}^{1,2}(H)$, DF can be interpreted as a stochastic process $(D_\tau F)_{\tau\in[0,T]}$ with values in $\mathrm{HS}(U,H)$. In particular, for $\tau \in [0,T]$ and $u \in \mathcal{U}$, $(D_\tau F)[u]$ can be interpreted as an H-valued random variable. Note, however, that the distributions of the $\mathrm{HS}(U,H)$-valued random variables $D_\tau F$, $\tau \in [0,T]$, are not uniquely determined for all $t \in [0,T]$, since $DF \in L^2(\mathbb{P}\otimes\lambda;\mathrm{HS}(U,H))$ is a $\mathbb{P}\otimes\lambda$-equivalence class of functions on $\Omega\times[0,T]$.

Remark 4.2.1. For $\Phi \in L^2([0,T],\lambda;\mathrm{HS}(U,\mathbb{R}))$, let the real-valued stochastic process $(X(t))_{t\in[0,T]}$ be defined by $X(t) = \int_0^t \Phi(\sigma)\mathrm{d}W(\sigma)$, $t \in [0,T]$. Obviously, it holds for all $t,\tau \in [0,T]$ that $X(t) \in \mathscr{S}(\mathbb{R})$ and that

$$D_\tau\big(X(t)\big) = \mathbb{1}_{[0,t]}(\tau)\Phi(\tau).$$

In particular, $D_\tau X(t) = 0$, $\mathbb{P}\otimes\mathrm{d}\tau$-a.e. on $\Omega\times(t,T]$. More generally, if $F \in \mathbb{D}^{1,2}(H)$ is $\mathcal{F}_t$-measurable, then it holds that $D_\tau F = 0$, $\mathbb{P}\otimes\mathrm{d}\tau$-a.e. on $\Omega\times(t,T]$. Indeed, for an $\mathcal{F}_t$-measurable $F \in \mathbb{D}^{1,2}(H)$ let $(F_n)_{n\in\mathbb{N}}$ be a sequence of $\mathcal{F}_t$-measurable $\mathscr{S}(H)$-valued random variables such that $F_n \to F$ in $\mathbb{D}^{1,2}(H)$, as $n\to\infty$. For each $n\in\mathbb{N}$, let the following representation be given:

$$F_n = \sum_{j=1}^{N_n} f_j^{(n)}\big(W(\Phi_1^{(n)}),\dots,W(\Phi_{m_n}^{(n)})\big)h_j^{(n)}.$$

As the random variables F_n, $n\in\mathbb{N}$, are $\mathcal{F}_t$-measurable, we can replace each $\Phi_i^{(n)}$, $i \in \{1,\dots,m_n\}$, $n\in\mathbb{N}$ with $\mathbb{1}_{[0,t]}\Phi_i^n$. Therefore we obtain for $\tau\in[0,T]$ that

$$D_\tau F_n = \sum_{j=1}^{N_n}\sum_{i=1}^{m_n} \frac{\partial}{\partial x_i} f_j^{(n)}\big(W(\mathbb{1}_{[0,t]}\Phi_1^{(n)}),\dots,W(\mathbb{1}_{[0,t]}\Phi_{m_n}^{(n)})\big)\, h_j^{(n)} \otimes \mathbb{1}_{[0,t]}(\tau)\Phi_i^{(n)}(\tau).$$

This shows for each $n\in\mathbb{N}$ that $D_\tau F_n = 0$, $\mathbb{P}\otimes\mathrm{d}\tau$-a.e. on $\Omega\times(t,T]$. Consequently, it holds that $D_\tau F = 0$, $\mathbb{P}\otimes\mathrm{d}\tau$-a.e. on $\Omega\times(t,T]$. Compare with [22, page 1415].

The following result is a particular case of the so-called chain rule for the Malliavin derivative and the proof can be found in [39, Lemma 4.7].

Lemma 4.2.2. Let $\phi\colon H_1 \to H_2$ be a continuously Fréchet differentiable mapping, where $(H_1, \langle\cdot,\cdot\rangle_{H_1}, \|\cdot\|_{H_1})$ and $(H_2, \langle\cdot,\cdot\rangle_{H_2}, \|\cdot\|_{H_2})$ are arbitrary separable real Hilbert spaces. Assume that there exists a constant $C \in (0,\infty)$ such that

$$\|\phi(h)\|_{H_2} \le C\big(1 + \|h\|_{H_1}\big), \quad \|\phi'(h)\|_{L(H_1,H_2)} \le C, \quad \forall h \in H_1.$$

Then for all $F \in \mathbb{D}^{1,2}(H_1)$ it holds that $\phi(F) \in \mathbb{D}^{1,2}(H_2)$ and

$$D\phi(F) = \phi'(F)DF.$$

The next result is the so-called Malliavin integration-by-parts formula and its proof can be found in [39, Proposition 4.3].

Lemma 4.2.3. For all $\Phi \in L^2(\Omega \times [0,T], \mathbb{P}\otimes\lambda; \mathrm{HS}(U,H))$ and all $F \in \mathbb{D}^{1,2}(H)$ it holds that

$$\mathbb{E}\Big[\Big\langle F, \int_0^T \Phi(t)\,\mathrm{d}W(t)\Big\rangle_H\Big] = \mathbb{E}\Big[\big\langle DF, \Phi\big\rangle_{L^2([0,T],\lambda;\mathrm{HS}(U,H))}\Big].$$

As an application of the Malliavin integration-by-parts formula, a short proof of the stochastic Fubini theorem is given in [39, Theorem 4.18]. We will use the stochastic Fubini theorem later in this chapter.

Lemma 4.2.4. Let $(E, \mathcal{E}, \mu)$ be a measure space with finite measure μ and let $\Phi\colon \Omega\times[0,T]\times E \to \mathrm{HS}(U,H)$ be a $(\mathcal{P}_T \otimes \mathcal{E})$-$\mathcal{B}(\mathrm{HS}(U,H))$-measurable mapping. Under the condition

$$\int_E \Big(\mathbb{E}\Big[\int_0^T \|\Phi(t,x)\|^2_{\mathrm{HS}(U,H)}\,\mathrm{d}t\Big]\Big)^{\frac{1}{2}} \mu(\mathrm{d}x) < \infty,$$

it holds $\mathbb{P}$-almost surely that

$$\int_0^T \int_E \Phi(t,x)\,\mu(\mathrm{d}x)\mathrm{d}W(t) = \int_E \int_0^T \Phi(t,x)\,\mathrm{d}W(t)\mu(\mathrm{d}x). \tag{4.2.1}$$

In particular, all integrals in (4.2.1) are well-defined.

4.3 Malliavin regularity of mild solutions of SEEs

In this section we investigate the Malliavin differentiability of the mild solution $(X(t))_{t\in[0,T]}$ of SPDE (4.1.1) and the regularity properties of its derivative. Moreover, we formulate the relation between the Malliavin derivatives of the H-valued random variables $X(t)$, $t \in [0,T]$, and the Malliavin derivative of the $L^2(\mu; H)$-valued random variable $X(\cdot)$, where μ is a finite Borel measure on $[0,T]$.

The following result is based on [22, Proposition 3.5] and complements some of the assertions therein in a way that is suitable for our purpose.

Lemma 4.3.1. Let Assumption 4.1.1 be fulfilled and let $X \in L^2(\mathbb{P}; C([0,T], H))$ be the mild solution of (4.1.1) with continuous sample paths. Here and below we set $\frac{2}{\vartheta} = \infty$ if $\vartheta = 0$. Then

(i) it holds for every $t \in [0,T]$ that $X(t) \in \mathbb{D}^{1,2}(H)$,

(ii) the mapping $[0,T] \ni t \mapsto X(t) \in \mathbb{D}^{1,2}(H)$ is continuous, and

(iii) there exists a measurable mapping $\eta\colon \Omega \times [0,T] \times [0,T] \to \mathrm{HS}(U,H)$ such that for all $t \in [0,T]$ it holds that

$$\eta(\cdot, t) = D_\cdot(X(t))$$

as an equality in $L^2(\Omega \times [0,T], \mathbb{P} \otimes \lambda; \mathrm{HS}(U,H))$ and for all $p \in [2, 2/\vartheta)$ we have that

$$\sup_{s\in[0,T)} \mathbb{E}\Big(\sup_{t\in(s,T]} (t-s)^{p\vartheta/2} \, \|\eta(s,t)\|^p_{\mathrm{HS}(U,H)} \Big) < \infty.$$

Proof. For the sake of clarity, we divide the proof into several steps.
Step I: Observe that item (i) is a direct consequence of [22, Proposition 3.5(ii)]. Next note that [22, Proposition 3.5] further implies the existence of a measurable mapping

$$\eta\colon \Omega \times [0,T] \times [0,T] \to \mathrm{HS}(U,H)$$

and a Lebesgue null set $N \in \mathcal{B}([0,T])$ with the following properties:

(a) For all $t \in [0,T] \setminus N$ it holds that

$$\eta(\cdot, t) = D_\cdot(X(t))$$

as an equality in $L^2(\Omega \times [0,T], \mathbb{P} \otimes \lambda; \mathrm{HS}(U,H))$.

(b) For all $\tau \in [0,T]$, $\omega \in \Omega$ the mapping

$$(\tau, T] \ni t \mapsto \eta(\omega, \tau, t) \in \mathrm{HS}(U,H),$$

is continuous.

(c) For all $p \in [2, \infty)$ it holds that

$$\sup_{s\in[0,T)} \mathbb{E}\Big(\sup_{r\in(s,T]} (r-s)^{p\vartheta/2} \, \|\eta(s,r)\|^p_{\mathrm{HS}(U,H)} \Big) < \infty.$$

The verification of items (ii) and (iii) above is based on item (i), items (a)–(c), and the auxiliary results derived in Steps II-V below.

Step II: We show that for all $t \in [0,T]$, $p \in [2, 2/\vartheta)$ it holds that $\eta(\cdot,t) \in L^p(\mathbb{P}\otimes\lambda; \mathrm{HS}(U,H))$ and that the mapping

$$[0,T] \ni t \mapsto \eta(\cdot,t) \in L^p(\mathbb{P}\otimes\lambda; \mathrm{HS}(U,H))$$

is continuous. First note that, since for all $t \in [0,T]$ it holds that $D_\tau(X(t)) = 0$ $\mathbb{P}\otimes \mathrm{d}\tau$ almost everywhere on $\Omega \times (t,T]$ (see Remark 4.2.1 above), we can, without loss of generality, assume that for all $(\omega,\tau,t) \in \Omega\times[0,T]\times[0,T]$ with $\tau > t$ it holds that $\eta(\omega,\tau,t) = 0$. Therefore, for all $t \in [0,T]$ and $p \in [2, 2/\vartheta)$, we can write

$$\begin{aligned}
\mathbb{E}\int_0^T \|\eta(\tau,t)\|^p_{\mathrm{HS}(U,H)}\,\mathrm{d}\tau &= \mathbb{E}\int_0^T \big\|\mathbb{1}_{[0,t)}(\tau)\,\eta(\tau,t)\big\|^p_{\mathrm{HS}(U,H)}\,\mathrm{d}\tau \\
&\le \sup_{s\in[0,T]} \mathbb{E}\Big(\sup_{r\in(s,T]} (r-s)^{p\vartheta/2}\,\|\eta(s,r)\|^p_{\mathrm{HS}(U,H)}\Big)\int_0^T \mathbb{1}_{[0,t)}(\tau)(t-\tau)^{-p\vartheta/2}\mathrm{d}\tau \\
&\le \sup_{s\in[0,T]} \mathbb{E}\Big(\sup_{r\in(s,T]} (r-s)^{p\vartheta/2}\,\|\eta(s,r)\|^p_{\mathrm{HS}(U,H)}\Big)\int_0^T (T-\tau)^{-p\vartheta/2}\mathrm{d}\tau \\
&= \sup_{s\in[0,T]} \mathbb{E}\Big(\sup_{r\in(s,T]} (r-s)^{p\vartheta/2}\,\|\eta(s,r)\|^p_{\mathrm{HS}(U,H)}\Big)\frac{T^{1-p\vartheta/2}}{1-p\vartheta/2} < \infty,
\end{aligned} \tag{4.3.1}$$

by item (c) above. This proves that $\eta(\cdot,t) \in L^p(\mathbb{P}\otimes\lambda; \mathrm{HS}(U,H))$, for $t \in [0,T]$ and $p \in [2,2/\vartheta)$. Now let $t_0 \in (0,T]$ and $p \in [2,2/\vartheta)$ be fixed. Next we prove that

$$\lim_{t\to t_0} \mathbb{E}\int_0^T \|\eta(\tau,t) - \eta(\tau,t_0)\|^p_{\mathrm{HS}(U,H)}\,\mathrm{d}\tau = 0. \tag{4.3.2}$$

Observe that, due to item (b) above it holds $\mathbb{P}$-almost surely for all $\tau \in [0,T]$ that

$$\lim_{t\to t_0} \mathbb{1}_{[0,t\wedge t_0)}(\tau)\,\|\eta(\tau,t) - \eta(\tau,t_0)\|^p_{\mathrm{HS}(U,H)} = 0. \tag{4.3.3}$$

In order to prove the continuity (4.3.2), we have to show the left- and right-continuity separately. We start first with the right-continuity: For all $t \in [t_0,T]$ we have that

$$\begin{aligned}
\mathbb{E}\int_0^T \|\eta(\tau,t) - \eta(\tau,t_0)\|^p_{\mathrm{HS}(U,H)}\,\mathrm{d}\tau &= \mathbb{E}\int_0^T \mathbb{1}_{[0,t_0)}(\tau)\,\|\eta(\tau,t) - \eta(\tau,t_0)\|^p_{\mathrm{HS}(U,H)}\,\mathrm{d}\tau \\
&\quad + \mathbb{E}\int_0^T \mathbb{1}_{[t_0,t)}(\tau)\,\|\eta(\tau,t) - \eta(\tau,t_0)\|^p_{\mathrm{HS}(U,H)}\,\mathrm{d}\tau.
\end{aligned} \tag{4.3.4}$$

Moreover, for all $t \in [t_0, T]$ and $\tau \in [0, T]$ it holds that

$$\begin{aligned}
\mathbb{1}_{[0,t_0)}(\tau)\, \|\eta(\tau,t)-\eta(\tau,t_0)\|^p_{\mathrm{HS}(U,H)} &\le \mathbb{1}_{[0,t_0)}(\tau)\, 2^{p-1}\Big(\|\eta(\tau,t)\|^p_{\mathrm{HS}(U,H)} + \|\eta(\tau,t_0)\|^p_{\mathrm{HS}(U,H)} \Big) \\
&\le \mathbb{1}_{[0,t_0)}(\tau)\, 2^{p-1}\Big(\sup_{r\in(\tau,T]} (r-\tau)^{p\vartheta/2}\, \|\eta(\tau,r)\|^p_{\mathrm{HS}(U,H)} \Big) \\
&\qquad \cdot \big((t-\tau)^{-p\vartheta/2} + (t_0-\tau)^{-p\vartheta/2} \big) \\
&\le \mathbb{1}_{[0,t_0)}(\tau)\, 2^{p}\,\Big(\sup_{r\in(\tau,T]} (r-\tau)^{p\vartheta/2}\, \|\eta(\tau,r)\|^p_{\mathrm{HS}(U,H)} \Big) \\
&\qquad \cdot (t_0-\tau)^{-p\vartheta/2} \in L^1(\mathbb{P}\otimes \mathrm{d}\tau;\mathbb{R}),
\end{aligned} \tag{4.3.5}$$

by the fact that $(t-\tau)^{-p\vartheta/2} \le (t_0-\tau)^{-p\vartheta/2}$ and item (c) above. Indeed, it holds that

$$\begin{aligned}
&\mathbb{E}\int_0^T \mathbb{1}_{[0,t_0)}(\tau)\Big(\sup_{r\in(\tau,T]} (r-\tau)^{p\vartheta/2}\, \|\eta(\tau,r)\|^p_{\mathrm{HS}(U,H)} \Big)(t_0-\tau)^{-p\vartheta/2}\mathrm{d}\tau \\
&\quad \le \sup_{s\in[0,T)} \mathbb{E}\Big(\sup_{r\in(s,T]} (r-s)^{p\vartheta/2}\, \|\eta(s,r)\|^p_{\mathrm{HS}(U,H)} \Big) \int_0^T (T-\tau)^{-p\vartheta/2}\mathrm{d}\tau < \infty.
\end{aligned}$$

The pointwise convergence in (4.3.3), the boundedness property (4.3.5), and the dominated convergence theorem thus imply that

$$\lim_{t\searrow t_0} \mathbb{E}\int_0^T \mathbb{1}_{[0,t_0)}(\tau)\, \|\eta(\tau,t)-\eta(\tau,t_0)\|^p_{\mathrm{HS}(U,H)}\, \mathrm{d}\tau = 0. \tag{4.3.6}$$

For the second summand in (4.3.4), observe that for all $t \in (t_0, T]$ we have that

$$\begin{aligned}
\mathbb{E}\int_0^T \mathbb{1}_{[t_0,t)}(\tau)\, \|\eta(\tau,t)-\eta(\tau,t_0)\|^p_{\mathrm{HS}(U,H)}\, \mathrm{d}\tau &= \mathbb{E}\int_0^T \mathbb{1}_{[t_0,t)}(\tau)\, \|\eta(\tau,t)\|^p_{\mathrm{HS}(U,H)}\, \mathrm{d}\tau \\
&\le \sup_{s\in[0,T)} \mathbb{E}\Big(\sup_{r\in(s,T]} (r-s)^{p\vartheta/2}\, \|\eta(s,r)\|^p_{\mathrm{HS}(U,H)} \Big) \\
&\qquad \cdot \int_0^T \mathbb{1}_{[t_0,t)}(\tau)(t-\tau)^{-p\vartheta/2}\mathrm{d}\tau \\
&= \sup_{s\in[0,T)} \mathbb{E}\Big(\sup_{r\in(s,T]} (r-s)^{p\vartheta/2}\, \|\eta(s,r)\|^p_{\mathrm{HS}(U,H)} \Big) \\
&\qquad \cdot \frac{(t-t_0)^{1-p\vartheta/2}}{1-p\vartheta/2} \longrightarrow 0 \quad \text{as } t\searrow t_0.
\end{aligned} \tag{4.3.7}$$

This together with (4.3.6) proves the right-continuity in (4.3.2). Now we show the left-continuity: For all $t \in [\frac{t_0}{2}, t_0]$ we have that

$$\begin{aligned} \mathbb{E} \int_0^T \|\eta(\tau,t) - \eta(\tau,t_0)\|^p_{\mathrm{HS}(U,H)} \,\mathrm{d}\tau &= \mathbb{E} \int_0^T \mathbb{1}_{[0,2t-t_0)}(\tau)\, \|\eta(\tau,t) - \eta(\tau,t_0)\|^p_{\mathrm{HS}(U,H)} \,\mathrm{d}\tau \\ &\quad + \mathbb{E} \int_0^T \mathbb{1}_{[2t-t_0,t_0)}(\tau)\, \|\eta(\tau,t) - \eta(\tau,t_0)\|^p_{\mathrm{HS}(U,H)} \,\mathrm{d}\tau. \end{aligned} \tag{4.3.8}$$

Note that, since $t \in [\frac{t_0}{2}, t_0]$, it holds for all $\tau \in [0, 2t - t_0)$ that $\frac{t_0-\tau}{2} \leq t - \tau$. Thus with a similar argument as in (4.3.5), we obtain that

$$\begin{aligned} \mathbb{1}_{[0,2t-t_0)}(\tau)\, \|\eta(\tau,t) - \eta(\tau,t_0)\|^p_{\mathrm{HS}(U,H)} &\leq \mathbb{1}_{[0,2t-t_0)}(\tau)\, 2^{p-1} \Big(\sup_{r\in(\tau,T]} (r-\tau)^{p\vartheta/2} \, \|\eta(\tau,r)\|^p_{\mathrm{HS}(U,H)} \Big) \\ &\quad \cdot \big((t-\tau)^{-p\vartheta/2} + (t_0-\tau)^{-p\vartheta/2} \big) \\ &\leq \mathbb{1}_{[0,2t-t_0)}(\tau)\, 2^{p-1} \Big(\sup_{r\in(\tau,T]} (r-\tau)^{p\vartheta/2} \, \|\eta(\tau,r)\|^p_{\mathrm{HS}(U,H)} \Big) \\ &\quad \cdot \Big(\big(\frac{t_0-\tau}{2}\big)^{-p\vartheta/2} + (t_0-\tau)^{-p\vartheta/2} \Big) \\ &\leq \mathbb{1}_{[0,2t-t_0)}(\tau)\, 2^{p-1} \Big(\sup_{r\in(\tau,T]} (r-\tau)^{p\vartheta/2} \, \|\eta(\tau,r)\|^p_{\mathrm{HS}(U,H)} \Big) \\ &\quad \cdot (1 + 2^{p\vartheta/2})(t_0-\tau)^{-p\vartheta/2} \in L^1(\mathbb{P}\otimes \mathrm{d}\tau; \mathbb{R}). \end{aligned} \tag{4.3.9}$$

The pointwise convergence in (4.3.3), the boundedness property (4.3.9), and the dominated convergence theorem thus imply that

$$\lim_{t\nearrow t_0} \mathbb{E} \int_0^T \mathbb{1}_{[0,2t-t_0)}(\tau)\, \|\eta(\tau,t) - \eta(\tau,t_0)\|^p_{\mathrm{HS}(U,H)} \,\mathrm{d}\tau = 0. \tag{4.3.10}$$

Finally, for the second summand in (4.3.8) observe that for all $t \in [0, t_0)$ it holds that

$$\begin{aligned} &\mathbb{E} \int_0^T \mathbb{1}_{[2t-t_0,t_0)}(\tau)\, \|\eta(\tau,t) - \eta(\tau,t_0)\|^p_{\mathrm{HS}(U,H)} \,\mathrm{d}\tau \\ &\leq 2^{p-1} \sup_{s\in[0,T)} \mathbb{E}\big(\sup_{r\in(s,T]} (r-s)^{p\vartheta/2} \, \|\eta(s,r)\|^p_{\mathrm{HS}(U,H)} \big) \\ &\quad \cdot \int_0^T \Big[\mathbb{1}_{[2t-t_0,t)}(\tau)(t-\tau)^{-p\vartheta/2} + \mathbb{1}_{[2t-t_0,t_0)}(\tau)(t_0-\tau)^{-p\vartheta/2} \Big] \mathrm{d}\tau \\ &= 2^{p-1} \sup_{s\in[0,T)} \mathbb{E}\big(\sup_{r\in(s,T]} (r-s)^{p\vartheta/2} \, \|\eta(s,r)\|^p_{\mathrm{HS}(U,H)} \big) \\ &\quad \cdot \Big(\frac{(t_0-t)^{1-p\vartheta/2}}{1-p\vartheta/2} + \frac{(2t_0-2t)^{1-p\vartheta/2}}{1-p\vartheta/2} \Big) \longrightarrow 0 \quad \text{as } t \nearrow t_0. \end{aligned} \tag{4.3.11}$$

This together with (4.3.10) proves the left-continuity in (4.3.2). Therefore the continuity (4.3.2) is true.

Step III: We show that for all $t_0 \in N$ there exists a sequence $(t_j)_{j\in\mathbb{N}} \subset [0,T] \setminus N$ with $\lim_{j\to\infty} t_j = t_0$ and

$$\forall F \in \mathbb{D}^{1,2}(H)\colon \lim_{j\to\infty} \langle X(t_j), F\rangle_{\mathbb{D}^{1,2}(H)} = \langle X(t_0), F\rangle_{\mathbb{D}^{1,2}(H)}\,. \tag{4.3.12}$$

Indeed, as a consequence of items (a) and (b), we have that

$$\begin{aligned}
\sup_{t\in[0,T]\setminus N} \mathbb{E}\int_0^T \|D_\tau(X(t))\|^2_{\mathrm{HS}(U,H)}\,\mathrm{d}\tau &= \sup_{t\in[0,T]\setminus N} \mathbb{E}\int_0^t \|D_\tau(X(t))\|^2_{\mathrm{HS}(U,H)}\,\mathrm{d}\tau \\
&= \sup_{t\in[0,T]\setminus N} \mathbb{E}\int_0^t \|\eta(\tau,t)\|^2_{\mathrm{HS}(U,H)}\,\mathrm{d}\tau \\
&\le \sup_{t\in[0,T]\setminus N} \int_0^t \mathbb{E}\big[(t-\tau)^\vartheta\,(t-\tau)^{-\vartheta}\,\|\eta(\tau,t)\|^2_{\mathrm{HS}(U,H)}\big]\mathrm{d}\tau \\
&\le \sup_{s\in[0,T)} \mathbb{E}\Big(\sup_{r\in(s,T]} (r-s)^\vartheta\,\|\eta(s,r)\|^2_{\mathrm{HS}(U,H)}\Big)\int_0^t (t-\tau)^{-\vartheta}\mathrm{d}\tau \\
&\le \sup_{s\in[0,T)} \mathbb{E}\Big(\sup_{r\in(s,T]} (r-s)^\vartheta\,\|\eta(s,r)\|^2_{\mathrm{HS}(U,H)}\Big)\int_0^T (T-\tau)^{-\vartheta}\mathrm{d}\tau \\
&< \infty.
\end{aligned}$$

This boundedness property together with Lemma 4.1.2 implies that

$$\sup_{t\in[0,T]\setminus N} \|X(t)\|_{\mathbb{D}^{1,2}(H)} < \infty,$$

which means that the family $X(t)$, $t \in [0,T] \setminus N$, is bounded in $\mathbb{D}^{1,2}(H)$. Next let $t_0 \in N$ be arbitrary and choose $(t_j)_{j\in\mathbb{N}} \subset [0,T] \setminus N$ with $t_j \to t_0$. As $(X(t_j))_{j\in\mathbb{N}}$ is bounded in $\mathbb{D}^{1,2}(H)$, by weak compactness in Hilbert spaces; see, e.g., [57, Theorem III.3.7], there exists a subsequence $(X(t_{j_k}))_{k\in\mathbb{N}}$ and $Y \in \mathbb{D}^{1,2}(H)$ with

$$\forall F \in \mathbb{D}^{1,2}(H)\colon \lim_{j\to\infty} \langle X(t_{j_k}), F\rangle_{\mathbb{D}^{1,2}(H)} = \langle Y, F\rangle_{\mathbb{D}^{1,2}(H)}\,. \tag{4.3.13}$$

Due to Lemma 4.1.2 it follows that

$$\begin{aligned}
\|X(t_0) - Y\|^2_{L^2(\mathbb{P};H)} &= \langle X(t_0) - Y,\, X(t_0) - Y\rangle_{L^2(\mathbb{P};H)} \\
&= \langle X(t_0),\, X(t_0) - Y\rangle_{L^2(\mathbb{P};H)} - \langle Y,\, X(t_0) - Y\rangle_{L^2(\mathbb{P};H)} \\
&= \lim_{k\to\infty} \langle X(t_{j_k}),\, X(t_0) - Y\rangle_{L^2(\mathbb{P};H)} - \lim_{k\to\infty} \langle X(t_{j_k}),\, X(t_0) - Y\rangle_{L^2(\mathbb{P};H)} \\
&= 0,
\end{aligned}$$

i.e., $X(t_0) = Y$ in $L^2(\mathbb{P};H)$. The equality holds even in $\mathbb{D}^{1,2}(H)$ as $X(t_0), Y \in \mathbb{D}^{1,2}(H)$. This and (4.3.13) imply (4.3.12).

Step IV: As an extension of property (a), we show that for **all** $t \in [0,T]$ it holds that

$$\eta(\cdot,t) = D.(X(t)) \qquad \text{in } L^2\big(\mathbb{P}\otimes\lambda; \mathrm{HS}(U,H)\big).$$

Due to property (a) it is enough to verify the equality above for $t \in N$. For this let $t \in N$ be arbitrary. By Step III there exists a sequence $(t_j)_{j\in\mathbb{N}} \subset [0,T]\setminus N$ with $\lim_{j\to\infty} t_j = t$ and

$$\forall F \in \mathbb{D}^{1,2}(H)\colon \lim_{j\to\infty} \langle X(t_j),\, F\rangle_{\mathbb{D}^{1,2}(H)} = \langle X(t),\, F\rangle_{\mathbb{D}^{1,2}(H)}\,.$$

Hence we obtain for all $\Phi \in L^2\big(\mathbb{P}\otimes\lambda; \mathrm{HS}(U,H)\big)$ that

$$\begin{aligned}\lim_{j\to\infty} \big\langle D\big(X(t_j)\big), \Phi\big\rangle_{L^2(\mathbb{P}\otimes\lambda;\mathrm{HS}(U,H))} &= \lim_{j\to\infty} \langle X(t_j), D^*\Phi\rangle_{\mathbb{D}^{1,2}(H)} \\ &= \langle X(t), D^*\Phi\rangle_{\mathbb{D}^{1,2}(H)} = \big\langle D\big(X(t)\big), \Phi\big\rangle_{L^2(\mathbb{P}\otimes\lambda;\mathrm{HS}(U,H))}\,,\end{aligned} \tag{4.3.14}$$

where $D^* \in L\big(L^2(\mathbb{P}\otimes\lambda;\mathrm{HS}(U,H)), \mathbb{D}^{1,2}(H)\big)$ is the Hilbert-adjoint operator of $D \in L\big(\mathbb{D}^{1,2}(H), L^2(\mathbb{P}\otimes\lambda;\mathrm{HS}(U,H))\big)$. On the other hand, for all $\Phi \in L^2\big(\mathbb{P}\otimes\lambda;\mathrm{HS}(U,H)\big)$ it holds that

$$\begin{aligned}\lim_{j\to\infty} \big\langle D\big(X(t_j)\big), \Phi\big\rangle_{L^2(\mathbb{P}\otimes\lambda;\mathrm{HS}(U,H))} &= \lim_{j\to\infty} \big\langle \eta(\cdot,t_j),\, \Phi\big\rangle_{L^2(\mathbb{P}\otimes\lambda;\mathrm{HS}(U,H))} \\ &= \big\langle \eta(\cdot,t),\, \Phi\big\rangle_{L^2(\mathbb{P}\otimes\lambda;\mathrm{HS}(U,H))},\end{aligned} \tag{4.3.15}$$

where the first equality above holds due to property (a) and the second equality due to the continuity property proven in Step II. Combining (4.3.14) and (4.3.15) yields that

$$\eta(\cdot,t) = D.\big(X(t)\big) \qquad \text{in } L^2\big(\mathbb{P}\otimes\lambda; \mathrm{HS}(U,H)\big).$$

Step V: We are now able to verify items (ii) and (iii). First, observe that item (ii) is a consequence of Lemma 4.1.2, Step II, and Step IV. Moreover, note that item (iii) follows from item (c) and Step IV. □

The next result demonstrates the interchange of integration and the Malliavin derivative.

Lemma 4.3.2. Let Assumption 4.1.1 be fulfilled, $X : \Omega\times[0,T] \to H$ be the mild solution of (4.1.1) with continuous sample paths, let μ be a finite Borel measure on $[0,T]$, let $\phi \in L^2(\mu; H)$. Furthermore, let $\mathcal{U} \subset U$ be an ONB, let $u \in \mathcal{U}$, and let $\eta\colon \Omega\times[0,T]\times[0,T] \to \mathrm{HS}(U,H)$ be a measurable mapping which satisfies the properties formulated in item (iii) of Lemma 4.3.1. Then

(i) it holds that

$$\int_{[0,T]} \langle \phi(t),\, X(t)\rangle_H \; \mu(\mathrm{d}t) \in \mathbb{D}^{1,2}(\mathbb{R}),$$

(ii) it holds $\mathbb{P}\otimes \mathrm{d}\tau$-a.e. that

$$\int_{[0,T]} \Big| \big\langle \phi(t), \big(\eta(\tau,t)\big)[u]\big\rangle_H \Big| \, \mu(\mathrm{d}t) < \infty,$$

and

(iii) it holds $\mathbb{P}\otimes \mathrm{d}\tau$-a.e. that

$$\Big(D_\tau \int_{[0,T]} \langle \phi(t),\, X(t)\rangle_H \;\mu(\mathrm{d}t)\Big)[u] = \int_{[0,T]} \big\langle \phi(t), \big(\eta(\tau,t)\big)[u]\big\rangle_H \;\mu(\mathrm{d}t).$$

Proof. We start with item (ii). For this, note that

$$\begin{aligned}
\mathbb{E}\int_0^T & \Big(\int_{[0,T]} \Big|\big\langle \phi(t),\, \big(\eta(\tau,t)\big)[u]\big\rangle_H\Big| \mu(\mathrm{d}t)\Big)\mathrm{d}\tau \\
&\le T\, \|\phi\|^2_{L^2(\mu;H)}\; \mathbb{E}\int_0^T \int_{[0,T]} \Big\|\big(\eta(\tau,t)\big)[u]\Big\|_H^2 \mu(\mathrm{d}t)\mathrm{d}\tau \\
&\le T\, \|\phi\|^2_{L^2(\mu;H)} \int_{[0,T]}\int_0^t (t-\tau)^{-\vartheta}\, \mathbb{E}\Big((t-\tau)^{\vartheta}\, \big\|\big(\eta(\tau,t)\big)\big\|^2_{\mathrm{HS}(U,H)}\Big)\, \mathrm{d}\tau\, \mu(\mathrm{d}t) \qquad (4.3.16)\\
&\le T\, \|\phi\|^2_{L^2(\mu;H)}\; \mu([0,T]) \Big[\sup_{s\in[0,T)} \mathbb{E}\big(\sup_{r\in(s,T]} (r-s)^{\vartheta}\, \|\eta(s,r)\|^2_{\mathrm{HS}(U,H)}\big)\Big] \\
&\quad \cdot \int_0^T (T-\tau)^{-\vartheta}\mathrm{d}\tau < \infty.
\end{aligned}$$

This proves item (ii). Next note that due to the predictability of $\big(X(t)\big)_{t\in[0,T]}$, the mapping $\Omega\times[0,T]\ni(\omega,t)\mapsto\langle\phi(t),X(\omega,t)\rangle_H$ is measurable. Moreover, observe that

$$\begin{aligned}
\mathbb{E}\Big(\Big(\int_{[0,T]} \big|\,\langle \phi(t),\, X(t)\rangle_H\,\big|\, \mu(\mathrm{d}t)\Big)^2\Big) &\le \mathbb{E}\Big(\int_{[0,T]} \|\phi(t)\|_H^2\; \mu(\mathrm{d}t) \int_{[0,T]} \|X(t)\|_H^2\; \mu(\mathrm{d}t)\Big) \\
&= \|\phi\|^2_{L^2(\mu;H)}\, \mathbb{E}\Big(\int_{[0,T]} \|X(t)\|_H^2\; \mu(\mathrm{d}t)\Big) \\
&= \|\phi\|^2_{L^2(\mu;H)} \int_{[0,T]} \mathbb{E}\big(\,\|X(t)\|_H^2\,\big)\mu(\mathrm{d}t) \qquad (4.3.17)\\
&\le \|\phi\|^2_{L^2(\mu;H)} \int_{[0,T]} \Big(\sup_{s\in[0,T]} \mathbb{E}\big(\,\|X(s)\|_H^2\,\big)\Big)\mu(\mathrm{d}t) \\
&= \mu\big([0,T]\big)\, \|\phi\|^2_{L^2(\mu;H)} \sup_{s\in[0,T]} \mathbb{E}\big(\,\|X(s)\|_H^2\,\big) < \infty,
\end{aligned}$$

where the last equality holds due to Lemma 4.1.2. On the other hand, due to the path-wise continuity of X we know that $\langle\phi,\, X(\omega)\rangle_{L^2(\mu;H)}$ exists for every $\omega\in\Omega$. Now Fubini's theorem implies that $\langle\phi,\, X\rangle_{L^2(\mu;H)}\in L^2(\mathbb{P};\mathbb{R})$. For $n\in\mathbb{N}$, let $t_j=\frac{jT}{n}$, $j\in\{0,\dots,n\}$ and let

$\big(X_n(t)\big)_{t\in[0,T]}$ be defined as

$$X_n(t) = \sum_{j=0}^{n-1} \mathbb{1}_{[t_j,t_{j+1})}(t)X(t_j) + \mathbb{1}_{\{T\}}(t)X(T).$$

Then it holds by Hölder's inequality that

$$\begin{aligned}
&\left\| \int_{[0,T]} \langle \phi(t),\, X_n(t)\rangle_H \,\mu(\mathrm{d}t) - \int_{[0,T]} \langle \phi(t),\, X(t)\rangle_H \,\mu(\mathrm{d}t) \right\|^2_{L^2(\mathbb{P};\mathbb{R})} \\
&= \mathbb{E}\bigg(\Big(\int_{[0,T]} \big\langle \phi(t),\, X_n(t) - X(t) \big\rangle_H \mu(\mathrm{d}t) \Big)^2 \bigg) \\
&\le \mathbb{E}\bigg(\int_{[0,T]} \|\phi(t)\|_H^2 \,\mu(\mathrm{d}t) \cdot \int_{[0,T]} \|X_n(t) - X(t)\|_H^2 \,\mu(\mathrm{d}t) \bigg) \\
&\le \|\phi\|^2_{L^2(\mu;H)} \int_{[0,T]} \mathbb{E}\Big(\|X_n(t) - X(t)\|_H^2 \Big) \mu(\mathrm{d}t) \longrightarrow 0 \qquad \text{as} \qquad n \to \infty.
\end{aligned} \tag{4.3.18}$$

Indeed, the convergence holds due to the definition of X_n and the uniform continuity of the mapping $[0,T] \ni t \to X(t) \in L^2(\mathbb{P};H)$; see Lemma 4.1.2. Moreover it holds that

$$\begin{aligned}
\int_{[0,T]} \langle \phi(t),\, X_n(t)\rangle_H \,\mu(\mathrm{d}t) &= \sum_{j=0}^{n-1} \int_{[t_j,t_{j+1})} \langle \phi(t),\, X(t_j)\rangle_H \,\mu(\mathrm{d}t) + \mu(\{T\}) \,\langle \phi(T),\, X(T)\rangle_H \\
&= \sum_{j=0}^{n-1} \left\langle \int_{[t_j,t_{j+1})} \phi(t)\mu(\mathrm{d}t),\, X(t_j) \right\rangle_H + \mu(\{T\}) \,\langle \phi(T),\, X(T)\rangle_H \,.
\end{aligned}$$

Hence, by Lemma 4.3.1(i), the fact that $\int_{[0,T]} \langle \phi(t),\, X_n(t)\rangle_H \,\mu(\mathrm{d}t) \in \mathbb{D}^{1,2}(\mathbb{R})$, and due to the general properties of the Malliavin derivative it holds $\mathbb{P} \otimes \mathrm{d}\tau$-a.e. that

$$\begin{aligned}
\bigg(D_\tau\Big(\int_{[0,T]} \langle \phi(t),\, X_n(t)\rangle_H \,\mu(\mathrm{d}t) \Big)\bigg)[u] &= \sum_{j=0}^{n-1} \left\langle \int_{[t_j,t_{j+1})} \phi(t)\mu(\mathrm{d}t),\, \Big(D_\tau\big(X(t_j)\big)\Big)[u] \right\rangle_H \\
&\quad + \mu(\{T\}) \Big\langle \phi(T),\, \Big(D_\tau\big(X(T)\big)\Big)[u] \Big\rangle_H \\
&= \sum_{j=0}^{n-1} \int_{[t_j,t_{j+1})} \Big\langle \phi(t),\, \Big(D_\tau\big(X(t_j)\big)\Big)[u] \Big\rangle_H \,\mu(\mathrm{d}t) \\
&\quad + \mu(\{T\}) \Big\langle \phi(T),\, \Big(D_\tau\big(X(T)\big)\Big)[u] \Big\rangle_H \\
&= \int_{[0,T]} \left\langle \phi(t),\, \sum_{j=0}^{n-1} \mathbb{1}_{[t_j,t_{j+1})}(t) \Big(D_\tau\big(X(t_j)\big)\Big)[u] \right\rangle_H \mu(\mathrm{d}t) \\
&\quad + \mu(\{T\}) \Big\langle \phi(T),\, \Big(D_\tau\big(X(T)\big)\Big)[u] \Big\rangle_H \,.
\end{aligned}$$

To complete the proof of this lemma it remains to show that

$$\lim_{n\to\infty}\Big(D_\tau\Big(\int_{[0,T]}\langle\phi(t),X_n(t)\rangle_H\,\mu(\mathrm{d}t)\Big)\Big)=\int_{[0,T]}\big\langle\phi(t),\,\big(\eta(\tau,t)\big)[\cdot]\big\rangle_H\,\mu(\mathrm{d}t)\quad\text{in } L^2(\mathbb{P}\otimes\mathrm{d}\tau;\mathrm{HS}(U,\mathbb{R})). \tag{4.3.19}$$

We show first that $\int_{[0,T]}\langle\phi(t),(\eta(\tau,t))[\cdot]\rangle_H\,\mu(\mathrm{d}t)\in L^2(\mathbb{P}\otimes\mathrm{d}\tau;\mathrm{HS}(U,\mathbb{R}))$. Indeed we have analogously to (4.3.16) that

$$\begin{aligned}&\mathbb{E}\int_0^T\sum_{u\in\mathcal{U}}\Big|\int_{[0,T]}\big\langle\phi(t),\big(\eta(\tau,t)\big)[u]\big\rangle_H\,\mu(\mathrm{d}t)\Big|^2\mathrm{d}\tau\\&\le\mathbb{E}\int_0^T\sum_{u\in\mathcal{U}}\Big(\|\phi\|^2_{L^2(\mu;H)}\cdot\int_{[0,T]}\Big\|\big(\eta(\tau,t)\big)[u]\Big\|^2_H\mu(\mathrm{d}t)\Big)\mathrm{d}\tau\\&=\|\phi\|^2_{L^2(\mu;H)}\int_{[0,T]}\int_0^T\Big(\mathbb{E}\sum_{u\in\mathcal{U}}\Big\|\big(\eta(\tau,t)\big)[u]\Big\|^2_H\Big)\mathrm{d}\tau\mu(\mathrm{d}t)<\infty.\end{aligned} \tag{4.3.20}$$

Next note that the definition of X_n and the (uniform) continuity in Lemma 4.3.1(ii) imply that

$$\begin{aligned}&\Big\|D_\cdot\Big(\int_{[0,T]}\langle\phi(t),X_n(t)\rangle_H\,\mu(\mathrm{d}t)\Big)-\int_{[0,T]}\big\langle\phi(t),\big(\eta(\cdot,t)\big)[\cdot]\big\rangle_H\,\mu(\mathrm{d}t)\Big\|^2_{L^2(\mathbb{P}\otimes\lambda;\mathrm{HS}(U,\mathbb{R}))}\\&=\mathbb{E}\int_0^T\sum_{u\in\mathcal{U}}\Big|\int_{[0,T]}\Big\langle\phi(t),\sum_{j=0}^{n-1}\mathbb{1}_{[t_j,t_{j+1})}(t)\Big(D_\tau\big(X(t_j)\big)-\eta(\tau,t)\Big)[u]\Big\rangle_H\,\mu(\mathrm{d}t)\Big|^2\,\mathrm{d}\tau\\&\le\mathbb{E}\int_0^T\sum_{u\in\mathcal{U}}\Big(\int_{[0,T]}\|\phi(s)\|_H^2\,\mu(\mathrm{d}s)\cdot\int_{[0,T]}\sum_{j=0}^{n-1}\mathbb{1}_{[t_j,t_{j+1})}(t)\,\Big\|\Big(D_\tau\big(X(t_j)\big)-\eta(\tau,t)\Big)[u]\Big\|_H^2\,\mu(\mathrm{d}t)\Big)\,\mathrm{d}\tau\\&=\|\phi\|^2_{L^2(\mu;H)}\,\mathbb{E}\int_0^T\sum_{u\in\mathcal{U}}\sum_{j=0}^{n-1}\int_{[t_j,t_{j+1})}\Big\|\Big(D_\tau\big(X(t_j)\big)-\eta(\tau,t)\Big)[u]\Big\|_H^2\,\mu(\mathrm{d}t)\mathrm{d}\tau\\&=\|\phi\|^2_{L^2(\mu;H)}\int_{[0,T]}\sum_{j=0}^{n-1}\mathbb{1}_{[t_j,t_{j+1})}(t)\,\mathbb{E}\int_0^T\big\|D_\tau\big(X(t_j)\big)-\eta(\tau,t)\big\|^2_{\mathrm{HS}(U,H)}\,\mathrm{d}\tau\,\mu(\mathrm{d}t)\\&=\|\phi\|^2_{L^2(\mu;H)}\int_{[0,T]}\sum_{j=0}^{n-1}\mathbb{1}_{[t_j,t_{j+1})}(t)\,\mathbb{E}\int_0^T\|\eta(\tau,t_j)-\eta(\tau,t)\|^2_{\mathrm{HS}(U,H)}\,\mathrm{d}\tau\,\mu(\mathrm{d}t)\longrightarrow 0.\end{aligned}$$

For the uniform continuity of the mapping $[0,T]\ni t\mapsto\eta(\cdot,t)\in L^2(\mathbb{P}\otimes\lambda;\mathrm{HS}(U,H))$ see Step II in the proof of Lemma 4.3.1. The combination of (4.3.18), (4.3.19), and the closedness of the Malliavin derivative prove items (i) and (iii). □

Our next lemma introduces a useful isometry which we use later in this chapter to investigate the situation when we apply the derivative operator on the square-integrable sample paths of mild solutions of SEEs.

Lemma 4.3.3. Let $T\in(0,\infty)$ and μ be a finite Borel measure on $[0,T]$. Let U,H be separable real Hilbert spaces and $\mathcal{U}$ be an orthonormal basis of U. Then there exists an isometric isomorphism $i\colon \mathrm{HS}(U, L^2(\mu; H)) \to L^2(\mu; \mathrm{HS}(U,H))$ such that for any $\Psi \in \mathrm{HS}(U, L^2(\mu; H))$ it holds that

$$\forall u \in \mathcal{U} : \Big(\big(i(\Psi)(r)\big)[u] = \big(\Psi[u]\big)(r), \qquad \mu\text{-a.e. } r \in [0,T] \Big). \tag{4.3.21}$$

Proof. Without loss of generality we assume that the Hilbert space U is infinite-dimensional and that $(u_j)_{j\in\mathbb{N}}$ is a counting for the orthonormal basis $\mathcal{U}$. First we show that the set $\mathcal{H} = \big\{ \sum_{j=1}^n \phi_j \otimes g_j \otimes h_j \colon \phi_j \in \mathrm{HS}(U,\mathbb{R}), g_j \in L^2(\mu;\mathbb{R}), h_j \in H, n \in \mathbb{N} \big\}$ is dense in $\mathrm{HS}(U, L^2(\mu; H))$. For this let $\epsilon > 0$ be given. For $\Psi \in \mathrm{HS}(U, L^2(\mu; H))$ let $n \in \mathbb{N}$ be such that

$$\Big\| \Psi - \sum_{j=1}^n \langle u_j, \cdot \rangle_U \otimes \Psi[u_j] \Big\|_{\mathrm{HS}(U, L^2(\mu;H))} < \frac{\epsilon}{2}. \tag{4.3.22}$$

Moreover note that, due to [51, Lemma A.1.4] and an application of the dominated convergence theorem, we have that for each $\Psi[u_j] \in L^2(\mu; H)$, $j \in \{1, \dots, n\}$, there exist $g_k^j \in L^2(\mu;\mathbb{R})$, $h_k^j \in H$, and $n_j \in \mathbb{N}$, $k \in \{1, \dots, n_j\}$, such that

$$\Big\| \Psi[u_j] - \sum_{k=1}^{n_j} g_k^j \otimes h_k^j \Big\|_{L^2(\mu;H)} < \frac{\epsilon}{2n}. \tag{4.3.23}$$

Now note that using the triangular inequality, (4.3.22) and (4.3.23) we obtain that

$$\begin{aligned}
&\Big\| \Psi - \sum_{j=1}^n \sum_{k=1}^{n_j} \langle u_j, \cdot \rangle_U \otimes g_k^j \otimes h_k^j \Big\|_{\mathrm{HS}(U, L^2(\mu;H))} \\
&\le \Big\| \Psi - \sum_{j=1}^n \langle u_j, \cdot \rangle_U \otimes \Psi[u_j] \Big\|_{\mathrm{HS}(U, L^2(\mu;H))} \\
&\quad + \Big\| \sum_{j=1}^n \langle u_j, \cdot \rangle_U \otimes \Psi[u_j] - \sum_{j=1}^n \sum_{k=1}^{n_j} \langle u_j, \cdot \rangle_U \otimes g_k^j \otimes h_k^j \Big\|_{\mathrm{HS}(U, L^2(\mu;H))} \\
&< \frac{\epsilon}{2} + \Big\| \sum_{j=1}^n \langle u_j, \cdot \rangle_U \otimes \Big(\Psi[u_j] - \sum_{k=1}^{n_j} g_k^j \otimes h_k^j \Big) \Big\|_{\mathrm{HS}(U, L^2(\mu;H))} \\
&\le \frac{\epsilon}{2} + \sum_{j=1}^n \Big\| \Psi[u_j] - \sum_{k=1}^{n_j} g_k^j \otimes h_k^j \Big\|_{L^2(\mu;H)} \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,
\end{aligned}$$

which proves that the set $\mathcal{H}$ is dense in $\mathrm{HS}(U, L^2(\mu; H))$. Let the mapping $i\colon \mathcal{H} \to L^2(\mu; \mathrm{HS}(U,H))$ be defined as

$$i(\phi \otimes g \otimes h) = g \otimes \phi \otimes h, \qquad \text{for } \phi \in \mathrm{HS}(U,\mathbb{R}), g \in L^2(\mu;\mathbb{R}), h \in H.$$

Obviously the mapping i is continuous and injective and for each element of $\mathcal{H}$ the equality (4.3.21) holds. Therefore it can be extended to an isomorphism between $\mathrm{HS}(U, L^2(\mu; H))$ and $L^2(\mu; \mathrm{HS}(U,H))$, considering the fact that the image space $i(\mathcal{H})$ is also dense in $L^2(\mu; \mathrm{HS}(U,H))$.

Now let $\Psi \in \mathrm{HS}(U, L^2(\mu; H))$ be arbitrary. Then there exists a sequence $\big(\Psi_n\big)_{n\in\mathbb{N}} \subset \mathcal{H}$ such that it converges to Ψ in $\mathrm{HS}(U, L^2(\mu; H))$. In particular,we obtain that

$$\forall u \in \mathcal{U}\colon \qquad \Big\| \Psi[u] - \Psi_{n_k}[u] \Big\|_{L^2(\mu;H)} \to 0, \quad \text{as } k \to \infty.$$

Therefore it holds that

$$\forall u \in \mathcal{U}\colon \quad \bigg(\exists (n_k^u)_{k\in\mathbb{N}} \subset \mathbb{N}\colon \quad \Psi[u](r) = \lim_{k\to\infty} \Psi_{n_k^u}[u](r) \quad \mu\text{-a.e. } r \in [0,T] \bigg).$$

Using a diagonal procedure for choosing above subsequences and due to countability of $\mathcal{U}$ one can, without loss of generality, assume that there exists a sequence $\big(\Psi_n\big)_{n\in\mathbb{N}} \subset \mathcal{H}$ such that for all $u \in \mathcal{U}$ and μ-a.e. $r \in [0,T]$ it holds that

$$\Psi[u](r) = \lim_{n\to\infty} \Psi_n[u](r).$$

On the other hand due to construction of the mapping i we have that $i(\Psi_n)$ converges to $i(\Psi)$ in $L^2(\mu; \mathrm{HS}(U,H))$ and one can analogously show (after passing to a subsequence) that for all $u \in \mathcal{U}$ and μ-a.e. $r \in [0,T]$ it holds that

$$i\big(\Psi\big)(r)[u] = \lim_{n\to\infty} i\big(\Psi_n\big)(r)[u].$$

Considering the fact that for all $u \in \mathcal{U}$ and μ-a.e. $r \in [0,T]$, $\Psi_n[u](r) = i\big(\Psi_n\big)(r)[u]$ we obtain for all $u \in \mathcal{U}$ and μ-a.e. $r \in [0,T]$ that

$$\Psi[u](r) = i\big(\Psi\big)(r)[u].$$

This together with the fact that the mapping i is an isometric isomorphism completes the proof. □

The next lemma presents a relation between the Malliavin derivatives of mild solutions of SEEs and the Malliavin derivative of its square-integrable sample paths.

Lemma 4.3.4. Let Assumption 4.1.1 be fulfilled, let $X : \Omega \times [0,T] \to H$ be the mild solution of (4.1.1) with continuous sample paths, let $\eta\colon \Omega \times [0,T] \times [0,T] \to H$ be a measurable mapping which satisfies the properties formulated in item (iii) of Lemma 4.3.1, and let μ and $i\colon \mathrm{HS}(U, L^2(\mu; H)) \to L^2(\mu; \mathrm{HS}(U,H))$ be the finite Borel measure and the isometric isomorphism introduced in Lemma 4.3.3, respectively. Then

(i) it holds that $X \in \mathbb{D}^{1,2}(L^2(\mu; H))$ and

(ii) the following equality holds $\mathbb{P} \otimes \mathrm{d}\tau$-a.e. as an equality in $L^2(\mu(\mathrm{d}t); \mathrm{HS}(U,H))$:

$$\big(i(D_\tau X)\big)(t) = \eta(\tau, t).$$

Proof. Throughout this proof let $(\phi_{jk})_{j\in J,k\in K}$ be an orthonormal basis of $L^2(\mu;H)$ of the form $\phi_{jk}=\psi_j\otimes e_k$, where $(\psi_j)_{j\in J}$ and $(e_k)_{k\in K}$ are orthonormal bases of $L^2(\mu;\mathbb{R})$ and H, respectively. To simplify notation, we restrict the exposition to the case where both $L^2(\mu;\mathbb{R})$ and H are infinite-dimensional spaces and assume that $J=K=\mathbb{N}$. We remark that the proof can be extended to other cases in a straightforward way. First we show that

$$X=\lim_{n\to\infty}\sum_{j=1}^{n}\sum_{k=1}^{n}\langle\phi_{jk},X\rangle_{L^2(\mu;H)}\,\phi_{jk}\qquad\text{in } L^2(\mathbb{P};L^2(\mu;H)).\tag{4.3.24}$$

For this note that the above equality holds for every $\omega\in\Omega$ in $L^2(\mu;H)$ due to the continuity of $\big(X(t)\big)_{t\in[0,T]}$. Therefore it is enough to show that the sequence $\big(\sum_{j=1}^{n}\sum_{k=1}^{n}\langle\phi_{jk},X\rangle_{L^2(\mu;H)}\,\phi_{jk}\big)_{n\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\mathbb{P};L^2(\mu;H))$. Indeed, for all $m,n\in\mathbb{N}$ with $m\le n$ it holds that

$$\begin{aligned}
&\Big\|\Big(\sum_{j=1}^{n}\sum_{k=1}^{n}-\sum_{j=1}^{m}\sum_{k=1}^{m}\Big)\langle\phi_{jk},X\rangle_{L^2(\mu;H)}\,\phi_{jk}\Big\|^2_{L^2(\mathbb{P};L^2(\mu;H))}\\
&=\Big\|\Big(\sum_{j=1}^{n}\sum_{k=m+1}^{n}+\sum_{j=m+1}^{n}\sum_{k=1}^{m}\Big)\langle\phi_{jk},X\rangle_{L^2(\mu;H)}\,\phi_{jk}\Big\|^2_{L^2(\mathbb{P};L^2(\mu;H))}\\
&\le 2\Big\|\sum_{j=1}^{n}\sum_{k=m+1}^{n}\langle\phi_{jk},X\rangle_{L^2(\mu;H)}\,\phi_{jk}\Big\|^2_{L^2(\mathbb{P};L^2(\mu;H))}+2\Big\|\sum_{j=m+1}^{n}\sum_{k=1}^{m}\langle\phi_{jk},X\rangle_{L^2(\mu;H)}\,\phi_{jk}\Big\|^2_{L^2(\mathbb{P};L^2(\mu;H))}\\
&=2\,\mathbb{E}\Big(\Big\|\sum_{k=m+1}^{n}\sum_{j=1}^{n}\langle\phi_{jk},X\rangle_{L^2(\mu;H)}\,\phi_{jk}\Big\|^2_{L^2(\mu;H)}\Big)+2\,\mathbb{E}\Big(\Big\|\sum_{k=1}^{m}\sum_{j=m+1}^{n}\langle\phi_{jk},X\rangle_{L^2(\mu;H)}\,\phi_{jk}\Big\|^2_{L^2(\mu;H)}\Big)\\
&=2\,\mathbb{E}\sum_{k=m+1}^{n}\sum_{j=1}^{n}\Big(\langle\phi_{jk},X\rangle_{L^2(\mu;H)}\Big)^2+2\,\mathbb{E}\sum_{k=1}^{m}\sum_{j=m+1}^{n}\Big(\langle\phi_{jk},X\rangle_{L^2(\mu;H)}\Big)^2\\
&\le 2\,\mathbb{E}\sum_{k=m+1}^{n}\sum_{j=1}^{\infty}\Big(\int_{[0,T]}\psi_j(t)\,\langle e_k,X(t)\rangle_H\;\mu(\mathrm{d}t)\Big)^2\\
&\quad+2\,\mathbb{E}\sum_{k=1}^{\infty}\sum_{j=m+1}^{n}\Big(\int_{[0,T]}\psi_j(t)\,\langle e_k,X(t)\rangle_H\;\mu(\mathrm{d}t)\Big)^2.
\end{aligned}\tag{4.3.25}$$

Now we continue with the above calculation by applying Parseval's identity and two times the dominated convergence theorem to obtain that

$$\begin{aligned}
&\Big\| \Big(\sum_{j=1}^{n} \sum_{k=1}^{n} - \sum_{j=1}^{m} \sum_{k=1}^{m} \Big) \left\langle \phi_{jk}, X \right\rangle_{L^2(\mu;H)} \phi_{jk} \Big\|^2_{L^2(\mathbb{P};L^2(\mu;H))} \\
&\quad \leq 2\,\mathbb{E} \sum_{k=m+1}^{n} \int_{[0,T]} \langle e_k, X(t) \rangle_H^2 \,\mu(\mathrm{d}t) + 2\,\mathbb{E} \sum_{k=1}^{\infty} \sum_{j=m+1}^{n} \Big(\int_{[0,T]} \psi_j(t) \langle e_k, X(t) \rangle_H \,\mu(\mathrm{d}t) \Big)^2 \\
&\quad \leq 2\,\mathbb{E} \int_{[0,T]} \sum_{k=m+1}^{n} \langle e_k, X(t) \rangle_H^2 \,\mu(\mathrm{d}t) + 2\,\mathbb{E} \sum_{k=1}^{\infty} \sum_{j=m+1}^{n} \Big(\int_{[0,T]} \psi_j(t) \langle e_k, X(t) \rangle_H \,\mu(\mathrm{d}t) \Big)^2 \\
&\qquad \longrightarrow 0, \quad \text{as } m, n \longrightarrow \infty.
\end{aligned} \tag{4.3.26}$$

Indeed it holds that

$$\begin{aligned}
&\sum_{k=m+1}^{n} \langle e_k, X(t) \rangle_H^2 \leq \|X(t)\|_H^2 \in L^1\big(\mathbb{P} \otimes \mu(\mathrm{d}t); \mathbb{R}\big), \qquad \text{and that} \\
&\sum_{j=m+1}^{n} \Big(\int_{[0,T]} \psi_j(t) \langle e_k, X(t) \rangle_H \,\mu(\mathrm{d}t) \Big)^2 \leq \int_{[0,T]} \langle e_k, X(t) \rangle_H^2 \,\mu(\mathrm{d}t) \in L^1\big(\mathbb{P} \otimes (\sum_{k=1}^{\infty} \delta_k); \mathbb{R}\big),
\end{aligned}$$

and that the corresponding terms converge pointwise. This proves (4.3.24). On the other hand, for all $j, k \in \mathbb{N}$ and $u \in U$ it holds by Lemma 4.3.2(iii) $\mathbb{P} \otimes \mathrm{d}\tau$-a.e. that

$$\Big(D_\tau \big(\left\langle \phi_{jk}, X \right\rangle_{L^2(\mu;H)} \big) \Big)[u] = \int_{[0,T]} \big\langle \phi_{jk}(t), \big(\eta(\tau, t)\big)[u] \big\rangle_H \,\mu(\mathrm{d}t). \tag{4.3.27}$$

Using this we show next that $X \in \mathbb{D}^{1,2}(L^2(\mu; H))$ and

$$D_\tau X = \lim_{n \to \infty} \sum_{j=1}^{n} \sum_{k=1}^{n} D_\tau \Big(\left\langle \phi_{jk}, X \right\rangle_{L^2(\mu;H)} \Big) \phi_{jk}, \qquad \text{in } L^2\big(\mathbb{P} \otimes \mathrm{d}\tau; \mathrm{HS}(U, L^2(\mu; H))\big). \tag{4.3.28}$$

To this end, note that because of (4.3.24) and the closedness of the operator D, it is enough to show that the sequence $\sum_{j=1}^{n} \sum_{k=1}^{n} \Big(D\Big(\left\langle \phi_{jk}, X \right\rangle_{L^2(\mu;H)} \Big) \phi_{jk} \Big)$, $n \in \mathbb{N}$, is a Cauchy sequence in $L^2\big(\mathbb{P} \otimes$

$\lambda; \mathrm{HS}(U, L^2(\mu;H))\big)$. For this let $m, n \in \mathbb{N}$ with $m \le n$ and observe that

$$\begin{aligned}
&\Big\| \Big(\sum_{j=1}^{n} \sum_{k=1}^{n} - \sum_{j=1}^{m} \sum_{k=1}^{m} \Big) D\Big(\langle \phi_{jk}, X \rangle_{L^2(\mu;H)} \Big) \phi_{jk} \Big\|^2_{L^2(\mathbb{P}\otimes\lambda; \mathrm{HS}(U, L^2(\mu;H)))} \\
&= \Big\| \Big(\sum_{j=1}^{n} \sum_{k=m+1}^{n} + \sum_{j=m+1}^{n} \sum_{k=1}^{m} \Big) D\Big(\langle \phi_{jk}, X \rangle_{L^2(\mu;H)} \Big) \phi_{jk} \Big\|^2_{L^2(\mathbb{P}\otimes\lambda; \mathrm{HS}(U, L^2(\mu;H)))} \\
&\le 2 \Big\| \sum_{j=1}^{n} \sum_{k=m+1}^{n} D\Big(\langle \phi_{jk}, X \rangle_{L^2(\mu;H)} \Big) \phi_{jk} \Big\|^2_{L^2(\mathbb{P}\otimes\lambda; \mathrm{HS}(U, L^2(\mu;H)))} \\
&\quad + 2 \Big\| \sum_{j=m+1}^{n} \sum_{k=1}^{m} D\Big(\langle \phi_{jk}, X \rangle_{L^2(\mu;H)} \Big) \phi_{jk} \Big\|^2_{L^2(\mathbb{P}\otimes\lambda; \mathrm{HS}(U, L^2(\mu;H)))} \\
&= 2\mathbb{E} \int_0^T \Big(\sum_{u\in\mathcal{U}} \Big\| \sum_{j=1}^{n} \sum_{k=m+1}^{n} \Big(D_\tau \langle \phi_{jk}, X \rangle_{L^2(\mu;H)} \Big)[u]\, \phi_{jk} \Big\|^2_{L^2(\mu;H)} \Big) \,\mathrm{d}\tau \\
&\quad + 2\mathbb{E} \int_0^T \Big(\sum_{u\in\mathcal{U}} \Big\| \sum_{j=m+1}^{n} \sum_{k=1}^{m} \Big(D_\tau \langle \phi_{jk}, X \rangle_{L^2(\mu;H)} \Big)[u]\, \phi_{jk} \Big\|^2_{L^2(\mu;H)} \Big) \,\mathrm{d}\tau \\
&= 2\mathbb{E} \int_0^T \Big(\sum_{u\in\mathcal{U}} \sum_{j=1}^{n} \sum_{k=m+1}^{n} \Big| \Big(D_\tau \langle \phi_{jk}, X \rangle_{L^2(\mu;H)} \Big)[u] \Big|^2 \Big) \,\mathrm{d}\tau \\
&\quad + 2\mathbb{E} \int_0^T \Big(\sum_{u\in\mathcal{U}} \sum_{j=m+1}^{n} \sum_{k=1}^{m} \Big| \Big(D_\tau \langle \phi_{jk}, X \rangle_{L^2(\mu;H)} \Big)[u] \Big|^2 \Big) \,\mathrm{d}\tau.
\end{aligned} \tag{4.3.29}$$

Now we continue with the above calculation by applying (4.3.27) and two times the dominated convergence theorem to obtain that

$$\begin{aligned}
&\Big\| \sum_{j,k=1}^{n} D\Big(\langle \phi_{jk}, X \rangle_{L^2(\mu;H)} \Big) \phi_{jk} - \sum_{j,k=1}^{m} D\Big(\langle \phi_{jk}, X \rangle_{L^2(\mu;H)} \Big) \phi_{jk} \Big\|^2_{L^2(\mathbb{P}\otimes\lambda; \mathrm{HS}(U, L^2(\mu;H)))} \\
&\le 2\mathbb{E} \int_0^T \sum_{u\in\mathcal{U}} \sum_{j=1}^{n} \sum_{k=m+1}^{n} \Big(\int_{[0,T]} \langle \phi_{jk}(t),\, (\eta(\tau,t))[u] \rangle_H \,\mu(\mathrm{d}t) \Big)^2 \mathrm{d}\tau \\
&\quad + 2\mathbb{E} \int_0^T \sum_{u\in\mathcal{U}} \sum_{j=m+1}^{n} \sum_{k=1}^{m} \Big(\int_{[0,T]} \langle \phi_{jk}(t),\, (\eta(\tau,t))[u] \rangle_H \,\mu(\mathrm{d}t) \Big)^2 \mathrm{d}\tau \\
&\le 2\mathbb{E} \int_0^T \sum_{u\in\mathcal{U}} \sum_{k=m+1}^{n} \sum_{j=1}^{\infty} \Big(\int_{[0,T]} \psi_j(t) \langle e_k,\, (\eta(\tau,t))[u] \rangle_H \,\mu(\mathrm{d}t) \Big)^2 \mathrm{d}\tau \\
&\quad + 2\mathbb{E} \int_0^T \sum_{u\in\mathcal{U}} \sum_{j=m+1}^{n} \sum_{k=1}^{\infty} \Big(\int_{[0,T]} \psi_j(t) \langle e_k,\, (\eta(\tau,t))[u] \rangle_H \,\mu(\mathrm{d}t) \Big)^2 \mathrm{d}\tau \\
&\le 2\mathbb{E} \int_0^T \sum_{u\in\mathcal{U}} \sum_{k=m+1}^{n} \int_{[0,T]} \langle e_k,\, (\eta(\tau,t))[u] \rangle_H^2 \,\mu(\mathrm{d}t)\,\mathrm{d}\tau \\
&\quad + 2\mathbb{E} \int_0^T \sum_{u\in\mathcal{U}} \sum_{k=1}^{\infty} \sum_{j=m+1}^{n} \Big(\int_{[0,T]} \psi_j(t) \langle e_k,\, (\eta(\tau,t))[u] \rangle_H \,\mu(\mathrm{d}t) \Big)^2 \mathrm{d}\tau \\
&\quad \longrightarrow 0, \quad \text{as } m, n \longrightarrow \infty.
\end{aligned} \tag{4.3.30}$$

Indeed, the corresponding pointwise convergences are clear and for the boundedness of the first summand above note that

$$\begin{aligned}\sum_{k=m+1}^{n}\int_{[0,T]}\left\langle e_k,\,\big(\eta(\tau,t)\big)[u]\right\rangle_H^2\,\mu(\mathrm{d}t) &= \int_{[0,T]}\sum_{k=m+1}^{n}\left\langle e_k,\,\big(\eta(\tau,t)\big)[u]\right\rangle_H^2\,\mu(\mathrm{d}t)\\ &\le \int_{[0,T]}\left\|\big(\eta(\tau,t)\big)[u]\right\|_H^2\mu(\mathrm{d}t) \quad \text{and}\end{aligned}$$

$$\mathbb{E}\int_0^T\sum_{u\in\mathcal{U}}\int_{[0,T]}\left\|\big(\eta(\tau,t)\big)[u]\right\|_H^2\mu(\mathrm{d}t)\mathrm{d}\tau<\infty,$$

similarly to (4.3.20). For the second summand in (4.3.30) note that

$$\sum_{j=m+1}^{n}\left(\int_{[0,T]}\psi_j(t)\left\langle e_k,\,\big(\eta(\tau,t)\big)[u]\right\rangle_H\mu(\mathrm{d}t)\right)^2\le\int_{[0,T]}\left\langle e_k,\,\big(\eta(\tau,t)\big)[u]\right\rangle_H^2\mu(\mathrm{d}t)\quad\text{and}$$

$$\begin{aligned}\mathbb{E}\int_0^T\sum_{u\in\mathcal{U}}\sum_{k=1}^{\infty}\int_{[0,T]}\left|\left\langle e_k,\,\big(\eta(\tau,t)\big)[u]\right\rangle_H^2\right|\mu(\mathrm{d}t)\,\mathrm{d}\tau &\le \mathbb{E}\int_0^T\sum_{u\in\mathcal{U}}\sum_{k=1}^{\infty}\int_{[0,T]}\left\langle e_k,\,\big(\eta(\tau,t)\big)[u]\right\rangle_H^2\mu(\mathrm{d}t)\,\mathrm{d}\tau\\ &= \mathbb{E}\int_0^T\sum_{u\in\mathcal{U}}\int_{[0,T]}\left\|\big(\eta(\tau,t)\big)[u]\right\|_H^2\mu(\mathrm{d}t)\,\mathrm{d}\tau\\ &<\infty.\end{aligned}$$

This proves (4.3.28) and also item (i). To verify item (ii), we first show that there exists a sequence $(n_l)_{l\in\mathbb{N}}\subset\mathbb{N}$ with $n_l\nearrow\infty$ such that $\mathbb{P}\otimes\mathrm{d}\tau$-a.e. it holds for μ-a.e. $t\in[0,T]$ and all $u\in\mathcal{U}$ that

$$\eta(\tau,t)[u]=\lim_{l\to\infty}\sum_{j=1}^{n_l}\sum_{k=1}^{n_l}\left\langle\phi_{jk},\,\big(\eta(\tau,\cdot)\big)[u]\right\rangle_{L^2(\mu;H)}\phi_{jk}(t). \tag{4.3.31}$$

To this end, note that it is enough to show that the sequence $\sum_{j,k=1}^{n}\left\langle\phi_{jk},\,\big(\eta(\tau,\cdot)\big)[\cdot]\right\rangle_{L^2(\mu;H)}\phi_{jk}$, $n\in\mathbb{N}$, is a Cauchy sequence in $L^2(\mathbb{P}\otimes\mathrm{d}\tau;\mathrm{HS}(U,L^2(\mu;H)))$, which follows from a similar reasoning as in the proof of Lemma 4.3.2(iii) and in (4.3.30). Moreover, note that (4.3.28) implies that there exists a subsequence $(n_{l'})_{l'\in\mathbb{N}}\subset(n_l)_{l\in\mathbb{N}}$ with $n_{l'}\nearrow\infty$ such that it holds $\mathbb{P}\otimes\mathrm{d}\tau$-a.e. that

$$D_\tau X=\lim_{l'\to\infty}\sum_{j=1}^{n_{l'}}\sum_{k=1}^{n_{l'}}\Big(D_\tau\left\langle\phi_{jk},X\right\rangle_{L^2(\mu;H)}\Big)\cdot\phi_{jk}\qquad\text{in }\mathrm{HS}\big(U,L^2(\mu;H)\big). \tag{4.3.32}$$

Now due to the continuity of the mapping $i\colon\mathrm{HS}(U,L^2(\mu;H))\to L^2(\mu;\mathrm{HS}(U,H))$ there exists a subsequence $(n_{l'_q})_{q\in\mathbb{N}}\subset(n_{l'})_{l'\in\mathbb{N}}$ with $n_{l'_q}\nearrow\infty$ such that it holds $\mathbb{P}\otimes\mathrm{d}\tau$-a.e. for μ-a.e. $t\in[0,T]$

and all $u \in \mathcal{U}$ that

$$\begin{aligned}\Big[i\big(D_\tau X\big)\Big](t)[u] &= \left[i\bigg(\lim_{l'\to\infty}\sum_{j=1}^{n_{l'}}\sum_{k=1}^{n_{l'}} D_\tau\Big(\big\langle\phi_{jk}, X\big\rangle_{L^2(\mu;H)}\Big)\,\phi_{jk}\bigg)\right](t)[u] \\ &= \left[\lim_{l'\to\infty} i\bigg(\sum_{j=1}^{n_{l'}}\sum_{k=1}^{n_{l'}} D_\tau\Big(\big\langle\phi_{jk}, X\big\rangle_{L^2(\mu;H)}\Big)\,\phi_{jk}\bigg)\right](t)[u] \\ &= \lim_{q\to\infty}\left[i\bigg(\sum_{j=1}^{n_{l'_q}}\sum_{k=1}^{n_{l'_q}} D_\tau\Big(\big\langle\phi_{jk}, X\big\rangle_{L^2(\mu;H)}\Big)\,\phi_{jk}\bigg)(t)[u]\right].\end{aligned} \tag{4.3.33}$$

Without loss of generality we can assume that all three sequences $(n_l)_{l\in\mathbb{N}}$, $(n'_l)_{l'\in\mathbb{N}}$, and $(n_{l'_q})_{q\in\mathbb{N}}$ are equal. The combination of (4.3.31), (4.3.33), and Lemma 4.3.3 implies that it holds $\mathbb{P}\otimes\mathrm{d}\tau$-a.e. for μ-a.e. $t \in [0,T]$ and all $u \in \mathcal{U}$ that

$$\begin{aligned}\Big[i\big(D_\tau X\big)\Big](t)[u] &= \lim_{l\to\infty}\left[i\bigg(\sum_{j=1}^{n_l}\sum_{k=1}^{n_l} D_\tau\Big(\big\langle\phi_{jk}, X\big\rangle_{L^2(\mu;H)}\Big)\,\phi_{jk}\bigg)(t)[u]\right] \\ &= \lim_{l\to\infty}\left[\bigg(\sum_{j=1}^{n_l}\sum_{k=1}^{n_l} D_\tau\Big(\big\langle\phi_{jk}, X\big\rangle_{L^2(\mu;H)}\Big)\,\phi_{jk}\bigg)[u](t)\right] \\ &= \lim_{l\to\infty}\left[\sum_{j=1}^{n_l}\sum_{k=1}^{n_l} D_\tau\Big(\big\langle\phi_{jk}, X\big\rangle_{L^2(\mu;H)}\Big)[u]\,\phi_{jk}(t)\right] \\ &= \lim_{l\to\infty}\left[\sum_{j=1}^{n_l}\sum_{k=1}^{n_l} \big\langle\phi_{jk}, \big(\eta(\tau,\cdot)\big)[u]\big\rangle_{L^2(\mu;H)}\,\phi_{jk}(t)\right] \\ &= \eta(\tau,t)[u].\end{aligned}$$

This proves item (ii). □

4.4 A weak convergence result

In this section we present, after a preparatory lemma, our main result Theorem 4.4.3. For $T \in (0,\infty)$ and a Hilbert space H, if μ is a finite Borel measure on $[0,T]$ and A is a linear operator on H, then the operator $e^{(\cdot)A}\colon H \to L^2(\mu;H)$ denotes the linear operator which maps any element $h \in H$ to the element of $L^2(\mu;H)$ given by the path $[0,T] \ni t \mapsto e^{tA}h \in H$. With that said, the integrals in the following lemma are $L^2(\mu;H)$-valued integrals.

Lemma 4.4.1. Let Assumption 4.1.1 be fulfilled, let $X : \Omega\times[0,T] \to H$ be the mild solution of (4.1.1) with continuous sample paths, and let μ be a finite Borel measure on $[0,T]$. Then the following equality holds $\mathbb{P}$-a.s. in $L^2(\mu;H)$:

$$X(\cdot) = e^{(\cdot)A}\xi + \int_0^T \mathbb{1}_{[\tau,T]}(\cdot)\,e^{(\cdot-\tau)A}F\big(X(\tau)\big)\,\mathrm{d}\tau + \int_0^T \mathbb{1}_{[\tau,T]}(\cdot)\,e^{(\cdot-\tau)A}B\big(X(\tau)\big)\,\mathrm{d}W(\tau). \tag{4.4.1}$$

Proof. Let $(\Psi(t))_{t\in[0,T]}$ be defined as the stochastic integral

$$\Psi(t) = \int_0^t e^{(t-\tau)A} B\big(X(\tau)\big)\,\mathrm{d}W(\tau) \quad \forall t \in [0,T] \ \mathbb{P}\text{-a.s.}$$

To prove equality (4.4.1), it is enough to show that it holds $\mathbb{P}$-a.s. in $L^2(\mu;H)$

$$\Psi(\cdot) = \int_0^T \mathbb{1}_{[\tau,T]}(\cdot)\, e^{(\cdot-\tau)A} B\big(X(\tau)\big)\,\mathrm{d}W(\tau). \tag{4.4.2}$$

The non-stochastic-integral part can be proven in a similar and easier way. Let $(\phi_k)_{k\in\mathbb{N}}$ be an orthonormal basis of $L^2(\mu;H)$. Observe that for all $k\in\mathbb{N}$ and all $t\in[0,T]$ it holds by [51, Lemma 2.4.1] that

$$\langle \Psi(t), \phi_k(t)\rangle_H = \int_0^T \Big\langle \mathbb{1}_{[\tau,T]}(t)\, e^{(t-\tau)A} B\big(X(\tau)\big)[\cdot],\, \phi_k(t)\Big\rangle_H \,\mathrm{d}W(\tau) \quad \mathbb{P}\text{-a.s.}$$

Furthermore, it holds by Assumption 4.1.1(i) and (4.1.3) that

$$\begin{aligned}
&\int_{[0,T]} \bigg(\mathbb{E}\int_0^T \Big\| \Big\langle \mathbb{1}_{[\tau,T]}(t)\, e^{(t-\tau)A} B\big(X(\tau)\big)[\cdot],\, \phi_k(t)\Big\rangle_H \Big\|^2_{\mathrm{HS}(U,\mathbb{R})} \mathrm{d}\tau\bigg)^{1/2} \mu(\mathrm{d}t) \\
&\quad\le \int_{[0,T]} \bigg(\mathbb{E}\int_0^T \Big\| \mathbb{1}_{[\tau,T]}(t)\, e^{(t-\tau)A} B\big(X(\tau)\big)\Big\|^2_{\mathrm{HS}(U,H)} \|\phi_k(t)\|_H^2 \mathrm{d}\tau\bigg)^{1/2} \mu(\mathrm{d}t) \\
&\quad\le L \int_{[0,T]} \bigg(\mathbb{E}\int_0^T \Big\| \mathbb{1}_{(\tau,T]}(t)\,(t-\tau)^{-\vartheta}\Big[1+\|X(\tau)\|_H\Big]^2 \mathrm{d}\tau\bigg)^{1/2} \cdot \big\|\phi_k(t)\big\|_H \,\mu(\mathrm{d}t) \\
&\quad\le L\cdot \mathbb{E}\Big[\sup_{t\in[0,T]} \Big(2+2\big\|X(t)\big\|_H^2\Big)\Big] \cdot \int_0^T (T-\tau)^{-\vartheta}\mathrm{d}\tau \cdot \int_{[0,T]} \big\|\phi_k(t)\big\|_H \,\mu(\mathrm{d}t) \\
&\quad\le L\cdot \mu([0,T])^{1/2}\cdot \big\|\phi_k\big\|_{L^2(\mu;H)} \cdot \mathbb{E}\Big[\sup_{t\in[0,T]} \Big(2+2\big\|X(t)\big\|_H^2\Big)\Big] \cdot \frac{T^{1-\vartheta}}{1-\vartheta} < \infty.
\end{aligned} \tag{4.4.3}$$

Therefore we can apply the stochastic Fubini theorem, Lemma 4.2.4, and obtain that

$$\begin{aligned}
\Psi(\cdot) &= \lim_{n\to\infty} \sum_{k=1}^n \langle \Psi,\, \phi_k\rangle_{L^2(\mu;H)}\ \phi_k \\
&= \lim_{n\to\infty} \sum_{k=1}^n \int_{[0,T]} \int_0^T \Big\langle \mathbb{1}_{[\tau,T]}(t)\, e^{(t-\tau)A} B\big(X(\tau)\big)[\cdot],\, \phi_k(t)\Big\rangle_H \,\mathrm{d}W(\tau)\,\mu(\mathrm{d}t)\,\phi_k \\
&= \lim_{n\to\infty} \sum_{k=1}^n \int_0^T \bigg(\int_{[0,T]} \Big\langle \mathbb{1}_{[\tau,T]}(t)\, e^{(t-\tau)A} B\big(X(\tau)\big)[\cdot],\, \phi_k(t)\Big\rangle_H \,\mu(\mathrm{d}t)\,\phi_k\bigg)\mathrm{d}W(\tau) \\
&= \lim_{n\to\infty} \int_0^T \sum_{k=1}^n \bigg(\int_{[0,T]} \Big\langle \mathbb{1}_{[\tau,T]}(t)\, e^{(t-\tau)A} B\big(X(\tau)\big)[\cdot],\, \phi_k(t)\Big\rangle_H \,\mu(\mathrm{d}t)\,\phi_k\bigg)\mathrm{d}W(\tau) \qquad \mathbb{P}\text{-a.s.}
\end{aligned} \tag{4.4.4}$$

To verify (4.4.2), it remains to show that the following holds in $L^2\big(\mathbb{P}\otimes \mathrm{d}\tau; \mathrm{HS}(U, L^2(\mu;H))\big)$

$$\lim_{n\to\infty}\sum_{k=1}^{n}\int_{[0,T]}\Big\langle \mathbb{1}_{[\tau,T]}(t)\, e^{(t-\tau)A}B\big(X(\tau)\big)[\cdot],\, \phi_k(t)\Big\rangle_H\, \mu(\mathrm{d}t)\,\phi_k = \mathbb{1}_{[\tau,T]}(\cdot)\, e^{(\cdot-\tau)A}B\big(X(\tau)\big). \tag{4.4.5}$$

To this end, note that the sequence $\big(\sum_{k=1}^{n}\int_{[0,T]}\langle \mathbb{1}_{[\tau,T]}(t)\, e^{(t-\tau)A}B(X(\tau))[\cdot],\, \phi_k(t)\rangle_H\, \mu(\mathrm{d}t)\,\phi_k\big)$, $n\in\mathbb{N}$, is a Cauchy sequence in $L^2(\mathbb{P}\otimes\mathrm{d}\tau; \mathrm{HS}(U, L^2(\mu;H)))$ and therefore converges in $L^2(\mathbb{P}\otimes \mathrm{d}\tau; \mathrm{HS}(U, L^2(\mu;H)))$. Indeed, for all $m,n\in\mathbb{N}$ with $m\le n$ it holds that

$$\begin{aligned}
&\mathbb{E}\int_0^T \Big\| \sum_{k=m}^{n}\int_{[0,T]}\Big\langle \mathbb{1}_{[\tau,T]}(t)\, e^{(t-\tau)A}B\big(X(\tau)\big)[\cdot],\, \phi_k(t)\Big\rangle_H\, \mu(\mathrm{d}t)\,\phi_k\Big\|^2_{\mathrm{HS}(U,L^2(\mu;H))}\mathrm{d}\tau\\
&= \mathbb{E}\int_0^T \sum_{u\in\mathcal{U}}\Big\| \sum_{k=m}^{n}\int_{[0,T]}\Big\langle \mathbb{1}_{[\tau,T]}(t)\, e^{(t-\tau)A}B\big(X(\tau)\big)[u],\, \phi_k(t)\Big\rangle_H\, \mu(\mathrm{d}t)\,\phi_k\Big\|^2_{L^2(\mu;H)}\mathrm{d}\tau \\
&= \mathbb{E}\int_0^T \sum_{u\in\mathcal{U}}\sum_{k=m}^{n}\Big(\int_{[0,T]}\Big\langle \mathbb{1}_{[\tau,T]}(t)\, e^{(t-\tau)A}B\big(X(\tau)\big)[u],\, \phi_k(t)\Big\rangle_H\, \mu(\mathrm{d}t)\Big)^2\mathrm{d}\tau.
\end{aligned} \tag{4.4.6}$$

Moreover, the following boundedness holds

$$\begin{aligned}
&\sum_{u\in\mathcal{U}}\sum_{k=m}^{n}\Big(\int_{[0,T]}\Big\langle \mathbb{1}_{[\tau,T]}(t)\, e^{(t-\tau)A}B\big(X(\tau)\big)[u],\, \phi_k(t)\Big\rangle_H\, \mu(\mathrm{d}t)\Big)^2\\
&\quad\le \sum_{u\in\mathcal{U}}\sum_{k=1}^{\infty}\Big(\int_{[0,T]}\Big\langle \mathbb{1}_{[\tau,T]}(t)\, e^{(t-\tau)A}B\big(X(\tau)\big)[u],\, \phi_k(t)\Big\rangle_H\, \mu(\mathrm{d}t)\Big)^2\\
&\quad= \Big\|\mathbb{1}_{[\tau,T]}(\cdot)\, e^{(\cdot-\tau)A}B\big(X(\tau)\big)\Big\|^2_{\mathrm{HS}(U,L^2(\mu;H))} \in L^1\big(\mathbb{P}\otimes\mathrm{d}\tau;\mathbb{R}\big),
\end{aligned}$$

and for $u\in\mathcal{U}$ we have that

$$\lim_{m,n\to\infty}\sum_{k=m}^{n}\Big(\int_{[0,T]}\Big\langle \mathbb{1}_{[\tau,T]}(t)\, e^{(t-\tau)A}B\big(X(\tau)\big)[u],\, \phi_k(t)\Big\rangle_H\, \mu(\mathrm{d}t)\Big)^2 = 0.$$

The dominated convergence theorem thus implies that

$$\lim_{m,n\to\infty}\mathbb{E}\int_0^T \Big\| \sum_{k=m}^{n}\int_{[0,T]}\Big\langle \mathbb{1}_{[\tau,T]}(t)\, e^{(t-\tau)A}B\big(X(\tau)\big)[\cdot],\, \phi_k(t)\Big\rangle_H\, \mu(\mathrm{d}t)\,\phi_k\Big\|^2_{\mathrm{HS}(U,L^2(\mu;H))}\mathrm{d}\tau = 0.$$

This proves (4.4.5) and consequently (4.4.4) can be rewritten as

$$\begin{aligned}
\Psi(\cdot) &= \lim_{n\to\infty}\int_0^T\sum_{k=1}^{n}\Big(\int_{[0,T]}\Big\langle \mathbb{1}_{[\tau,T]}(t)\, e^{(t-\tau)A}B\big(X(\tau)\big)[\cdot],\, \phi_k(t)\Big\rangle_H\, \mu(\mathrm{d}t)\,\phi_k\Big)\mathrm{d}W(\tau)\\
&= \int_0^T \lim_{n\to\infty}\sum_{k=1}^{n}\Big(\int_{[0,T]}\Big\langle \mathbb{1}_{[\tau,T]}(t)\, e^{(t-\tau)A}B\big(X(\tau)\big)[\cdot],\, \phi_k(t)\Big\rangle_H\, \mu(\mathrm{d}t)\,\phi_k\Big)\mathrm{d}W(\tau)\\
&= \int_0^T \mathbb{1}_{[\tau,T]}(\cdot)\, e^{(\cdot-\tau)A}B\big(X(\tau)\big)\,\mathrm{d}W(\tau) \quad \mathbb{P}\text{-a.s.}
\end{aligned}$$

This concludes the assertion of the Lemma. □

We are now ready to state and prove our main result. For this we need the following assumption:

Assumption 4.4.2. In addition to Assumption 4.1.1, let $A\colon D(A) \subset H \to H$ be a diagonal linear operator with eigenbasis $(e_n)_{n\in\mathbb{N}} \subset H$ and associated sequence of eigenvalues $(\lambda_n)_{n\in\mathbb{N}} \in \mathbb{R}$ such that $\sup_{n\in\mathbb{N}}(\lambda_n) < 0$, and let $\big(H_r, \langle\cdot,\cdot\rangle_{H_r}, \|\cdot\|_{H_r}\big)$, $r \in \mathbb{R}$ be a family of interpolation spaces associated to $-A$. Moreover, suppose that $B \in \operatorname{Lip}\big(H, \operatorname{HS}(U, H_{-\vartheta/2})\big)$. Let also $(P_N)_{N\in\mathbb{N}} \subset L(H)$ satisfy $P_N(v) = \sum_{n=1}^N \langle e_n, v\rangle_H \, e_n$ for all $v \in H$ and $N \in \mathbb{N}$. Note that for $x \in L^2(\mu; H)$, by $P_N x$ we mean the path defined by

$$P_N x(s) = P_N(x(s)) \qquad s \in [0,T].$$

Note that Assumption 4.4.2 and Lemma 2.2.14 imply that it holds for all $r \in [0, e)$ that

$$\sup_{t\in(0,\infty)} \big\|(-tA)^r e^{tA}\big\|_{L(H)} \le \sup_{x\in(0,\infty)} \Big[\frac{x^r}{e^x}\Big] \le \Big[\frac{r}{e}\Big]^r \le 1. \tag{4.4.7}$$

Theorem 4.4.3. Let Assumption 4.4.2 be fulfilled, let $X : \Omega \times [0,T] \to H$ be the mild solution of (4.1.1) with continuous sample paths, let $X(0) = \xi \in H_\rho$ with $\rho \in [0, 1-\vartheta)$, $\eta\colon \Omega \times [0,T] \times [0,T] \to H$ be a meausrable mapping which satisfies the properties formulated in item (iii) of Lemma 4.3.1. Moreover, let μ be a finite Borel measure on $[0,T]$, $N \in \mathbb{N}$, $q < \frac{1}{\rho+\vartheta/2}$, $p < \frac{2}{\vartheta}$ with $\frac{1}{p} + \frac{1}{q} = 1$, $q \le 2 \le p$, where we set $\frac{1}{\rho+\vartheta/2} = \infty$ if $\rho = \vartheta = 0$ and $\frac{2}{\vartheta} = \infty$ if $\vartheta = 0$. Furthermore, let $f \in C^2(L^q(\mu; H), \mathbb{R})$ be such that

$$M_f \colon = \max\Big\{ \sup_{x\in L^2(\mu;H)} \big\|f'(x)\big\|_{L(L^2(\mu;H),\mathbb{R})}, \sup_{x\in L^q(\mu;H)} \big\|f''(x)\big\|_{L^{(2)}(L^q(\mu;H),\mathbb{R})} \Big\} < \infty.$$

Then it holds that

$$\Big|\mathbb{E}\left[f(X) - f(P_N X)\right]\Big| \le M_{\xi,F,B,f,\mu,T,A} \cdot \|\operatorname{Id} - P_N\|_{L(H,H_{-\rho})}, \tag{4.4.8}$$

where

$$\begin{aligned} M_{\xi,F,B,f,\mu,T,A} = {} & M_f \sup_{t\in[0,T]} \big\|e^{tA}\big\|_{L(H)} \, \|\xi\|_{H_\rho} \, \mu([0,T])^{1/2} + M_f \, \|F\|_{\operatorname{Lip}(H,H_{-\vartheta})} \, \mu([0,T])^{1/2} \\ & \cdot \mathbb{E}\Big[\sup_{t\in[0,T]} \|X(t)\|_H\Big] \frac{T^{1-(\rho+\vartheta)}}{1-(\rho+\vartheta)} \\ & + \Big(\|B\|_{\operatorname{Lip}(H,\operatorname{HS}(U,H_{-\vartheta/2}))} \sup_{t\in[0,T]} \mathbb{E}\Big[\|X(t)\|_H^q\Big]^{1/q} \mu([0,T])^{1/q} \cdot \frac{1}{1-(\rho+\vartheta/2)q} T^{1-(\rho+\vartheta/2)q}\Big) \\ & \cdot M_f^2 \Big(\sup_{s\in[0,T)} \mathbb{E}\Big(\sup_{t\in(s,T]} (t-s)^{p\vartheta/2} \, \|\eta(s,t)\|_{\operatorname{HS}(U,H)}^p \Big)\Big)^{1/p} \mu([0,T])^{1/p} \, \frac{T^{1-p\vartheta/2}}{1-p\vartheta/2}. \end{aligned}$$

Proof. Due to Lemma 4.4.1 it holds $\mathbb{P}$-a.s. in $L^2(\mu;H)$ that

$$P_N X(\cdot) = e^{(\cdot)A} P_N \xi + \int_0^T \mathbb{1}_{[\tau,T]} e^{(\cdot-\tau)A} P_N F\big(X(\tau)\big)\,\mathrm{d}\tau + \int_0^T \mathbb{1}_{[\tau,T]} e^{(\cdot-\tau)A} P_N B\big(X(\tau)\big)\,\mathrm{d}W(\tau).$$

Note that the differentiability assumption on f ensures that the mapping $(0,1) \ni \theta \to f\big(P_N X + \theta(\mathrm{Id} - P_N)X\big) \in \mathbb{R}$ is differentiable and an application of the fundamental theorem of calculus and then Fubini's theorem yield that

$$\begin{aligned}
\mathbb{E}\big[f(X) - f(P_N X)\big] &= \mathbb{E}\Big[\int_0^1 \big\langle f'\big(\theta X + (1-\theta)P_N X\big),\, \big(\mathrm{Id} - P_N\big)X\big\rangle_{L^2(\mu;H)} \mathrm{d}\theta\Big] \\
&= \int_0^1 \mathbb{E}\Big[\big\langle f'\big(\theta X + (1-\theta)P_N X\big),\, \big(\mathrm{Id} - P_N\big)X\big\rangle_{L^2(\mu;H)}\Big] \mathrm{d}\theta.
\end{aligned}$$

Next consider that for $\theta \in [0,1]$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned}
&\big\langle f'\big(\theta X + (1-\theta)P_N X\big),\, \big(\mathrm{Id} - P_N\big)X\big\rangle_{L^2(\mu;H)} \\
&\quad = \big\langle f'\big(\theta X + (1-\theta)P_N X\big),\, e^{(\cdot)A}\big(\mathrm{Id} - P_N\big)\xi\big\rangle_{L^2(\mu;H)} \\
&\quad + \big\langle f'\big(\theta X + (1-\theta)P_N X\big),\, \int_0^T \mathbb{1}_{[\tau,T]} e^{(\cdot-\tau)A}\big(\mathrm{Id} - P_N\big)F\big(X(\tau)\big)\,\mathrm{d}\tau\big\rangle_{L^2(\mu;H)} \\
&\quad + \Big\langle f'\big(\theta X + (1-\theta)P_N X\big),\, \int_0^T \mathbb{1}_{[\tau,T]} e^{(\cdot-\tau)A}\big(\mathrm{Id} - P_N\big)B\big(X(\tau)\big)\,\mathrm{d}W(\tau)\Big\rangle_{L^2(\mu;H)} \qquad (4.4.9) \\
&\le \sup_{x \in L^2(\mu;H)} \big\|f'(x)\big\|_{L(L^2(\mu;H),\mathbb{R})} \Big(\big\|e^{(\cdot)A}\big(\mathrm{Id} - P_N\big)\xi\big\|_{L^2(\mu;H)} \\
&\qquad + \Big\| \int_0^T \mathbb{1}_{[\tau,T]} e^{(\cdot-\tau)A}\big(\mathrm{Id} - P_N\big)F\big(X(\tau)\big)\,\mathrm{d}\tau \Big\|_{L^2(\mu;H)} \Big) \\
&\quad + \Big\langle f'\big(\theta X + (1-\theta)P_N X\big),\, \int_0^T \mathbb{1}_{[\tau,T]} e^{(\cdot-\tau)A}\big(\mathrm{Id} - P_N\big)B\big(X(\tau)\big)\,\mathrm{d}W(\tau)\Big\rangle_{L^2(\mu;H)}.
\end{aligned}$$

For the sake of clarity, we divide the proof into several steps and analyze the three terms above separately.

Step I: For the first term above we have that

$$\begin{aligned}
\mathbb{E}\Big[\big\|e^{(\cdot)A}\big(\mathrm{Id} - P_N\big)\xi\big\|_{L^2(\mu;H)}\Big] &= \mathbb{E}\Big[\Big(\int_{[0,T]} \big\|e^{rA}\,(-A)^{-\rho}\big(\mathrm{Id} - P_N\big)\,(-A)^{\rho}\xi\big\|_H^2\; \mu(\mathrm{d}r)\Big)^{1/2}\Big] \\
&\le \|\mathrm{Id} - P_N\|_{L(H,H_{-\rho})} \sup_{t\in[0,T]} \big\|e^{tA}\big\|_{L(H)}\; \|\xi\|_{H_\rho}\; \mu([0,T])^{1/2}.
\end{aligned} \qquad (4.4.10)$$

Step II: For the second term in (4.4.9), we apply Jensen's inequality to obtain that

$$
\begin{aligned}
&\mathbb{E}\Big[\Big\| \int_0^T \mathbb{1}_{[\tau,T]}(\cdot) e^{(\cdot-\tau)A}\big(\mathrm{Id}-P_N\big)F\big(X(\tau)\big)\,\mathrm{d}\tau\Big\|_{L^2(\mu;H)}\Big] \\
&\quad = \mathbb{E}\Bigg[\Big(\int_{[0,T]}\Big\|\int_0^T \mathbb{1}_{[\tau,T]}(r)e^{(r-\tau)A}\big(\mathrm{Id}-P_N\big)F\big(X(\tau)\big)\,\mathrm{d}\tau\Big\|_H^2\,\mu(\mathrm{d}r)\Big)^{1/2}\Bigg] \\
&\quad \le \mathbb{E}\Bigg[\Big(\int_{[0,T]}\Big(\int_0^T \Big\|\mathbb{1}_{[\tau,T]}(r)\,(-A)^{\rho+\vartheta}e^{(r-\tau)A}\,(-A)^{-\rho}\big(\mathrm{Id}-P_N\big)\,(-A)^{-\vartheta}F\big(X(\tau)\big)\Big\|_H\,\mathrm{d}\tau\Big)^2\mu(\mathrm{d}r)\Big)^{1/2}\Bigg] \\
&\quad \le \mathbb{E}\Bigg[\Big(\int_{[0,T]}\Big(\int_0^T \mathbb{1}_{(\tau,T]}(r)\,(r-\tau)^{-(\rho+\vartheta)}\,\|\mathrm{Id}-P_N\|_{L(H,H_{-\rho})}\,\big\|F\big(X(\tau)\big)\big\|_{H_{-\vartheta}}\,\mathrm{d}\tau\Big)^2\mu(\mathrm{d}r)\Big)^{1/2}\Bigg] \\
&\quad \le \big\|\big(\mathrm{Id}-P_N\big)\big\|_{L(H,H_{-\rho})}\,\|F\|_{\mathrm{Lip}(H,H_{-\vartheta})}\,\mathbb{E}\Bigg[\Big(\int_{[0,T]}\Big(\int_0^T (T-\tau)^{-(\rho+\vartheta)}\,\|X(\tau)\|_H\,\mathrm{d}\tau\Big)^2\mu(\mathrm{d}r)\Big)^{1/2}\Bigg] \\
&\quad \le \|\mathrm{Id}-P_N\|_{L(H,H_{-\rho})}\,\|F\|_{\mathrm{Lip}(H,H_{-\vartheta})}\,\mathbb{E}\Big[\mu([0,T])^{1/2}\int_0^T (T-\tau)^{-(\rho+\vartheta)}\,\|X(\tau)\|_H\,\mathrm{d}\tau\Big] \\
&\quad \le \|\mathrm{Id}-P_N\|_{L(H,H_{-\rho})}\,\|F\|_{\mathrm{Lip}(H,H_{-\vartheta})}\,\mu([0,T])^{1/2}\,\mathbb{E}\Big[\sup_{t\in[0,T]}\|X(t)\|_H\Big]\int_0^T (T-\tau)^{-(\rho+\vartheta)}\mathrm{d}\tau \\
&\quad = \|\mathrm{Id}-P_N\|_{L(H,H_{-\rho})}\,\|F\|_{\mathrm{Lip}(H,H_{-\vartheta})}\,\mu([0,T])^{1/2}\,\mathbb{E}\Big[\sup_{t\in[0,T]}\|X(t)\|_H\Big]\,\frac{T^{1-(\rho+\vartheta)}}{1-(\rho+\vartheta)}.
\end{aligned}
\tag{4.4.11}
$$

Step III: For the last term in (4.4.9), note that $X \in \mathbb{D}^{1,2}(L^2(\mu;H))$ by Lemma 4.3.4(i). Therefore by the Malliavin chain rule, and Lemma 4.2.2, it holds that $f'\big(\theta X + (1-\theta)P_N X\big) \in \mathbb{D}^{1,2}(L^2(\mu;H))$, and that means we can apply the Malliavin integration by part formula, Lemma 4.2.3, and obtain that

$$
\begin{aligned}
&\mathbb{E}\Big[\Big\langle f'\big(\theta X + (1-\theta)P_N X\big), \int_0^T \mathbb{1}_{[\tau,T]}e^{(\cdot-\tau)A}\big(\mathrm{Id}-P_N\big)B\big(X(\tau)\big)\,\mathrm{d}W(\tau)\Big\rangle_{L^2(\mu;H)}\Big] \\
&\quad = \mathbb{E}\int_0^T \Big\langle D_\tau f'\big(\theta X + (1-\theta)P_N X\big), \mathbb{1}_{[\tau,T]}e^{(\cdot-\tau)A}\big(\mathrm{Id}-P_N\big)B\big(X(\tau)\big)\Big\rangle_{\mathrm{HS}(U,L^2(\mu;H))}\mathrm{d}\tau.
\end{aligned}
\tag{4.4.12}
$$

We use now the isomerty $i\colon \operatorname{HS}(U,L^2(\mu;H))\to L^2(\mu;\operatorname{HS}(U,H))$, introduced in Lemma 4.3.3, to continue our calculation and get that

$$\begin{aligned}
&\mathbb{E}\int_0^T \Big\langle D_\tau f'\big(\theta X+(1-\theta)P_NX\big), \mathbb{1}_{[\tau,T]}e^{(\cdot-\tau)A}\big(\operatorname{Id}-P_N\big)B\big(X(\tau)\big)\Big\rangle_{\operatorname{HS}(U,L^2(\mu;H))}\mathrm{d}\tau\\
&=\mathbb{E}\int_0^T \Big\langle i\Big(D_\tau f'\big(\theta X+(1-\theta)P_NX\big)\Big),\ i\Big(\mathbb{1}_{[\tau,T]}e^{(\cdot-\tau)A}\big(\operatorname{Id}-P_N\big)B\big(X(\tau)\big)\Big)\Big\rangle_{L^2(\mu;\operatorname{HS}(U,H))}\mathrm{d}\tau\\
&=\mathbb{E}\int_0^T\int_{[0,T]} \Big\langle i\Big(D_\tau f'\big(\theta X+(1-\theta)P_NX\big)\Big)(r),\ i\Big(\mathbb{1}_{[\tau,T]}e^{(\cdot-\tau)A}\big(\operatorname{Id}-P_N\big)B\big(X(\tau)\big)\Big)(r)\Big\rangle_{\operatorname{HS}(U,H)}\mu(\mathrm{d}r)\mathrm{d}\tau\\
&\le\bigg(\mathbb{E}\int_0^T\int_{[0,T]}\Big\|i\Big(D_\tau f'\big(\theta X+(1-\theta)P_NX\big)\Big)(r)\Big\|^p_{\operatorname{HS}(U,H)}\mu(\mathrm{d}r)\,\mathrm{d}\tau\bigg)^{1/p}\\
&\quad\cdot\bigg(\mathbb{E}\int_0^T\int_{[0,T]}\Big\|i\Big(\mathbb{1}_{[\tau,T]}e^{(\cdot-\tau)A}\big(\operatorname{Id}-P_N\big)B\big(X(\tau)\big)\Big)(r)\Big\|^q_{\operatorname{HS}(U,H)}\mu(\mathrm{d}r)\,\mathrm{d}\tau\bigg)^{1/q}\\
&=:C_1(\theta)\cdot C_2,
\end{aligned}\tag{4.4.13}$$

where the inequality above holds by Hölder with respect to $\mathbb{P}\otimes\mathrm{d}\tau\otimes\mu$. We analyze the term C_2 first. For this, observe that it holds $\mathbb{P}\otimes\mathrm{d}\tau$-a.e. for μ-a.e. $r\in[0,T]$ that

$$i\Big(\mathbb{1}_{[\tau,T]}e^{(\cdot-\tau)A}\big(\operatorname{Id}-P_N\big)B\big(X(\tau)\big)\Big)(r)=\mathbb{1}_{[\tau,T]}(r)e^{(r-\tau)A}\big(\operatorname{Id}-P_N\big)B\big(X(\tau)\big),$$

and therefore we obtain that

$$\begin{aligned}
C_2&=\bigg(\mathbb{E}\int_0^T\int_{[0,T]}\Big\|\mathbb{1}_{[\tau,T]}(r)e^{(r-\tau)A}\big(\operatorname{Id}-P_N\big)B\big(X(\tau)\big)\Big\|^q_{\operatorname{HS}(U,H)}\mu(\mathrm{d}r)\,\mathrm{d}\tau\bigg)^{1/q}\\
&=\bigg(\mathbb{E}\int_0^T\int_{[0,T]}\Big\|\mathbb{1}_{[\tau,T]}(r)\,(-A)^{\rho+\vartheta/2}\,e^{(r-\tau)A}\,(-A)^{-\rho}\big(\operatorname{Id}-P_N\big)\,(-A)^{-\vartheta/2}B\big(X(\tau)\big)\Big\|^q_{\operatorname{HS}(U,H)}\mu(\mathrm{d}r)\,\mathrm{d}\tau\bigg)^{1/q}\\
&\le\bigg(\int_0^T\int_{[0,T]}\mathbb{1}_{(\tau,T]}(r)\ (r-\tau)^{-(\rho+\vartheta/2)q}\ \mathbb{E}\Big(\big\|B\big(X(\tau)\big)\big\|^q_{\operatorname{HS}(U,H_{-\vartheta/2})}\Big)\,\mu(\mathrm{d}r)\,\mathrm{d}\tau\bigg)^{1/q}\ \|\operatorname{Id}-P_N\|_{L(H,H_{-\rho})}\\
&\le\|B\|_{\operatorname{Lip}(H,\operatorname{HS}(U,H_{-\vartheta/2}))}\sup_{t\in[0,T]}\mathbb{E}\Big[\,\|X(t)\|^q_H\Big]^{1/q}\ \|\operatorname{Id}-P_N\|_{L(H,H_{-\rho})}\\
&\quad\cdot\bigg(\int_{[0,T]}\int_0^T\tau^{-(\rho+\vartheta/2)q}\ \mathrm{d}\tau\,\mu(\mathrm{d}r)\bigg)^{1/q}\\
&\le\|B\|_{\operatorname{Lip}(H,\operatorname{HS}(U,H_{-\vartheta/2}))}\sup_{t\in[0,T]}\mathbb{E}\Big[\,\|X(t)\|^q_H\Big]^{1/q}\ \|\operatorname{Id}-P_N\|_{L(H,H_{-\rho})}\\
&\quad\cdot\mu([0,T])^{1/q}\frac{1}{1-(\rho+\vartheta/2)q}T^{1-(\rho+\vartheta/2)q}.
\end{aligned}\tag{4.4.14}$$

Next we try to bound the term $C_1(\theta)$ in (4.4.13) above. Using Minkowski's integral inequality [53, Theorem 13.14] we get that

$$\begin{aligned}
C_1(\theta) &= \left(\mathbb{E}\int_0^T \int_{[0,T]} \left\| i\Big(D_\tau f'\big(\theta X + (1-\theta)P_N X\big)\Big)(r)\right\|_{\mathrm{HS}(U,H)}^p \mu(\mathrm{d}r)\,\mathrm{d}\tau\right)^{1/p} \\
&= \left(\mathbb{E}\int_0^T \int_{[0,T]} \left[\sum_{u\in\mathcal{U}} \left\| i\Big(D_\tau f'\big(\theta X + (1-\theta)P_N X\big)\Big)(r)[u]\right\|_H^2\right]^{p/2} \mu(\mathrm{d}r)\,\mathrm{d}\tau\right)^{1/p} \\
&\le \left(\mathbb{E}\int_0^T \left(\sum_{u\in\mathcal{U}} \left[\int_{[0,T]} \left\| i\Big(D_\tau f'\big(\theta X + (1-\theta)P_N X\big)\Big)(r)[u]\right\|_H^p \mu(\mathrm{d}r)\right]^{2/p}\right)^{p/2} \mathrm{d}\tau\right)^{1/p} \\
&= \left(\mathbb{E}\int_0^T \left(\sum_{u\in\mathcal{U}} \left[\int_{[0,T]} \left\| \Big(D_\tau f'\big(\theta X + (1-\theta)P_N X\big)\Big)[u](r)\right\|_H^p \mu(\mathrm{d}r)\right]^{2/p}\right)^{p/2} \mathrm{d}\tau\right)^{1/p},
\end{aligned} \tag{4.4.15}$$

where the last equality holds due to Lemma 4.3.3. Now we can apply the Malliavin chain rule, Lemma 4.2.2, and obtain that

$$C_1(\theta) =$$

$$\left(\mathbb{E}\int_0^T \left(\sum_{u\in\mathcal{U}} \left[\int_{[0,T]} \left\| \Big((f''\big(\theta X + (1-\theta)P_N X\big)\big(\theta D_\tau X + (1-\theta)P_N D_\tau X\big)\Big)[u](r)\right\|_H^p \mu(\mathrm{d}r)\right]^{2/p}\right)^{p/2} \mathrm{d}\tau\right)^{1/p}.$$

Since $f \in C^2(L^q(\mu; H), \mathbb{R})$, it holds $\mathbb{P}\otimes \mathrm{d}\tau$-a.e. and for all $u \in \mathcal{U}$ that

$$f''\big(\theta X + (1-\theta)P_N X\big)\left(\cdot,\, \theta\,(D_\tau X)[u] + (1-\theta)P_N(D_\tau X)[u]\right) \in L\big(L^q(\mu; H), \mathbb{R}\big),$$

and by $f''\big(\theta X + (1-\theta)P_N X\big)\left(\theta\, D_\tau X + (1-\theta)P_N D_\tau X\right)[u]$ we mean the image of $f''\big(\theta X + (1-\theta)P_N X\big)\left(\cdot,\, \theta\,(D_\tau X)[u] + (1-\theta)P_N(D_\tau X)[u]\right)$ under the Riesz isomorphism. For all $\tau \in [0,T]$ and $u \in U$, we know that $\theta\,(D_\tau X)[u] + (1-\theta)P_N(D_\tau X)[u] \in L^2(\mu; H)$. However to continue with the calculation in (4.4.15) we consider $\theta\,(D_\tau X)[u] + (1-\theta)P_N(D_\tau X)[u]$ as an element in $L^q(\mu; H)$. Therefore it holds that

$$\begin{aligned}
C_1(\theta) &= \left(\mathbb{E}\int_0^T \left(\left\| f''\big(\theta X + (1-\theta)P_N X\big)\big(\theta D_\tau X + (1-\theta)P_N D_\tau X\big)\right\|_{\mathrm{HS}(U, L^p(\mu;H))}^p \mathrm{d}\tau\right)^{1/p} \\
&= \left(\mathbb{E}\int_0^T \left(\left\| f''\big(\theta X + (1-\theta)P_N X\big)\big(\cdot,\, \theta D_\tau X + (1-\theta)P_N D_\tau X\big)\right\|_{\mathrm{HS}(U, L(L^q(\mu;H);\mathbb{R}))}^p \mathrm{d}\tau\right)^{1/p} \\
&\le \sup_{x\in L^q(\mu;H)} \|f''(x)\|_{L^{(2)}(L^q(\mu;H),\mathbb{R})} \left(\mathbb{E}\int_0^T \left(\sum_{u\in\mathcal{U}} \left\|\theta\,(D_\tau X)[u] + (1-\theta)P_N(D_\tau X)[u]\right\|_{L^q(\mu;H)}^2\right)^{p/2} \mathrm{d}\tau\right)^{1/p} \\
&= \sup_{x\in L^q(\mu;H)} \|f''(x)\| \left(\mathbb{E}\int_0^T \left(\sum_{u\in\mathcal{U}} \left[\int_{[0,T]} \left\|\Big(\theta\,(D_\tau X)[u] + (1-\theta)P_N\,(D_\tau X)[u]\Big)(r)\right\|_H^q \mu(\mathrm{d}r)\right]^{2/q}\right)^{p/2} \mathrm{d}\tau\right)^{1/p}.
\end{aligned}$$

Now consider that there exists a measurable mapping $Q\colon \big(\mathcal{U}\times[0,T],\mathcal{P}(\mathcal{U})\otimes\mathcal{B}([0,T])\big)\to\big(H,\mathcal{B}(H)\big)$ such that

$$\forall u\in\mathcal{U}:\Big(Q(u,r)=\Big(\theta\,(D_\tau X)[u]+(1-\theta)P_N\,(D_\tau X)[u]\Big)(r),\quad \mu\text{-a.e. } r\in[0,T]\Big) \tag{4.4.16}$$

Therefore using the above equality and then Minkowski's integral inequality [53, Theorem 13.14] we get that

$$\begin{aligned}&\bigg(\sum_{u\in\mathcal{U}}\Big[\int_{[0,T]}\Big\|\Big(\theta\,(D_\tau X)[u]+(1-\theta)P_N\,(D_\tau X)[u]\Big)(r)\Big\|_H^q\mu(\mathrm{d}r)\Big]^{2/q}\bigg)^{p/2}\\&\quad=\bigg(\sum_{u\in\mathcal{U}}\Big[\int_{[0,T]}\Big\|Q(u,r)\Big\|_H^q\mu(\mathrm{d}r)\Big]^{2/q}\bigg)^{p/2}\\&\quad\le\bigg(\int_{[0,T]}\Big[\sum_{u\in\mathcal{U}}\Big\|Q(u,r)\Big\|_H^2\Big]^{q/2}\mu(\mathrm{d}r)\bigg)^{p/q}.\end{aligned}\tag{4.4.17}$$

Moreover, we know from Lemma 4.3.3 that

$$\forall u\in\mathcal{U}:\Big(i\Big(\theta\,(D_\tau X)+(1-\theta)P_N\,(D_\tau X)\Big)(r)[u]=\Big(\theta\,(D_\tau X)[u]+(1-\theta)P_N\,(D_\tau X)[u]\Big)(r),\quad \mu\text{-a.e. } r\in[0,T]\Big) \tag{4.4.18}$$

Putting (4.4.16), (4.4.17) and (4.4.18) together we get that

$$\begin{aligned}C_1(\theta)&=M_f\Bigg(\mathbb{E}\int_0^T\bigg(\int_{[0,T]}\Big[\sum_{u\in\mathcal{U}}\Big\|i\Big(\theta\,(D_\tau X)+(1-\theta)P_N\,(D_\tau X)\Big)(r)[u]\Big\|_H^2\Big]^{q/2}\mu(\mathrm{d}r)\bigg)^{p/q}\mathrm{d}\tau\Bigg)^{1/p}\\&=M_f\Bigg(\mathbb{E}\int_0^T\bigg(\int_{[0,T]}\Big\|i\Big(\theta\,(D_\tau X)+(1-\theta)P_N\,(D_\tau X)\Big)(r)\Big\|_{\mathrm{HS}(U,H)}^q\mu(\mathrm{d}r)\bigg)^{p/q}\mathrm{d}\tau\Bigg)^{1/p}\\&\le M_f\Bigg(\mathbb{E}\int_0^T\bigg(\int_{[0,T]}\Big\|i\Big(\theta\,(D_\tau X)+(1-\theta)P_N\,(D_\tau X)\Big)(r)\Big\|_{\mathrm{HS}(U,H)}^p\mu(\mathrm{d}r)\bigg)\mathrm{d}\tau\Bigg)^{1/p},\end{aligned}$$

where the last inequality holds due to Jensen's inequality and the fact that $p>q$. Now we apply Lemma 4.3.4(ii) and obtain that

$$\begin{aligned}C_1(\theta)&\le M_f\Bigg(\mathbb{E}\int_0^T\bigg(\int_{[0,T]}\Big\|i\Big(\theta\,(D_\tau X)+(1-\theta)P_N\,(D_\tau X)\Big)(r)\Big\|_{\mathrm{HS}(U,H)}^p\mu(\mathrm{d}r)\bigg)\mathrm{d}\tau\Bigg)^{1/p}\\&=M_f\Bigg(\mathbb{E}\int_0^T\bigg(\int_{[0,T]}\Big\|\theta\,\eta(\tau,r)+(1-\theta)P_N\,\eta(\tau,r)\Big\|_{\mathrm{HS}(U,H)}^p\mu(\mathrm{d}r)\bigg)\mathrm{d}\tau\Bigg)^{1/p}.\end{aligned}$$

Now considering the inequality above and using Lemma 4.3.1(iii) we have that

$$
\begin{aligned}
\int_0^1 C_1(\theta)\,\mathrm{d}\theta &= \int_0^1 \left(\mathbb{E}\int_0^T \int_{[0,T]} \left\| i\Big(D_\tau f'\big(\theta X + (1-\theta)P_N X\big)\Big)(r)\right\|^p_{\mathrm{HS}(U,H)} \mu(\mathrm{d}r)\,\mathrm{d}\tau\right)^{1/p} \mathrm{d}\theta \\
&\leq M_f \int_0^1 \Bigg[\theta \left(\mathbb{E}\int_0^T \left(\int_{[0,T]} \|\eta(\tau,r)\|^p_{\mathrm{HS}(U,H)} \mu(\mathrm{d}r)\right)\mathrm{d}\tau\right)^{1/p} \\
&\quad + (1-\theta)\left(\mathbb{E}\int_0^T \left(\int_{[0,T]} \left\|P_N\,\eta(\tau,r)\right\|^p_{\mathrm{HS}(U,H)} \mu(\mathrm{d}r)\right)\mathrm{d}\tau\right)^{1/p}\Bigg]\mathrm{d}\theta \\
&\leq M_f \left(\mathbb{E}\int_0^T \left(\int_{[0,T]} \|\eta(\tau,r)\|^p_{\mathrm{HS}(U,H)} \mu(\mathrm{d}r)\right)\mathrm{d}\tau\right)^{1/p} \\
&\leq M_f \Big(\sup_{s\in[0,T)} \mathbb{E}\big(\sup_{t\in(s,T]} (t-s)^{p\vartheta/2}\,\|\eta(s,t)\|^p_{\mathrm{HS}(U,H)}\big)\Big)^{1/p} \cdot \left(\int_{[0,T]}\int_0^T \mathbb{1}_{[0,r]}(\tau)(r-\tau)^{-p\vartheta/2}\mathrm{d}\tau\mu(\mathrm{d}r)\right)^{1/p} \\
&\leq M_f \Big(\sup_{s\in[0,T)} \mathbb{E}\big(\sup_{t\in(s,T]} (t-s)^{p\vartheta/2}\,\|\eta(s,t)\|^p_{\mathrm{HS}(U,H)}\big)\Big)^{1/p} \cdot \left(\int_{[0,T]}\int_0^T (T-\tau)^{-p\vartheta/2}\mathrm{d}\tau\mu(\mathrm{d}r)\right)^{1/p} \\
&\leq M_f \left(\sup_{s\in[0,T)} \mathbb{E}\big(\sup_{t\in(s,T]} (t-s)^{p\vartheta/2}\,\|\eta(s,t)\|^p_{\mathrm{HS}(U,H)}\big)\right)^{1/p} \mu([0,T])^{1/p}\,\frac{T^{1-p\vartheta/2}}{1-p\vartheta/2} < \infty.
\end{aligned}
\tag{4.4.19}
$$

Putting (4.4.10), (4.4.11), (4.4.13), (4.4.14) and (4.4.19) together proves (4.4.8). □

Remark 4.4.4. According to [31, Theorem 5.1], under the assumptions in Theorem 4.4.3 there exists a unique predictable stochastic process $X\colon [0,T]\times\Omega \to H_{\rho/2}$ satisfying (4.1.2) with $\sup_{t\in[0,T]}\mathbb{E}\left[\|X(t)\|^2_{H_{\rho/2}}\right] < \infty$. Hence, we have that

$$
\begin{aligned}
\mathbb{E}\left[\|X(\cdot) - P_N X(\cdot)\|_{L^2(\mu;H)}\right] &= \mathbb{E}\left[\left\|(-A)^{-\rho/2}\left(\mathrm{Id} - P_N\right)(-A)^{\rho/2}X(\cdot)\right\|_{L^2(\mu;H)}\right] \\
&\leq \|\mathrm{Id} - P_N\|_{L(H,H_{-\rho/2})}\,\mathbb{E}\left[\|X(\cdot)\|_{L^2(\mu;H_{\rho/2})}\right] \\
&\leq \|\mathrm{Id} - P_N\|_{L(H,H_{-\rho/2})}\left(\mathbb{E}\Big[\int_{[0,T]} \|X(t)\|^2_{H_{\rho/2}}\,\mu(\mathrm{d}t)\Big]\right)^{1/2} \\
&\leq \|\mathrm{Id} - P_N\|_{L(H,H_{-\rho/2})}\,\mu([0,T])^{1/2}\left(\sup_{t\in[0,T]}\mathbb{E}\left[\|X(t)\|^2_{H_{\rho/2}}\right]\right)^{1/2}.
\end{aligned}
\tag{4.4.20}
$$

This shows that the weak convergence rate in Theorem 4.4.3 is twice the strong convergence rate in (4.4.20) above.

Bibliography

[1] A. Andersson, M. Kovács, and S. Larsson, *Weak error analysis for semilinear stochastic Volterra equations with additive noise*, Journal of Mathematical Analysis and Applications, 437 (2016), pp. 1283–1304.

[2] A. Andersson, R. Kruse, and S. Larsson, *Duality in refined Sobolev–Malliavin spaces and weak approximation of SPDE*, Stochastics and Partial Differential Equations Analysis and Computations, 4 (2016), pp. 113–149.

[3] A. Andersson and S. Larsson, *Weak convergence for a spatial approximation of the nonlinear stochastic heat equation*, Mathematics of Computation, 85 (2016), pp. 1335–1358.

[4] A. Andersson and F. Lindner, *Malliavin regularity and weak approximation of semilinear SPDEs with Lévy noise*, Discrete and Continuous Dynamical Systems - Series B, 24 (2019), pp. 42–71.

[5] V. Bally, L. Caramellino, and R. Cont, *Functional Itô Calculus and Functional Kolmogorov Equations*, in Stochastic Integration by Parts and Functional Itô Calculus, Advanced Courses in Mathematics - CRM Barcelona, Birkhäuser Cham, 2016, pp. 123–216.

[6] V. Bally and D. Talay, *The law of the Euler scheme for stochastic differential equations. I. Convergence rate of the distribution function*, Probability Theory Related Fields, 104 (1996), pp. 43–60.

[7] ——, *The law of the Euler scheme for stochastic differential equations. II. Convergence rate of the density*, Monte Carlo Methods and Applications, 2 (1996), pp. 93–128.

[8] C.-E. Bréhier and A. Debussche, *Kolmogorov equations and weak order analysis for SPDES with nonlinear diffusion coefficient*, Journal de Mathématiques Pures et Appliquées (9), (2018), pp. 193–254.

[9] C.-E. Bréhier, M. Hairer, and A. M. Stuart, *Weak error estimates for trajectories of SPDEs under spectral Galerkin discretization*, Journal of Computational Mathematics, 36 (2018), pp. 159–182.

[10] E. Clément, A. Kohatsu-Higa, and D. Lamberton, *A duality approach for the weak approximation of stochastic differential equations*, Annals of Applied Probability, 16 (2006), pp. 1124–1154.

[11] R. Cont and D.-A. Fournié, *Change of variable formulas for non-anticipative functionals on path space*, Journal of Functional Analysis, 259 (2010), pp. 1043–1072.

[12] R. Cont and D.-A. Fournié, *Functional Itô calculus and stochastic integral representation of martingales*, Annals of Probability, 41 (2013), pp. 109–133.

[13] D. Conus, A. Jentzen, and R. Kurniawan, *Weak convergence rates of spectral Galerkin approximations for SPDEs with nonlinear diffusion coefficients*, Annals of Applied Probability, 29 (2019), pp. 653–716.

[14] A. Cosso and F. Russo, *A regularization approach to functional Itô calculus and strong-viscosity solutions to path-dependent PDEs*, arXiv e-print arXiv:1401.5034, (2014).

[15] G. Da Prato, A. Jentzen, and M. Röckner, *A mild Itô formula for SPDEs*, Transactions of the American Mathematical Society, (2019), pp. 3755–3807.

[16] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, Cambridge University Press, 2. ed., 2014.

[17] A. Debussche, *Weak approximation of stochastic partial differential equations: the nonlinear case*, Mathematics of Computation, 80 (2011), pp. 89–117.

[18] A. Debussche and E. Faou, *Weak backward error analysis for SDEs*, SIAM Journal on Numerical Analysis, 50 (2012), pp. 1735–1752.

[19] A. Debussche and J. Printems, *Weak order for the discretization of the stochastic heat equation*, Mathematics of Computation, 78 (2009), pp. 845–863.

[20] B. Dupire, *Functional Itô calculus*, Bloomberg Portfolio Research Paper No. 2009-04-FRONTIERS (2009). *Available at* http://dx.doi.org/10.2139/ssrn.1435551 (Accessed: 25 October 2023).

[21] K. Frieler and C. Knoche, *Solutions of stochastic differential equations in infinite dimensional Hilbert spaces and their dependence on initial data*, Diploma Thesis, Bielefeld University, BiBoS-Preprint E02-04-083, (2001).

[22] M. Fuhrman and G. Tessitore, *Nonlinear Kolmogorov equations in infinite dimensional spaces: the backward stochastic differential equations approach and applications to optimal control*, Annals of Probability, 30 (2002), pp. 1397–1465.

[23] L. Gawarecki and V. Mandrekar, *Stochastic Differential Equations in Infinite Dimensions*, Springer-Verlag, 2011.

[24] M. Geissert, M. Kovács, and S. Larsson, *Rate of weak convergence of the finite element method for the stochastic heat equation with additive noise*, BIT, 49 (2009), pp. 343–356.

[25] M. B. Giles, *Multilevel Monte Carlo methods*, Acta Numerica, 24 (2015), p. 259–328.

[26] I. Gyöngy and N. V. Krylov, *On stochastic equations with respect to semimartingales II. Itô formula in Banach spaces*, Stochastics, 6 (1982), pp. 153–173.

[27] I. Gyöngy and N. V. Krylov, *Itô formula in Banach spaces*, in Stochastic Differential Systems, M. Arató, D. Vermes, and A. V. Balakrishnan, eds., vol. 36 of Lecture Notes in Control and Information Sciences, Springer Berlin Heidelberg, 1981, pp. 69–73.

[28] K. Itô, *Stochastic integral*, Proceedings of the Imperial Academy, 20 (1944), pp. 519 – 524.

[29] K. Itô, *On a stochastic integral equation*, Proceedings of the Japan Academy, 22 (1946), pp. 32–35.

[30] A. Jentzen, *Stochastic partial differential equations: Analysis and numerical approximations*, ETH Zürich Lecture Notes (2016). *Available at* https://www2.math.ethz.ch/education/bachelor/lectures/fs2016/math/numsol/NASPDE_181.pdf *(Accessed: 25 October 2023).*

[31] A. Jentzen and P. E. Kloeden, *Taylor Approximations for Stochastic Partial Differential Equations*, Society for Industrial and Applied Mathematics (SIAM), 2011.

[32] A. Jentzen and R. Kurniawan, *Weak Convergence Rates for Euler-Type Approximations of Semilinear Stochastic Evolution Equations with Nonlinear Diffusion Coefficients*, Foundations of Computational Mathematics, 21 (2021), pp. 445–536.

[33] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, 1999.

[34] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, 1992.

[35] M. Kovács and F. Lindner, *Weak error analysis via functional Itô calculus*, arXiv e-print arXiv:1603.08756, (2016).

[36] M. Kovács, F. Lindner, and R. L. Schilling, *Weak Convergence of Finite Element Approximations of Linear Stochastic Evolution Equations with Additive Lévy Noise*, SIAM/ASA Journal on Uncertainty Quantification, 3 (2015), pp. 1159–1199.

[37] M. Kovács, S. Larsson, and F. Lindgren, *Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise.*, BIT, 52 (2012), pp. 85–108.

[38] ——, *Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise II. Fully discrete schemes*, BIT, 53 (2013), pp. 497–525.

[39] R. Kruse, *Strong and Weak Approximation of Semilinear Stochastic Evolution Equations*, Springer International Publishing, 2014.

[40] D. Leão, A. Ohashi, and A. B. Simas, *A weak version of path-dependent functional Itô calculus*, The Annals of Probability, 46 (2018), pp. 3399–3441.

[41] F. Lindner and R. L. Schilling, *Weak Order for the Discretization of the Stochastic Heat Equation Driven by Impulsive Noise*, Potential Analysis, 38 (2013), pp. 345–379.

[42] W. Lui and M. Röckner, *Stochastic Partial Differential Equations: An Introduction*, Springer-Verlag, 2015.

[43] P. Malliavin, *Stochastic Analysis*, Springer-Verlag, 1997.

[44] G. N. Milstein, *Weak Approximation of Solutions of Systems of Stochastic Differential Equations*, Theory of Probability & Its Applications, 30 (1986), pp. 750–766.

[45] G. N. Milstein and M. V. Tretyakov, *Stochastic Numerics for Mathematical Physics*, Springer-Verlag, 2004.

[46] T. Müller-Gronbach, E. Novak, and K. Ritter, *Monte Carlo-Algorithmen*, Springer Berlin, Heidelberg, 2012.

[47] D. Nualart, *The Malliavin Calculus and Related Topics*, Springer-Verlag, 2006.

[48] H. Oberhauser, *The functional Itô formula under the family of continuous semimartingale measures*, Stochastics and Dynamics, 16 (2016).

[49] S. Peszat and J. Zabczyk, *Stochastic Partial Differential Equations with Lévy Noise*, Cambridge University Press, 2007.

[50] G. Pisier, *Martingales in Banach Spaces*, Cambridge University Press, 2016.

[51] C. Prévôt and M. Röckner, *A Concise Course on Stochastic Partial Differential Equations*, Springer-Verlag, 2007.

[52] M. Rosestolato, *Functional Itô calculus in Hilbert spaces and application to path-dependent Kolmogorov equations*, arXiv e-print arXiv:1606.06326, (2018).

[53] R. L. Schilling, *Measures, Integrals and Martingales*, Cambridge University Press, 2005.

[54] R. L. Schilling and L. Partzsch, *Brownian Motion: An Introduction to Stochastic Processes*, De Gruyter, 2012.

[55] D. Talay, *Probabilistic numerical methods for partial differential equations: Elements of analysis*, in Probabilistic Models for Nonlinear Partial Differential Equations, D. Talay and L. Tubaro, eds., vol. 1627 of Lecture Notes in Mathematics, Springer Berlin Heidelberg, 1996, pp. 148–196.

[56] D. Talay and L. Tubaro, *Expansion of the global error for numerical schemes solving stochastic differential equations*, Stochastic Analysis and Applications, 8 (1990), pp. 483–509.

[57] D. Werner, *Funktionalanalysis*, Springer-Verlag, 8. ed., 2018.

[58] E. Zeidler, *Applied Functional Analysis: Main Principles and Their Applications*, vol. 109 of Applied Mathematical Sciences, New York, 1995.