J. Evol. Equ. 20 (2020), 1093–1130

1424-3199/20/031093-38, *published online* October 30, 2019
https://doi.org/10.1007/s00028-019-00546-0

Journal of Evolution Equations

Fréchet differentiability of mild solutions to SPDEs with respect to the initial datum

Carlo Marinelli and Luca Scarpa

Abstract. We establish n-th-order Fréchet differentiability with respect to the initial datum of mild solutions to a class of jump diffusions in Hilbert spaces. In particular, the coefficients are Lipschitz-continuous, but their derivatives of order higher than one can grow polynomially, and the (multiplicative) noise sources are a cylindrical Wiener process and a quasi-left-continuous integer-valued random measure. As preliminary steps, we prove well-posedness in the mild sense for this class of equations, as well as first-order Gâteaux differentiability of their solutions with respect to the initial datum, extending previous results by Marinelli, Prévôt, and Röckner in several ways. The differentiability results obtained here are a fundamental step to construct classical solutions to non-local Kolmogorov equations with sufficiently regular coefficients by probabilistic means.

1. Introduction

Our goal is to obtain existence and uniqueness of mild solutions, and, especially, their differentiability with respect to the initial datum, to a class of stochastic evolution equations on Hilbert spaces of the form

$$\begin{cases} \mathrm{d}u(t) + Au(t)\,\mathrm{d}t = f(t,u(t))\,\mathrm{d}t + B(t,u(t))\,\mathrm{d}W(t) \\ \quad + \displaystyle\int_Z G(t,z,u(t-))\,\bar{\mu}(\mathrm{d}t,\mathrm{d}z), \\ u(0) = u_0. \end{cases} \tag{1.1}$$

Here, A is a linear m-accretive operator, W is a cylindrical Wiener process, $\bar{\mu}$ is a compensated integer-valued quasi-left-continuous random measure, and the coefficients f, B, G satisfy suitable measurability and Lipschitz continuity conditions. Precise assumptions on the data of the problem are stated in Sects. 2.1 and 3 below.

The results extend (and partially supersede) those obtained in [15] in several ways: (a) well-posedness is established here in much greater generality, in particular allowing $\bar{\mu}$ to be a quite general random measure, rather than just a compensated Poisson measure as in [15]. Moreover, using a more precise maximal estimate for stochastic convolutions, solutions are no longer needed to be sought in spaces of processes with finite second moment (yet more general well-posedness results are going to appear in [14]); (b) the sufficient conditions on the coefficients of (1.1) for the differentiability of

its solution with respect to the initial datum are the natural ones. For instance, roughly speaking, Fréchet differentiability of f, B, and G imply Fréchet differentiability of the solution map $u_0 \mapsto u$, while in [15] a C^1 condition on f, B, and G was needed. In fact, the proof in [15] was based on an implicit function theorem with parameters, for which the C^1 assumption seems indispensable, while here we use a direct approach based on the definition of derivative; (c) we study the n-th-order differentiability of the solution map for arbitrary natural n, instead of considering only first and second-order differentiability as in [15]. In this regard it is worth mentioning that we just assume that the derivatives of f, B, and G of order higher than one satisfy a polynomial growth condition. While this assumption causes non-trivial technical difficulties, it is more natural than much more restrictive boundedness conditions that are often found in the literature: a possible example of coefficients with non-bounded higher derivatives is given in Example 6.1 below.

There are several reasons to study the differentiability of solutions to stochastic equations in infinite dimensions with respect to the initial datum (or, more generally, with respect to parameters), among which the probabilistic construction of solutions to Kolmogorov equations is our main motivation. This vast and mature field of investigation is still very active, especially regarding stochastic equations with additive Wiener noise: see, e.g., [12] for classical results in the finite-dimensional case, [9] for basic results in the Hilbertian setting, and [4,6,7,20] for accounts of more recent developments. On the other hand, the case of equations with discontinuous noise, for which the associated Kolmogorov equations are of non-local type, is much less investigated, especially in the infinite-dimensional setting (see [15] for simple results and [19] for a special case). As an application of the above-mentioned differentiability results, we shall construct, in a forthcoming work, classical solutions to non-local Kolmogorov equations with sufficiently regular coefficients. As is well known, such results are essential to consider Kolmogorov equations motivated by applications that usually have less regular coefficients. In fact, a typical approach is, roughly speaking, to regularize the coefficients of the equation, thus obtaining a family of approximating Kolmogorov equations that are sufficiently simple to have classical solutions, and to obtain a solution to the original problem passing to the limit, in an appropriate sense, with respect to the regularization parameter. In this spirit, our ultimate goal is the extension of the results in [18] to non-local Kolmogorov equations associated with stochastic evolution equations with jumps in a generalized variational setting as considered in [17].

Since the literature on the problem at hand is very large, it is not easy to provide an accurate comparison of our results with existing ones, apart of the remarks already made. We should nonetheless mention the recent work [2], which considers a problem analogous to ours, but without discontinuous noise term and with coefficients with bounded derivatives of all orders. Here, the authors exploit the smoothing property of an analytic semigroup and study differentiability in negative order spaces.

The remaining text is organized as follows: in Sect. 2, after fixing some notation, we recall a characterization of Gâteaux and Fréchet differentiability, as well as some

maximal estimates for deterministic and stochastic convolutions, all of which are essential tools. Well-posedness of (1.1), i.e., existence and uniqueness of a mild solution and its continuous dependence on the initial datum, is proved in Sect. 3. The remaining sections are devoted to differentiability properties of the mild solution to (1.1) with respect to the initial datum: first-order Gâteaux and Fréchet differentiability are treated in Sect. 4 and Sect. 5, respectively, and n-th-order Fréchet differentiability is considered in Sect. 6.

2. Preliminaries

2.1. Notation

The spaces of linear bounded operators from a Banach space E to a further Banach space F will be denoted by $\mathscr{L}(E, F)$, and $\mathscr{L}^2(E, F)$ stands for the space of Hilbert–Schmidt operators from E to F if E and F are Hilbert spaces. The closed ball of radius $r > 0$ in E will be denoted by $B_r(E)$.

All stochastic elements will be defined on a fixed filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, with the filtration $\mathbb{F} := (\mathcal{F}_t)_{t\in[0,T]}$ complete and right-continuous, and $T > 0$ a fixed final time. Moreover, H will always denote a fixed real separable Hilbert space with norm $\|\cdot\|$. For any $p > 0$ and $[t_0, t_1] \subseteq [0, T]$, we shall use the notation $\mathbb{S}^p(t_0, t_1)$ for the space of adapted càdlàg H-valued processes Y such that

$$\|Y\|_{\mathbb{S}^p(t_0,t_1)} := \Big(\mathbb{E}\sup_{t\in[t_0,t_1]}\|Y(t)\|^p\Big)^{1/p} < +\infty,$$

and we set $\mathbb{S}^p := \mathbb{S}^p(0, T)$. We recall that these are Banach spaces if $p \geq 1$, and quasi-Banach spaces if $p \in]0, 1[$. In the latter case the triangle inequality is reversed, but one has

$$\|Y_1 + Y_1\|_{\mathbb{S}^p(t_0,t_1)} \leq 2^{1/p}\big(\|Y_1\|_{\mathbb{S}^p(t_0,t_1)} + \|Y_2\|_{\mathbb{S}^p(t_0,t_1)}\big),$$

to which we shall also refer, with a harmless abuse of terminology, as the triangle inequality. Moreover, $\mathbb{S}^p(t_0, t_1)$ is a complete metric space for every $p > 0$ when endowed with the distance

$$d_{p,t_0,t_1}(Y_1, Y_2) := \|Y_1 - Y_2\|^{1\wedge p}_{\mathbb{S}^p(t_0,t_1)},$$

as it follows from the inequality $|x + y|^p \leq |x|^p + |y|^p$, which holds true for every x, $y \in \mathbb{R}$ and $p \in]0, 1[$. For brevity we shall write $d_p := d_{p,0,T}$. Entirely analogously, $L^p(\Omega; H)$ endowed with the distance

$$(Y_1, Y_2) \mapsto \|Y_1 - Y_2\|^{1\wedge p}_{L^p(\Omega;H)}$$

is a complete metric space for every $p > 0$.

Let K be a real separable Hilbert space and W a cylindrical Wiener process on K. Let $(Z, \mathscr{Z})$ be a Blackwell measurable space and μ an integer-valued quasi-left-continuous random measure on $Z \times [0,T]$, independent of W, with dual predictable projection (compensator) ν, and $\bar{\mu} := \mu - \nu$. We recall that the assumption on $(Z, \mathscr{Z})$ as a Blackwell space is usually required in the literature on random measures (see [11, §1a]), and it ensures, for example, that $\mathscr{Z}$ is separable and generated by a countable algebra. We also recall that the quasi-left-continuity of μ implies that the random measure ν is non-atomic (see, e.g., [11, Corollary 1.19, p. 70]). A map $g : \Omega \times [0,T] \times Z \to H$ will be called predictable if it is $\mathscr{P} \otimes \mathscr{Z}$-measurable, where $\mathscr{P}$ stands for the predictable σ-algebra of $\mathbb{F}$ (the target space H is always assumed to be endowed with the Borel σ-algebra). Moreover, for any such predictable map g, we set, for any $p, q \in]0, \infty[$,

$$\|g\|_{L^q(\nu;H)} := \left(\int_{]0,T]\times Z} \|g\|^q \, \mathrm{d}\nu \right)^{1/q},$$

$$\|g\|_{L^p(\Omega;L^q(\nu;H))} := \left(\mathbb{E}\left(\int_{]0,T]\times Z} \|g\|^q \, \mathrm{d}\nu \right)^{p/q} \right)^{1/p}$$

and

$$\|g\|_{\mathbb{G}^p} := \begin{cases} \|g\|_{L^p(\Omega;L^2(\nu;H))} & \text{if } p \in]0,1], \\ \inf\limits_{g_1+g_2=g} \left(\|g_1\|_{L^p(\Omega;L^2(\nu;H))} + \|g_2\|_{L^p(\Omega;L^p(\nu;H))} \right) & \text{if } p \in]1,2[, \\ \|g\|_{L^p(\Omega;L^2(\nu;H))} + \|g\|_{L^p(\Omega;L^p(\nu;H))} & \text{if } p \in [2,\infty[, \end{cases}$$

where the infima are taken with respect to $\mathscr{P} \otimes \mathscr{Z}$-measurable maps g_1, g_2 only. One may actually show that $L^p(\Omega; L^q(\nu; H))$ as well as $\mathbb{G}^p$ are (quasi-)Banach space and that

$$\mathbb{G}^p = \begin{cases} L^p(\Omega; L^2(\nu; H)), & p \in]0,1], \\ L^p(\Omega; L^2(\nu; H)) + L^p(\Omega; L^p(\nu; H)), & p \in [1,2], \\ L^p(\Omega; L^2(\nu; H)) \cap L^p(\Omega; L^p(\nu; H)), & p \in [2,\infty[. \end{cases}$$

For a proof of this statement, as well as of other properties of such mixed-norm L^p spaces involving random measures (even in a more general setting), we refer to [10]. For us, however, it is enough to know that they are quasi-normed spaces, and the "norms" just introduced on spaces where the underlying measure is random is only a convenient notation. We shall also need to consider spaces where $]0,T] \times Z$ is replaced by $]t_0, t_1] \times Z$, with $0 \le t_0 \le t_1 \le T$, and the corresponding notation will be self-explanatory.

We shall use standard notation of stochastic calculus: we write, for instance, f^* and f_- to denote the maximal function and the left-limit function of a càdlàg function f, respectively. Further notation related to deterministic and stochastic convolutions, as well as to different notions of derivative for maps between infinite-dimensional spaces,

will be introduced where they first appear. For any $a, b > 0$, we use the notation $a \lesssim b$ to indicate that there exists a constant $c > 0$ such that $a \leq cb$. If c depends on some further quantities that we need to keep track of, we shall indicate them in a subscript. We use the classical notation $\wedge$ and $\vee$ for min and max, respectively.

2.2. Notions of derivative

Let E, F be Banach spaces, and G be a subspace of E. A function $\phi : E \to F$ is Gâteaux differentiable at $x_0 \in E$ along G if there exists a continuous linear map $L \in \mathscr{L}(G, F)$ such that

$$\lim_{\varepsilon \to 0} \frac{\phi(x_0 + \varepsilon h) - \phi(x_0)}{\varepsilon} = Lh \qquad \forall h \in G.$$

The linear map L, which is necessarily unique, will be denoted by $D_{\mathcal{G}}\phi(x_0)$ and is called the Gâteaux derivative of ϕ at x_0 (along the subspace G, if $G \neq E$). If $G = E$ and ϕ is also Lipschitz-continuous with Lipschitz constant L_ϕ, it easily follows from the definition that $\left\| D_{\mathcal{G}}\phi(x_0) \right\|_{\mathscr{L}(E,F)} \leq L_\phi$: indeed, for all $h \in E$ we have

$$\left\| D_{\mathcal{G}}\phi(x_0)h \right\|_F = \lim_{\varepsilon \to 0} \frac{\left\| \phi(x_0 + \varepsilon h) - \phi(x_0) \right\|_F}{\varepsilon} \leq L_\phi \frac{\left\| x_0 + \varepsilon h - x_0 \right\|_E}{\varepsilon} = L_\phi \left\| h \right\|_E.$$

The map ϕ is Fréchet-differentiable at $x_0 \in E$ along the subspace G if there exists a continuous linear map $L \in \mathscr{L}(G, F)$ such that

$$\lim_{h \to 0} \frac{\phi(x_0 + h) - \phi(x_0) - Lh}{\left\| h \right\|_G} = 0.$$

The (unique) map L will be denoted by $D\phi(x_0)$ and is called the Fréchet derivative of ϕ at x_0 (along the subspace G, in case $G \neq E$). It is well known that Fréchet differentiability implies Gâteaux differentiability, while the converse is not true. We shall often use the following characterization of Fréchet differentiability, of which we include a proof for the convenience of the reader.

Lemma 2.1. *A map $\phi : E \to F$ is Fréchet-differentiable at $x_0 \in E$ with $D\phi(x_0) = L$ if and only if for each bounded set $B \subset E$ one has*

$$\lim_{\varepsilon \to 0} \frac{\phi(x_0 + \varepsilon h) - \phi(x_0) - \varepsilon Lh}{\varepsilon} = 0 \tag{2.1}$$

uniformly with respect to $h \in B$.

Proof. Let ϕ be Fréchet-differentiable at x_0 with $D\phi(x_0) = L$, and set $R(h) := \phi(x_0 + h) - \phi(x_0) - Lh$. Then, $R(h)/\left\| h \right\| \to 0$ as $h \to 0$. Let B be a bounded set and M a real number such that B is included in the ball of E of radius M centered at zero. For any $\eta > 0$ there exists $\delta > 0$ such that $\left\| R(h) \right\| / \left\| h \right\| \leq \eta / M$ for every h with $\left\| h \right\| \leq \delta$. Therefore, for any ε such that $|\varepsilon| \leq \delta / M$, one has $\left\| \varepsilon h \right\| \leq \delta$ and

$$\left\| R(\varepsilon h) \right\| \leq \eta \frac{\left\| \varepsilon h \right\|}{M} = \eta |\varepsilon| \frac{\left\| h \right\|}{M} \leq \eta |\varepsilon|,$$

i.e., $\|R(\varepsilon h)\|/|\varepsilon| \to 0$ as $\varepsilon \to 0$ uniformly with respect to $h \in B$.

Let us now prove the converse implication: assume that (2.1) holds for every B, uniformly with respect to $h \in B$, and that, by contradiction, ϕ is not Fréchet differentiable at x_0, i.e., that $R(h)/\|h\|$ does not converge to zero as $h \to 0$. In particular, there exists a sequence $(k_n) \subset E\backslash\{0\}$ converging to zero such that $R(k_n)/\|k_n\|$ does not converge to zero. We claim that it cannot happen that

$$\sup_{h\in B} \frac{\phi(x_0+\varepsilon h)-\phi(x_0)-\varepsilon Lh}{\varepsilon} \longrightarrow 0$$

as $\varepsilon \to 0$. In fact, setting $\varepsilon_n := \|k_n\|$, $h_n = k_n/\|k_n\|$, and $B := (h_n)$, this would imply that $\varepsilon_n^{-1}\big(\phi(x_0+\varepsilon_n h_n)-\varphi(x_0)-\varepsilon_n L h_n\big)$ converges to zero as $n \to \infty$, which is equivalent to $R(k_n)/\|k_n\| \to 0$. □

By a simple scaling argument it is evident that it is sufficient to consider as bounded subset B the unit ball in E. One can thus say that $\phi : E \to F$ is Fréchet-differentiable at $x_0 \in E$ along a subspace $G \subseteq E$ if there exists a continuous linear map $L : G \to F$ such that

$$\lim_{\varepsilon\to 0} \frac{\phi(x_0+\varepsilon h)-\phi(x_0)-\varepsilon Lh}{\varepsilon} = 0 \qquad \text{uniformly on } \{h \in G : \|h\|_G \le 1.\}$$

For a comprehensive treatment of differential calculus for functions between topological vector spaces we refer to [1] for basic results in the case of Banach spaces and to [3,5] for the general case.

2.3. Estimates for deterministic and stochastic convolutions

Throughout this section S stands for a strongly continuous linear semigroup of contractions on H and $-A$ for its generator. Clearly, A is necessarily a linear maximal monotone operator.

Here and in the following we shall use $S * g$ to denote convolution of S and an H-valued measurable function g on $\mathbb{R}_+$, defined as

$$S * g : \mathbb{R}_+ \ni t \longmapsto \int_0^t S(t-s)g(s)\,\mathrm{d}s,$$

under the minimal assumption that $S(t-\cdot)g \in L^1(0,t;H)$ for all t in a set of interest, usually a bounded interval of $\mathbb{R}_+$.

The following estimate for convolutions is trivial, but sufficient for our purposes.

Lemma 2.2. *For every $p > 0$ and for every measurable adapted process $\phi : \Omega \times [0,T] \to H$ such that $\phi \in L^p(\Omega; L^1(0,T;H))$, it holds that $S * \phi \in \mathbb{S}^p(0,T)$ and*

$$\|S * \phi\|_{\mathbb{S}^p(0,T)} \le \|\phi\|_{L^p(\Omega;L^1(0,T;H))}.$$

Proof. Minkowski's inequality and contractivity of S immediately yield

$$\mathbb{E}\sup_{t\le T}\left\|\int_0^t S(t-s)\phi(s)\,\mathrm{d}s\right\|^p \le \mathbb{E}\left(\sup_{t\le T}\int_0^t \left\|S(t-s)\phi(s)\right\|\mathrm{d}s\right)^p$$
$$\le \mathbb{E}\left(\int_0^T \left\|\phi(s)\right\|\mathrm{d}s\right)^p. \qquad \square$$

We shall also need estimates for stochastic convolutions with respect to the cylindrical Wiener process W, for which we shall always use the following notation: for any $\mathscr{L}^2(K,H)$-valued process G, the stochastic convolution $S\diamond G$ is the process defined as

$$S\diamond G(t) := \int_0^t S(t-s)G(s)\,\mathrm{d}W(s), \qquad t\ge 0,$$

under a stochastic integrability assumption on $S(t-\cdot)G$. There is an extensive literature on maximal estimates for stochastic convolutions, mostly obtained through the so-called factorization method by Da Prato et al. [8], which requires $-A$ to generate a holomorphic semigroup. The following estimate instead requires A to be maximal monotone and can be proved by relatively elementary techniques of stochastic calculus (see, e.g., [13] for a proof in a more general context).

Proposition 2.3. *For every $p>0$ and for every $G\in L^p(\Omega;L^2(0,T;\mathscr{L}^2(K,H)))$ progressively measurable, the stochastic convolution $S\diamond G$ admits a modification in $\mathbb{S}^p(0,T)$ and*

$$\left\|S\diamond G\right\|_{\mathbb{S}^p(0,T)} \lesssim_p \left\|G\right\|_{L^p(\Omega;L^2(0,T;\mathscr{L}^2(K,H)))}.$$

Finally, a key role is played by the following maximal estimate for stochastic convolutions with respect to the compensated random measure $\bar\mu$. For a predictable H-valued process g, the stochastic convolution of g with respect to $\bar\mu$ will be denote by $S\diamond_\mu g$ and defined as

$$S\diamond_\mu g(t) := \int_{]0,t]}\int_Z S(t-s)g(s,z)\,\bar\mu(\mathrm{d}s,\mathrm{d}z), \qquad t\ge 0,$$

under a stochastic integrability assumption on $S(t-\cdot)g$ with respect to $\bar\mu$.

Lemma 2.4. *For every $p>0$ and for every $g\in\mathbb{G}^p$, the stochastic convolution $S\diamond_\mu g$ admits a càdlàg modification and*

$$\left\|S\diamond_\mu g\right\|_{\mathbb{S}^p} \lesssim \left\|g\right\|_{\mathbb{G}^p}.$$

A proof can be found in [16]. A generalization of this inequality to L^q-valued processes will appear in [14].

3. Well-posedness

This section is devoted to the proof of well-posedness of Eq. (1.1). We show existence and uniqueness of a mild solution, as well as its continuous dependence on the initial datum, in spaces of processes with finite moments of order $p \in]0, +\infty[$. Although only the case $p \geq 1$ is needed in the following sections on differentiability of the solution with respect to the initial datum, the general case $p > 0$ is necessary to deal with initial data or driving random measures admitting finite moments of order strictly less than one. An example is given by α-stable random measures with $\alpha < 1$.

The following assumptions (A0)–(A4) on the coefficients and the initial datum of (1.1) are in force throughout the paper.

(A0) The initial datum u_0 is an $\mathscr{F}_0$-measurable random variable with values in H;

(A1) A is a linear maximal monotone operator on H, and S is the strongly continuous semigroup of contractions generated by $-A$ on H;

(A2) The function $f : \Omega \times [0, T] \times H \to H$ is such that $f(\cdot, \cdot, x)$ is measurable and adapted for every $x \in H$, and there exists a constant $C_f > 0$ such that

$$\begin{aligned} \big\| f(\omega, t, x) \big\| &\leq C_f \big(1 + \|x\|\big), \\ \big\| f(\omega, t, x) - f(\omega, t, y) \big\| &\leq C_f \|x - y\| \end{aligned}$$

for all $\omega \in \Omega$, $t \in [0, T]$, and $x,\ y \in H$;

(A3) The function $B : \Omega \times [0, T] \times H \to \mathscr{L}^2(K, H)$ is such that $B(\cdot, \cdot, x)$ is progressively measurable for all $x \in H$, and there exists a constant $C_B > 0$ such that

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

for all $\omega \in \Omega$, $t \in [0, T]$, and $x,\ y \in H$;

(A4) The function $G : \Omega \times [0, T] \times Z \times H \to H$ is such that $G(\cdot, \cdot, \cdot, x)$ is $\mathscr{P} \otimes \mathscr{Z}$-measurable for all $x \in H$. Moreover,

(i) if $p \leq 1$ or $p \geq 2$, then there exists a $\mathscr{P} \otimes \mathscr{Z}$-measurable function $g : \Omega \times [0, T] \times Z \to \mathbb{R}$ such that

$$\begin{aligned} \big\| G(\omega, t, z, x) - G(\omega, t, z, y) \big\| &\leq g(\omega, t, z) \|x - y\|, \\ \big\| G(\omega, t, z, x) \big\| &\leq g(\omega, t, z) \big(1 + \|x\|\big) \end{aligned}$$

for all $\omega \in \Omega$, $t \in [0, T]$, $z \in Z$ and $x,\ y \in H$;

(ii) if $1 < p < 2$, then there exist functions $G_1, G_2 : \Omega \times [0, T] \times Z \times H \to H$, satisfying the same measurability properties of G, with $G = G_1 + G_2$, and $\mathscr{P} \otimes \mathscr{Z}$-measurable functions $g_1, g_2 : \Omega \times [0, T] \times Z \to \mathbb{R}$ such that, for $j \in \{1, 2\}$,

$$\begin{aligned} \big\| G_j(\omega, t, z, x) - G_j(\omega, t, z, y) \big\| &\leq g_j(\omega, t, z) \|x - y\|, \\ \big\| G_j(\omega, t, z, x) \big\| &\leq g_j(\omega, t, z) \big(1 + \|x\|\big) \end{aligned}$$

for all $\omega \in \Omega$, $t \in [0,T]$, $z \in Z$ and $x,\ y \in H$.

Further assumptions will be made when needed.

The concept of solution to (1.1) we shall work with is the following.

Definition 3.1. An H-valued adapted càdlàg process u is a mild solution to (1.1) if

(i) $S(t-\cdot)f(u) \in L^1(0,t;H)$ for all $t \in [0,T]$ $\mathbb{P}$-a.s.;
(ii) $S(t-\cdot)B(u) \in L^2(0,t;\mathscr{L}^2(K,H))$ for all $t \in [0,T]$ $\mathbb{P}$-a.s.;
(iii) there exists $p > 0$ such that $S(t-\cdot)G(u_-) \in \mathbb{G}_p(0,t)$ for all $t \in [0,T]$;
(iv) one has

$$u = S(\cdot)u_0 + S * f(u) + S \diamond B(u) + S \diamond_\mu G(u_-)$$

as an identity in the sense of modifications.

In order to formulate the well-posedness result in the mild sense for (1.1), it is convenient to introduce an assumption depending on a parameter $p \in]0,\infty[$:

(A5$_p$) Setting $g_1 := g_2 := g/2$ if $p \notin]1,2[$, there exists a continuous increasing function $\kappa : \mathbb{R}_+ \to \mathbb{R}_+$, with $\kappa(0) = 0$, such that

$$1_{\{p>1\}}\left(\int_{Z\times[t_0,t_1]} g_1^p(\omega,s,z)\,\mathrm{d}\nu\right)^{1/p} + \left(\int_{Z\times[t_0,t_1]} g_2^2(\omega,s,z)\,\mathrm{d}\nu\right)^{1/2} \leq \kappa(t_1-t_0) \qquad \forall \omega \in \Omega.$$

Theorem 3.2. *Let $p > 0$ and (A5$_p$) be satisfied. For any $u_0 \in L^p(\Omega;H)$, Eq. (1.1) admits a unique mild solution $u \in \mathbb{S}^p$ such that $\|u\|_{\mathbb{S}^p} \lesssim 1 + \|u_0\|_{L^p(\Omega;H)}$, with implicit constant independent of u_0. Moreover, the solution map $u_0 \mapsto u$ is Lipschitz-continuous from $L^p(\Omega;H)$ to $\mathbb{S}^p$.*

Proof. We are going to use a fixed-point argument in the metric space $(\mathbb{S}^p(0,T_0), d_{p,0,T_0})$, with T_0 sufficiently small. By a classical patching argument, this will imply existence and uniqueness of a solution in $\mathbb{S}^p(0,T)$. Let Γ be the map formally defined on $L^p(\Omega;H) \times \mathbb{S}^p$ as

$$\Gamma : (u_0,u) \longmapsto S(\cdot)u_0 + S * f(u) + S \diamond B(u) + S \diamond_\mu G(u_-).$$

Let us show that Γ is in fact well defined on $L^p(\Omega;H) \times \mathbb{S}^p$ and that its image is contained in $\mathbb{S}^p$: one has

$$\|\Gamma(u_0,u)\|_{\mathbb{S}^p} \lesssim \|S(\cdot)u_0\|_{\mathbb{S}^p} + \|S * f(u)\|_{\mathbb{S}^p} + \|S \diamond B(u)\|_{\mathbb{S}^p} + \|S \diamond_\mu G(u_-)\|_{\mathbb{S}^p}, \tag{3.1}$$

where $\|S(\cdot)u_0\|_{\mathbb{S}^p} \leq \|u_0\|_{L^p(\Omega;H)}$ by contractivity of the semigroup S; the elementary Lemma 2.2 and linear growth of f imply

$$\begin{aligned}\|S * f(u)\|_{\mathbb{S}^p} &\leq \|f(u)\|_{L^p(\Omega;L^1(0,T;H))} \leq C_f\big(1 + \|u\|_{L^p(\Omega;L^1(0,T;H))}\big) \\ &\lesssim_p TC_f\big(1 + \|u\|_{\mathbb{S}^p}\big);\end{aligned} \tag{3.2}$$

similarly, Proposition 2.3 yields

$$\begin{aligned}\big\|S \diamond B(u)\big\|_{\mathbb{S}^p} &\lesssim \big\|B(u)\big\|_{L^p(\Omega;L^2(0,T;\mathscr{L}^2(K,H)))} \le C_B\big(1+\big\|u\big\|_{L^p(\Omega;L^2(0,T;H))}\big)\\ &\lesssim_p T^{1/2}C_B\big(1+\big\|u\big\|_{\mathbb{S}^p}\big);\end{aligned} \tag{3.3}$$

finally, it follows by Proposition 2.4 that $\big\|S \diamond_\mu G(u_-)\big\|_{\mathbb{S}^p}^p \lesssim \big\|G(u_-)\big\|_{\mathbb{G}^p}^p$, where if $p \in]0,1] \cup [2,\infty[$,

$$\begin{aligned}\big\|G(u_-)\big\|_{\mathbb{G}^p}^p &= 1_{\{p>1\}}\,\mathbb{E}\int \big\|G(u_-)\big\|^p\,\mathrm{d}\nu + \mathbb{E}\bigg(\int \big\|G(u_-)\big\|^2\,\mathrm{d}\nu\bigg)^{p/2}\\ &\le 1_{\{p>1\}}\,\mathbb{E}\int g^p(1+\big\|u_-\big\|^p)\,\mathrm{d}\nu + \mathbb{E}\bigg(\int g^2(1+\big\|u_-\big\|^2)\,\mathrm{d}\nu\bigg)^{p/2}\\ &\lesssim \kappa^p(T)(1+\big\|u\big\|_{\mathbb{S}^p}^p),\end{aligned} \tag{3.4}$$

and, similarly, if $p \in]1,2[$,

$$\begin{aligned}\big\|G(u_-)\big\|_{\mathbb{G}^p}^p &= \inf_{\tilde g_1+\tilde g_2=G(u_-)}\Big(\big\|\tilde g_2\big\|_{L^p(\Omega;L^p(\nu;H))}^p + \big\|\tilde g_1\big\|_{L^p(\Omega;L^2(\nu;H))}^p\Big)\\ &\le \mathbb{E}\int \big\|G_1(u_-)\big\|^p\,\mathrm{d}\nu + \mathbb{E}\bigg(\int \big\|G_2(u_-)\big\|^2\,\mathrm{d}\nu\bigg)^{p/2}\\ &\le \mathbb{E}\int g_1^p(1+\big\|u_-\big\|^p)\,\mathrm{d}\nu + \mathbb{E}\bigg(\int g_2^2(1+\big\|u_-\big\|^2)\,\mathrm{d}\nu\bigg)^{p/2}\\ &\lesssim \kappa^p(T)(1+\big\|u\big\|_{\mathbb{S}^p}^p).\end{aligned} \tag{3.5}$$

Analogous arguments show that that $\Gamma(u_0,\cdot)$ is a contraction of $\mathbb{S}^p(0,T_0)$, with T_0 to be chosen later. In fact, one has, with a slightly simplified notation,

$$\begin{aligned}\big\|\Gamma u - \Gamma v\big\|_{\mathbb{S}^p}^{1\wedge p} &\le \big\|S * (f(u)-f(v))\big\|_{\mathbb{S}^p}^{1\wedge p} + \big\|S \diamond (B(u)-B(v))\big\|_{\mathbb{S}^p}^{1\wedge p}\\ &\quad + \big\|S \diamond_\mu (G(u_-)-G(v_-))\big\|_{\mathbb{S}^p}^{1\wedge p}\\ &=: A_1 + A_2 + A_3.\end{aligned}$$

Let us estimate the three terms separately. The Lipschitz continuity of f, B, and G yields

$$\begin{aligned}\big\|S * (f(u)-f(v))\big\|_{\mathbb{S}^p(0,T_0)} &\le \big\|f(u)-f(v)\big\|_{L^p(\Omega;L^1(0,T_0;H))}\\ &\le T_0 C_f\big\|u-v\big\|_{\mathbb{S}^p(0,T_0)},\\ \big\|S \diamond (B(u)-B(v))\big\|_{\mathbb{S}^p(0,T_0)} &\lesssim \big\|B(u)-B(v)\big\|_{L^p(\Omega;L^2(0,T_0;\mathscr{L}^2(K,H)))}\\ &\le T_0^{1/2} C_B\big\|u-v\big\|_{\mathbb{S}^p(0,T_0)},\\ \big\|S \diamond_\mu (G(u_-)-G(v_-))\big\|_{\mathbb{S}^p(0,T_0)} &\lesssim \big\|G(u_-)-G(v_-)\big\|_{\mathbb{G}^p(0,T_0)}\\ &\lesssim \kappa(T_0)\big\|u-v\big\|_{\mathbb{S}^p(0,T_0)},\end{aligned}$$

so that

$$A_1 = \big\| S * (f(u) - f(v)) \big\|^{1\wedge p}_{\mathbb{S}^p(0,T_0)} \lesssim (T_0 C_f)^{1\wedge p} \big\| u - v \big\|^{1\wedge p}_{\mathbb{S}^p(0,T_0)},$$
$$A_2 = \big\| S \diamond (B(u) - B(v)) \big\|^{1\wedge p}_{\mathbb{S}^p(0,T_0)} \lesssim \big(T_0^{1/2} C_B\big)^{1\wedge p} \big\| u - v \big\|^{1\wedge p}_{\mathbb{S}^p(0,T_0)},$$
$$A_3 = \big\| S \diamond (G(u_-) - G(v_-)) \big\|^{1\wedge p}_{\mathbb{S}^p(0,T_0)} \lesssim \kappa(T_0)^{1\wedge p} \big\| u - v \big\|^{1\wedge p}_{\mathbb{S}^p(0,T_0)}.$$

Since κ is continuous with $\kappa(0) = 0$, it follows that there exists $T_0 > 0$ and a constant $\eta \in]0, 1[$, which depends on T_0, such that

$$d_{p,0,T_0}(\Gamma u, \Gamma v) \leq \eta d_{p,0,T_0}(u, v),$$

hence, by the Banach–Caccioppoli contraction principle, for any $u_0 \in L^p(\Omega; H)$ there exists a fixed point u of the contraction $\Gamma(u_0, \cdot)$, which is thus the unique solution in $\mathbb{S}^p(0, T_0)$ to (1.1). Choosing T_0 such that $T = nT_0$, with $n \in \mathbb{N}$, and repeating the same argument on each interval $[kT_0, (k+1)T_0]$, with $k \in \{1, \dots, n-1\}$, a unique solution to (1.1) can be constructed on the whole interval $[0, T]$. Furthermore, for any $u_0 \in L^p(\Omega; H)$, by (3.1)–(3.5), the unique solution $u = \Gamma(u_0, u) \in \mathbb{S}^p(0, T)$ satisfies

$$\big\| u \big\|^p_{\mathbb{S}^p(0,T)} \lesssim 1 + \big\| u_0 \big\|^p_{L^p(\Omega;H)} + (T + T^{1/2} + \kappa^p(T)) \big\| u \big\|^p_{\mathbb{S}^p(0,T)},$$

where the implicit constant is independent of T. Hence, there is $T_0 \in (0, T)$ small enough such that

$$\big\| u \big\|^p_{\mathbb{S}^p(0,T_0)} \lesssim 1 + \big\| u_0 \big\|^p_{L^p(\Omega;H)}.$$

Performing now a patching argument as above on $[0, T_0], \dots, [(n-1)T_0, T]$ yields the desired estimate

$$\big\| u \big\|^p_{\mathbb{S}^p(0,T)} \lesssim 1 + \big\| u_0 \big\|^p_{L^p(\Omega;H)}.$$

The argument to show the Lipschitz continuity of $u_0 \mapsto u$ is similar: let $u_{01}, u_{02} \in L^p(\Omega; H)$, and $u_1, u_2 \in \mathbb{S}^p(0, T)$ be the unique solutions to (1.1) with initial datum u_{01} and u_{02}, respectively. Using a patching argument as above, it suffices to show that $u_0 \mapsto u$ is Lipschitz-continuous on $[0, T_0]$. To this purpose, one has

$$\begin{aligned} d_{p,0,T_0}(u_1, u_2) &= d_{p,0,T_0}\big(\Gamma(u_{01}, u_1), \Gamma(u_{02}, u_2)\big) \\ &\leq d_{p,0,T_0}\big(\Gamma(u_{01}, u_1), \Gamma(u_{02}, u_1)\big) + d_{p,0,T_0}\big(\Gamma(u_{02}, u_1), \Gamma(u_{02}, u_2)\big) \\ &\leq \big\| u_{01} - u_{02} \big\|^{1\wedge p}_{L^p(\Omega;H)} + \eta d_{p,0,T_0}(u_1, u_2), \end{aligned}$$

where $\eta < 1$ is a positive constant (that depends on T_0). Rearranging terms and performing a patching argument as above immediately yields the Lipschitz continuity of $u_0 \mapsto u$. □

Remark 3.3. It immediately follows from the Lipschitz continuity of the solution map that one also has, in the same notation used above,

$$\big\| u_1 - u_2 \big\|_{\mathbb{S}^p} \lesssim \big\| u_{01} - u_{02} \big\|_{L^p(\Omega;H)},$$

with implicit constant depending on T and p.

4. Gâteaux differentiability of the solution map

In the previous section we have shown that the solution map $u_0 \mapsto u$ is Lipschitz-continuous from $L^p(\Omega; H)$ to $\mathbb{S}^p$. We are now going to show that Gâteaux differentiability of the coefficients of (1.1) implies Gâteaux differentiability of the solution map. For some applications (e.g., to study Kolmogorov equations associated with stochastic PDEs) it is sufficient to consider non-random initial data and to consider first-order derivatives as linear maps from H to $\mathbb{S}^p$, i.e., roughly speaking, to consider only non-random directions of differentiability. However, the more general case of random initial data and random directions of differentiability considered here as well as in the next sections is conceptually not more difficult and, apart of being interesting in its own right because treated at the natural level of generality, it is necessary to study, for instance, higher-order stability issues of stochastic models with respect to perturbations of the initial datum.

We shall make the following additional assumption, which is assumed to hold throughout this section.

(G1) The maps $f(\omega, t, \cdot)$ and $B(\omega, t, \cdot)$ are Gâteaux differentiable for all $(\omega, t) \in \Omega \times [0, T]$, and the maps

$$G(\omega, t, z, \cdot), \quad G_1(\omega, t, z, \cdot), \quad G_2(\omega, t, z, \cdot)$$

are Gâteaux differentiable for all $(\omega, t, z) \in \Omega \times [0, T] \times Z$.

The Gâteaux derivatives of f, B and G (in their H-valued argument) are denoted by

$$\begin{aligned} D_{\mathcal{G}} f &: \Omega \times [0, T] \times H \longrightarrow \mathscr{L}(H, H), \\ D_{\mathcal{G}} B &: \Omega \times [0, T] \times H \longrightarrow \mathscr{L}(H, \mathscr{L}^2(K, H)), \\ D_{\mathcal{G}} G &: \Omega \times [0, T] \times Z \times H \longrightarrow \mathscr{L}(H, H). \end{aligned}$$

Recalling that f and B are Lipschitz-continuous in their H-valued argument, uniformly over $\Omega \times [0, T]$, we infer that

$$\begin{aligned} \left\| D_{\mathcal{G}} f(\omega, t, x_0) \right\|_{\mathscr{L}(H,H)} &\le C_f, \\ \left\| D_{\mathcal{G}} B(\omega, t, x_0) \right\|_{\mathscr{L}(H, \mathscr{L}^2(K,H))} &\le C_B \end{aligned}$$

for all $\omega \in \Omega, t \in [0, T]$, and $x_0 \in H$. Similarly, the Lipschitz continuity of G implies, if $p \notin]1, 2[$, that

$$\left\| D_{\mathcal{G}} G(\omega, t, z, x_0) \right\|_{\mathscr{L}(H,H)} \le g(\omega, t, z),$$

and, if $p \in]1, 2[$, that

$$\begin{aligned} \left\| D_{\mathcal{G}} G(\omega, t, z, x_0) \right\|_{\mathscr{L}(H,H)} &\le \left\| D_{\mathcal{G}} G_1(\omega, t, z, x_0) \right\|_{\mathscr{L}(H,H)} \\ &\quad + \left\| D_{\mathcal{G}} G_2(\omega, t, z, x_0) \right\|_{\mathscr{L}(H,H)} \\ &\le g_1(\omega, t, z) + g_2(\omega, t, z), \end{aligned}$$

for all $\omega \in \Omega$, $t \in [0, T]$, $z \in Z$, and $x_0 \in H$.

We begin with two general results that will be extensively used in the sequel. The first lemma is an immediate corollary of the well-posedness results.

Lemma 4.1. *Under the assumptions of Theorem 3.2, let $u \in \mathbb{S}^p$ be the unique mild solution to (1.1) with initial condition $u_0 \in L^p(\Omega; H)$. For any $h \in L^p(\Omega; H)$, the linear stochastic evolution equation*

$$dy + Ay\,dt = D_{\mathcal{G}} f(u)y\,dt + D_{\mathcal{G}} B(u)y\,dW + \int_Z D_{\mathcal{G}} G(u_-)y_-\,d\bar{\mu}, \qquad y(0) = h, \tag{4.1}$$

admits a unique mild solution $y \in \mathbb{S}^p$ that depends continuously on the initial datum h.

Proof. The linear maps $D_{\mathcal{G}} f(u)$ and $D_{\mathcal{G}} B(u)$ are bounded, uniformly over $\Omega \times [0, T]$, hence, a fortiori, Lipschitz-continuous. Analogously, the linear map $D_{\mathcal{G}} G(u_-)$ has norm (and, a fortiori, Lipschitz constant) bounded by $g_1 + g_2$ (with $g_1 := g_2 := g/2$ if $p \notin]1, 2[$) on $\Omega \times [0, T] \times Z$. Theorem 3.2 thus implies that, for any $h \in L^p(\Omega; H)$, (4.1) admits a unique mild solution $y \in \mathbb{S}^p$, which depends continuously on h. □

Note that since the equation for y is linear, it is immediate that the map $h \mapsto y$ is linear and continuous from $L^p(\Omega; H)$ to $\mathbb{S}^p$.

The next lemma will play a crucial role both in the proof of the Gâteaux differentiability of the solution map in this section and in the proof of its Fréchet differentiability in the next section, taking into account Lemma 2.1.

Lemma 4.2. *Under the assumptions of Theorem 3.2, let $h \in L^p(\Omega; H)$ and u, $u_\varepsilon \in \mathbb{S}^p$ the unique mild solutions to (1.1) with initial conditions u_0 and $u_0 + \varepsilon h$, respectively. Moreover, let $y \in \mathbb{S}^p$ be the unique mild solution to (4.1) with initial condition h. One has*

$$\begin{aligned}
&\left\|\varepsilon^{-1}(u_\varepsilon - u - \varepsilon y)\right\|_{\mathbb{S}^p} \\
&\quad \lesssim_p \left\|\varepsilon^{-1}\big(f(u + \varepsilon y) - f(u) - \varepsilon D_{\mathcal{G}} f(u)y\big)\right\|_{L^p(\Omega; L^1(0,T;H))} \\
&\qquad + \left\|\varepsilon^{-1}\big(B(u + \varepsilon y) - B(u) - \varepsilon D_{\mathcal{G}} B(u)y\big)\right\|_{L^p(\Omega; L^2(0,T;\mathscr{L}^2(K,H)))} \\
&\qquad + \left\|\varepsilon^{-1}\big(G(u_- + \varepsilon y_-) - G(u_-) - \varepsilon D_{\mathcal{G}} G(u_-)y_-\big)\right\|_{\mathbb{G}^p}
\end{aligned}$$

Proof. Let $[t_0, t_1] \subset [0, T]$, and consider the evolution equation

$$\mathrm{d}v + Av\,\mathrm{d}t = f(v)\,\mathrm{d}t + B(v)\,\mathrm{d}W + G(v_-)\,\mathrm{d}\bar{\mu}, \qquad v(t_0) = u(t_0).$$

One easily sees that it admits a unique mild solution v, which coincides with the restriction of u to $[t_0, t_1]$. In particular, for any $t \geq t_0$,

$$\begin{aligned}
u(t) = S(t - t_0)u(t_0) &+ \int_{t_0}^t S(t - s) f(u(s))\,\mathrm{d}s + \int_{t_0}^t S(t - s)B(u(s))\,\mathrm{d}W(s) \\
&+ \int_{t_0}^t \int_Z S(t - s)G(z, u(s-))\,\bar{\mu}(\mathrm{d}z, \mathrm{d}s).
\end{aligned} \tag{4.2}$$

A completely analogous flow property holds for u_ε and y. Then, one has, by the triangle inequality,

$$\begin{aligned}
&\big\|\varepsilon^{-1}(u_\varepsilon - u - \varepsilon y)\big\|_{\mathbb{S}^p(t_0,t_1)}\\
&\lesssim_p \big\|\varepsilon^{-1}(u_\varepsilon(t_0) - u(t_0) - \varepsilon y(t_0))\big\|_{L^p(\Omega;H)}\\
&\quad + \big\|S * \varepsilon^{-1}\big(f(u_\varepsilon) - f(u) - \varepsilon D_{\mathcal{G}} f(u)y\big)\big\|_{\mathbb{S}^p(t_0,t_1)}\\
&\quad + \big\|S \diamond \varepsilon^{-1}\big(B(u_\varepsilon) - B(u) - \varepsilon D_{\mathcal{G}} B(u)y\big)\big\|_{\mathbb{S}^p(t_0,t_1)}\\
&\quad + \big\|S \diamond_\mu \varepsilon^{-1}\big(G(u_{\varepsilon-}) - G(u_-) - \varepsilon D_{\mathcal{G}} G(u_-)y_-\big)\big\|_{\mathbb{S}^p(t_0,t_1)}\\
&=: I_0 + I_1 + I_2 + I_3,
\end{aligned}$$

where, by abuse of notation, the (deterministic and stochastic) convolutions are defined on $[t_0, t_1]$, in accordance with (4.2), and $u_{\varepsilon-} := (u_\varepsilon)_-$. We are going to estimate I_1, I_2 and I_3 separately. To simplify the notation, let us set, for a generic mapping ϕ,

$$[Q_{1,\varepsilon}\phi](u) := \frac{\phi(u_\varepsilon) - \phi(u + \varepsilon y)}{\varepsilon}, \qquad [Q_{2,\varepsilon}\phi](u) := \frac{\phi(u + \varepsilon y) - \phi(u)}{\varepsilon}$$

(with obvious modifications if u and y are replaced by u_- and y_-), and note that

$$\frac{\phi(u_\varepsilon) - \phi(u)}{\varepsilon} = [Q_{1,\varepsilon}\phi](u) + [Q_{2,\varepsilon}\phi](u)$$

(the formal operators $Q_{1,\varepsilon}$ and $Q_{2,\varepsilon}$ clearly depend also on y, but we do not need to explicitly denote this fact). Recalling the elementary estimate of Lemma 2.2, one has

$$\begin{aligned}
I_1 &\lesssim_p \big\|S * Q_{1,\varepsilon} f(u)\big\|_{\mathbb{S}^p(t_0,t_1)} + \big\|S * \big(Q_{2,\varepsilon} f(u) - D_{\mathcal{G}} f(u)y\big)\big\|_{\mathbb{S}^p(t_0,t_1)}\\
&\le \big\|Q_{1,\varepsilon} f(u)\big\|_{L^p(\Omega;L^1(t_0,t_1;H))} + \big\|Q_{2,\varepsilon} f(u) - D_{\mathcal{G}} f(u)y\big\|_{L^p(\Omega;L^1(t_0,t_1;H))}\\
&=: I_{11} + I_{12},
\end{aligned}$$

where, by the Lipschitz continuity of f,

$$I_{11} = \big\|Q_{1,\varepsilon} f(u)\big\|_{L^p(\Omega;L^1(t_0,t_1;H))} \lesssim (t_1 - t_0)\big\|\varepsilon^{-1}(u_\varepsilon - u - \varepsilon y)\big\|_{\mathbb{S}^p(t_0,t_1)}.$$

The terms I_2 and I_3 can be handled similarly, thanks to the maximal inequalities of Sect. 2.3:

$$\begin{aligned}
I_2 &\lesssim_p \big\|S \diamond Q_{1,\varepsilon} B(u)\big\|_{\mathbb{S}^p(t_0,t_1)} + \big\|S \diamond \big(Q_{2,\varepsilon} B(u) - D_{\mathcal{G}} B(u)y\big)\big\|_{\mathbb{S}^p(t_0,t_1)}\\
&\lesssim_p \big\|Q_{1,\varepsilon} B(u)\big\|_{L^p(\Omega;L^2(t_0,t_1;\mathscr{L}^2(K,H)))} + \big\|Q_{2,\varepsilon} B(u)\\
&\quad - D_{\mathcal{G}} B(u)y\big\|_{L^p(\Omega;L^2(t_0,t_1;\mathscr{L}^2(K,H)))}\\
&=: I_{21} + I_{22},
\end{aligned}$$

where

$$I_{21} = \big\|Q_{1,\varepsilon} B(u)\big\|_{L^p(\Omega;L^2(t_0,t_1;\mathscr{L}^2(K,H)))} \lesssim (t_1 - t_0)^{1/2}\big\|\varepsilon^{-1}(u_\varepsilon - u - \varepsilon y)\big\|_{\mathbb{S}^p(t_0,t_1)},$$

and

$$\begin{aligned} I_3 &\lesssim_p \left\| S \diamond_\mu Q_{1,\varepsilon} G(u_-) \right\|_{\mathbb{S}^p(t_0,t_1)} + \left\| S \diamond_\mu \left(Q_{2,\varepsilon} G(u_-) - D_{\mathcal{G}} G(u_-) y_- \right) \right\|_{\mathbb{S}^p(t_0,t_1)} \\ &\lesssim_p \left\| Q_{1,\varepsilon} G(u_-) \right\|_{\mathbb{G}^p(t_0,t_1)} + \left\| Q_{2,\varepsilon} G(u_-) - D_{\mathcal{G}} G(u_-) y_- \right\|_{\mathbb{G}^p(t_0,t_1)} \\ &=: I_{31} + I_{32}, \end{aligned}$$

where

$$I_{31} \le \kappa(t_1 - t_0) \left\| \varepsilon^{-1}(u_\varepsilon - u - \varepsilon y) \right\|_{\mathbb{S}^p(t_0,t_1)}.$$

Recalling that κ is continuous with $\kappa(0) = 0$, these estimates imply that for every $\sigma > 0$ there exists $\delta > 0$ such that, for any $t_0 < t_1$ with $t_1 - t_0 < \delta$, one has

$$\begin{aligned} \left\| \varepsilon^{-1}(u_\varepsilon - u - \varepsilon y) \right\|_{\mathbb{S}^p(t_0,t_1)} \lesssim \sigma \left\| \varepsilon^{-1}(u_\varepsilon - u - \varepsilon y) \right\|_{\mathbb{S}^p(t_0,t_1)} \\ + \left\| \varepsilon^{-1}(u_\varepsilon(t_0) - u(t_0) - \varepsilon y(t_0)) \right\|_{L^p(\Omega;H)} \\ + I_{12} + I_{22} + I_{32}. \end{aligned}$$

Fixing then σ sufficiently small and rearranging the terms yields

$$\begin{aligned} \left\| \varepsilon^{-1}(u_\varepsilon - u - \varepsilon y) \right\|_{\mathbb{S}^p(t_0,t_1)} \lesssim \left\| \varepsilon^{-1}(u_\varepsilon(t_0) - u(t_0) - \varepsilon y(t_0)) \right\|_{L^p(\Omega;H)} \\ + I_{12} + I_{22} + I_{32}. \end{aligned}$$

where the implicit constant depends on δ and I_{12}, I_{22}, I_{32} are "supported" on $[t_0, t_1]$. Let $t_0 = 0 < t_1 < \cdots < t_{N-1} < t_N = T$ be a subdivision of the interval $[0, T]$ such that $t_n - t_{n-1} < \delta$ for all n. Then, we have, for every $n \in \{1, \ldots, N\}$, with obvious meaning of the notation,

$$\begin{aligned} \left\| \varepsilon^{-1}(u_\varepsilon - u - \varepsilon y) \right\|_{\mathbb{S}^p(t_{n-1},t_n)} \lesssim \left\| \varepsilon^{-1}(u_\varepsilon(t_{n-1}) - u(t_{n-1}) - \varepsilon y(t_{n-1})) \right\|_{L^p(\Omega;H)} \\ + [I_{12} + I_{22} + I_{32}](t_{n-1}, t_n), \end{aligned}$$

where

$$\begin{aligned} &\left\| \varepsilon^{-1}(u_\varepsilon(t_{n-1}) - u(t_{n-1}) - \varepsilon y(t_{n-1})) \right\|_{L^p(\Omega;H)} \\ &\quad \le \left\| \varepsilon^{-1}(u_\varepsilon - u - \varepsilon y) \right\|_{\mathbb{S}^p(t_{n-2},t_{n-1})} \\ &\quad \lesssim \left\| \varepsilon^{-1}(u_\varepsilon(t_{n-2}) - u(t_{n-2}) - \varepsilon y(t_{n-2})) \right\|_{L^p(\Omega;H)} \\ &\qquad + [I_{12} + I_{22} + I_{32}](t_{n-2}, t_{n-1}). \end{aligned}$$

Backward recursion thus yields

$$\begin{aligned} \left\| \varepsilon^{-1}(u_\varepsilon - u - \varepsilon y) \right\|_{\mathbb{S}^p(0,T)} &\le \sum_{n=1}^{N} \left\| \varepsilon^{-1}(u_\varepsilon - u - \varepsilon y) \right\|_{\mathbb{S}^p(t_{n-1},t_n)} \\ &\lesssim \left\| \varepsilon^{-1}(u_\varepsilon(0) - u(0) - \varepsilon y(0)) \right\|_{L^p(\Omega;H)} \\ &\quad + \sum_{n=1}^{N} [I_{12} + I_{22} + I_{32}](t_{n-1}, t_n) \end{aligned}$$

where the first summand on the right-hand side is zero. To conclude the proof it suffices to show that

$$\sum_{n=1}^{N} I_{j2}(t_{n-1}, t_n) \lesssim I_{j2}(0, T)$$

for every $j \in \{1, 2, 3\}$. We shall show that this is true for I_{32}, as both other cases are entirely similar (in fact slightly simpler): it is enough to observe that, for any ϕ satisfying suitable measurability conditions and for any $q > 0$, the obvious inequality

$$\int_{Z\times[t_{n-1},t_n]} |\phi|^q \, \mathrm{d}\nu \le \int_{Z\times[0,T]} |\phi|^q \, \mathrm{d}\nu \qquad \forall n \in \{1, \dots, N\},$$

implies $\|\phi\|_{\mathbb{G}^p(t_{n-1},t)} \le \|\phi\|_{\mathbb{G}^p(0,T)}$, hence

$$\sum_{n=1}^{N} I_{32}(t_{n-1}, t_n) \le N I_{32}(0, T).$$

□

The main result of this section is the following. Note that since the (standard) definition of Gâteaux derivative requires a Banach space framework, we shall confine ourself to the case $p \in [1, +\infty[$.

Theorem 4.3. *Let $p \ge 1$ and $(\mathrm{A5}_p)$ be satisfied. Then, the solution map of (1.1) is Gâteaux differentiable from $L^p(\Omega; H)$ to $\mathbb{S}^p$, and its Gâteaux derivative at u_0 is $(h \mapsto y) \in \mathscr{L}(L^p(\Omega; H), \mathbb{S}^p)$, where y is the unique mild solution to (4.1).*

Proof. By Lemma 4.2, it is enough to show that

$$\begin{aligned} &\big\|\varepsilon^{-1}\big(f(u+\varepsilon y) - f(u) - \varepsilon D_{\mathcal{G}} f(u) y\big)\big\|_{L^p(\Omega;L^1(0,T;H))} \\ &\quad + \big\|\varepsilon^{-1}\big(B(u+\varepsilon y) - B(u) - \varepsilon D_{\mathcal{G}} B(u) y\big)\big\|_{L^p(\Omega;L^2(0,T;\mathscr{L}^2(K,H)))} \\ &\quad + \big\|\varepsilon^{-1}\big(G(u_- + \varepsilon y_-) - G(u_-) - \varepsilon D_{\mathcal{G}} G(u_-) y_-\big)\big\|_{\mathbb{G}^p} \end{aligned} \tag{4.3}$$

converges to zero as ε tends to zero. By assumption (G1) it immediately follows that, as $\varepsilon \to 0$,

$$\begin{aligned} &\big\|\varepsilon^{-1}\big(f(u+\varepsilon y) - f(u) - \varepsilon D_{\mathcal{G}} f(u) y\big)\big\| \longrightarrow 0, \\ &\big\|\varepsilon^{-1}\big(B(u+\varepsilon y) - B(u) - \varepsilon D_{\mathcal{G}} B(u) y\big)\big\|_{\mathscr{L}^2(K,H)} \longrightarrow 0 \end{aligned}$$

for a.a. $(\omega, t) \in \Omega \times [0, T]$. Moreover, recalling that the operator norms of $D_{\mathcal{G}} f$ and $D_{\mathcal{G}} B$ are bounded by the Lipschitz constants of f and B, respectively, the triangle inequality yields

$$\begin{aligned} &\big\|\varepsilon^{-1}\big(f(u+\varepsilon y) - f(u) - \varepsilon D_{\mathcal{G}} f(u) y\big)\big\| \\ &\quad + \big\|\varepsilon^{-1}\big(B(u+\varepsilon y) - B(u) - \varepsilon D_{\mathcal{G}} B(u) y\big)\big\|_{\mathscr{L}^2(K,H)} \lesssim \big(C_f + C_B\big)\|y\| \end{aligned}$$

for a.a. (ω, t). Since $y \in \mathbb{S}^p$, the right-hand side belongs to $L^p(\Omega; L^1(0,T))$ as well as to $L^p(\Omega; L^2(0,T))$; hence, the first two terms in (4.3) converge to zero as $\varepsilon \to 0$ by the dominated convergence theorem. Similarly, setting $G_1 := G_2 := G/2$ if $p \geq 2$, one has

$$\begin{aligned}&\big\|\varepsilon^{-1}\big(G(u_- + \varepsilon y_-) - G(u_-) - \varepsilon D_{\mathcal{G}} G(u_-) y_-\big)\big\|_{\mathbb{G}^p} \\ &\quad \lesssim_p \big\|\varepsilon^{-1}\big(G_1(u_- + \varepsilon y_-) - G_1(u_-) - \varepsilon D_{\mathcal{G}} G_1(u_-) y_-\big)\big\|_{L^p(\Omega; L^2(\nu; H))} \\ &\qquad + \big\|\varepsilon^{-1}\big(G_2(u_- + \varepsilon y_-) - G_2(u_-) - \varepsilon D_{\mathcal{G}} G_2(u_-) y_-\big)\big\|_{L^p(\Omega; L^p(\nu; H))},\end{aligned}$$

where the implicit constant is equal to 1 for $p \in [1, 2[$, and to 2 for $p \geq 2$. Since

$$\big\|\varepsilon^{-1}\big(G_j(u_- + \varepsilon y_-) - G_j(u_-) - \varepsilon D_{\mathcal{G}} G_j(u_-) y_-\big)\big\| \longrightarrow 0$$

as $\varepsilon \to 0$, as well as

$$\big\|\varepsilon^{-1}\big(G_j(u_- + \varepsilon y_-) - G_j(u_-) - \varepsilon D_{\mathcal{G}} G_j(u_-) y_-\big)\big\| \leq g_j \|y\|$$

for all $(t, z) \in [0, T] \times Z$, $\mathbb{P}$-almost surely, for both $j = 1$ and $j = 2$, one has, thanks to $(\mathrm{A5}_p)$ and the dominated convergence theorem, recalling that $y \in \mathbb{S}^p$,

$$\begin{aligned}\big\|\varepsilon^{-1}\big(G_1(u_- + \varepsilon y_-) - G_1(u_-) - \varepsilon D_{\mathcal{G}} G_1(u_-) y_-\big)\big\|_{L^2(\nu; H)} &\longrightarrow 0, \\ \big\|\varepsilon^{-1}\big(G_2(u_- + \varepsilon y_-) - G_2(u_-) - \varepsilon D_{\mathcal{G}} G_2(u_-) y_-\big)\big\|_{L^p(\nu; H)} &\longrightarrow 0\end{aligned}$$

$\mathbb{P}$-a.s. as $\varepsilon \to 0$. A further application of the dominated convergence theorem hence yields that the third term in (4.3) converges to zero as $\varepsilon \to 0$, thus completing the proof. □

5. Fréchet differentiability of the solution map

We are going to show that the Fréchet differentiability of the coefficients of (1.1) implies the Fréchet differentiability of the solution map. We shall work under the following assumption that is assumed to hold throughout this section.

(F) The maps $f(\omega, t, \cdot)$ and $B(\omega, t, \cdot)$ are Fréchet-differentiable for all $(\omega, t) \in \Omega \times [0, T]$, and the maps

$$G(\omega, t, z, \cdot), \quad G_1(\omega, t, z, \cdot), \quad G_2(\omega, t, z, \cdot)$$

are Fréchet-differentiable for all $(\omega, t, z) \in \Omega \times [0, T] \times Z$.

The Fréchet derivatives of f and B (in their H-valued argument), denoted by

$$\begin{aligned}Df &: \Omega \times [0, T] \times H \longrightarrow \mathscr{L}(H, H), \\ DB &: \Omega \times [0, T] \times H \longrightarrow \mathscr{L}(H, \mathscr{L}^2(K, H)),\end{aligned}$$

satisfy the boundedness properties

$$\begin{aligned}\big\|Df(\omega,t,x)\big\|_{\mathscr{L}(H,H)} &\le C_f,\\ \big\|DB(\omega,t,x)\big\|_{\mathscr{L}(H,\mathscr{L}^2(K,H))} &\le C_B\end{aligned}$$

for all $(\omega,t,x)\in\Omega\times[0,T]\times H$ (see Sect. 2.2). Similarly, and in complete analogy to the previous section, the Lipschitz continuity assumptions on G, G_1 and G_2 imply that,

$$\begin{aligned}\big\|DG(\omega,t,z,x)\big\|_{\mathscr{L}(H,H)} &\le g(\omega,t,z), \quad p\ge 2,\\ \big\|DG_j(\omega,t,z,x)\big\|_{\mathscr{L}(H,H)} &\le g_j(\omega,t,z), \quad p\in[1,2[, \quad j=1,2.\end{aligned}$$

The main result of this section is the following theorem, which states that the solution map is Fréchet-differentiable along subspaces of vectors with finite higher moments.

Theorem 5.1. *Let $q>p\ge 1$. If $(\mathrm{A5}_p)$ and $(\mathrm{A5}_q)$ hold, then the solution map of (1.1) is Fréchet-differentiable from $L^p(\Omega;H)$ to $\mathbb{S}^p$ along $L^q(\Omega;H)$ and its Fréchet derivative at $u_0\in L^p(\Omega;H)$ is the map $h\mapsto y\in\mathscr{L}(L^q(\Omega;H),\mathbb{S}^p)$, where y is the unique mild solution to the stochastic evolution equation*

$$dy+Ay\,dt=Df(u)y\,dt+DB(u)y\,dW+\int_Z DG(u_-)y_-\,d\bar{\mu}, \qquad y(0)=h. \tag{5.1}$$

Proof. For any $h\in L^q(\Omega;H)$, Eq. (5.1) admits a unique mild solution $y\in\mathbb{S}^q$, as it follows immediately by the boundedness properties of the Fréchet derivatives of f, B and G, and by hypothesis $(\mathrm{A5}_q)$. Therefore, the map $h\mapsto y$ is well defined from $L^q(\Omega;H)$ to $\mathbb{S}^q$, and it is obviously linear and continuous. To prove that this map is the Fréchet derivative of the solution map $u_0\mapsto u$, thanks to the characterization of Fréchet differentiability of Lemma 2.1, it is enough to show that

$$\lim_{\varepsilon\to 0}\big\|\varepsilon^{-1}(u_\varepsilon-u-\varepsilon y)\big\|_{\mathbb{S}^p}=0$$

uniformly over h belonging to bounded subsets of $L^q(\Omega;H)$. By Lemma 4.2, for this it suffices to show that each term in (4.3) converges to zero uniformly with respect to h belonging to the unit ball of $L^q(\Omega;H)$. Since $h\mapsto y\in\mathscr{L}(L^q(\Omega;H),\mathbb{S}^q)$, it is evident that if h belongs to $B_1(L^q(\Omega;H))$, then $y(h)$ belongs to $B_R(\mathbb{S}^q)$, where $R:=\big\|h\mapsto y\big\|_{\mathscr{L}(L^q(\Omega;H),\mathbb{S}^q)}$. Hence, denoting by I_j, $j=1,2,3$, the terms appearing in (4.3), by homogeneity

$$\sup_{h\in B_1(L^q(\Omega;H))}(I_1+I_2+I_3)(y(h))\le\sup_{y\in B_R(\mathbb{S}^q)}(I_1+I_2+I_3)(y).$$

Hence, it suffices to show that I_1, I_2 and I_3 converge to zero uniformly with respect to y bounded in $\mathbb{S}^q$. That is, we need to show that, for any $R>0$ and $\vartheta>0$, there exists $\varepsilon_0=\varepsilon_0(R,\vartheta)$ such that $|\varepsilon|<\varepsilon_0$ implies $I_j(y)<\vartheta$ for all $y\in B_R(\mathbb{S}^q)$ and $j\in\{1,2,3\}$.

For any measurable $E \subset \Omega$, one clearly has

$$\begin{aligned} I_1 &= \big\|\varepsilon^{-1}\big(f(u+\varepsilon y) - f(u) - \varepsilon Df(u)y\big)\big\|_{L^p(\Omega;L^1(0,T;H))} \\ &\le \big\|\varepsilon^{-1}\big(f(u+\varepsilon y) - f(u) - \varepsilon Df(u)y\big)\big\|_{L^p(E;L^1(0,T;H))} \\ &\quad + \big\|\varepsilon^{-1}\big(f(u+\varepsilon y) - f(u) - \varepsilon Df(u)y\big)\big\|_{L^p(E^c;L^1(0,T;H))} \\ &=: I_1(E) + I_1(E^c), \end{aligned}$$

where, by the Lipschitz continuity of f,

$$\begin{aligned} I_1(E)^p &= \mathbb{E}\,1_E\bigg(\int_0^T \big\|(f(u+\varepsilon y) - f(u))/\varepsilon - Df(u)y\big\|\,\mathrm{d}t\bigg)^p \\ &\le \big(2TC_f\big)^p\,\mathbb{E}\,1_E\,(y_T^*)^p. \end{aligned}$$

The set $Y := \{(y_T^*)^p : y \in B_R(\mathbb{S}^q)\}$ is bounded in $L^{q/p}(\Omega)$, with $q > p$, hence uniformly integrable on $(\Omega, \mathscr{F}, \mathbb{P})$. In particular, for any $\vartheta > 0$ there exists $\sigma > 0$ such that, for any $E \in \mathscr{F}$ with $\mathbb{P}(E) < \sigma$, one has

$$\mathbb{E}\,1_E\,(y_T^*)^p < \bigg(\frac{\vartheta}{2C_fT}\bigg)^p, \qquad \forall y \in B_R(\mathbb{S}^q),$$

hence $I_1(E) \le \vartheta$.

Let $y \in B_R(\mathbb{S}^q)$ be arbitrary but fixed. Markov's inequality yields, for any $n > 0$,

$$\mathbb{P}\big(y_T^* > n\big) \le \frac{\mathbb{E}(y_T^*)^q}{n^q} \le \frac{R^q}{n^q}.$$

Therefore, there exists $n > 0$ such that, setting $E := \{y_T^* > n\}$, one has $I_1(E) < \vartheta$. It is important to note that n depends on R, but not on y, while E depends on y. The Fréchet differentiability hypothesis on f amounts to saying that, for any $x \in H$ and $n \in \mathbb{N}$,

$$\lim_{\varepsilon\to 0} \sup_{z\in B_n(H)} \big\|(f(x+\varepsilon z) - f(x))/\varepsilon - Df(x)z\big\| = 0.$$

In particular, one has

$$\lim_{\varepsilon\to 0} \sup_{z\in B_n(H)} \big\|(f(u(\omega,t)+\varepsilon z) - f(u(\omega,t)))/\varepsilon - Df(u(\omega,t))z\big\| = 0$$

for a.a. $(\omega, t) \in E^c \times [0,T]$, where, by the Lipschitz continuity of f,

$$\sup_{z\in B_n(H)} \big\|(f(u(\omega,t)+\varepsilon z) - f(u(\omega,t)))/\varepsilon - Df(u(\omega,t))z\big\| \lesssim 2nC_f,$$

for a.a. $(\omega, t) \in E^c \times [0,T]$. Therefore, by the dominated convergence theorem,

$$\lim_{\varepsilon\to 0} \bigg\| \sup_{z\in B_n(H)} \big\|(f(u(\omega,t)+\varepsilon z) - f(u(\omega,t)))/\varepsilon - Df(u(\omega,t))z\big\| \bigg\|_{L^p(E^c;L^1(0,T))} = 0,$$

that is, for any $\vartheta > 0$ there exists ε_1 depending only on ϑ and n such that

$$\left\| \sup_{z\in B_n(H)} \left\| (f(u+\varepsilon z) - f(u))/\varepsilon - Df(u)z \right\| \right\|_{L^p(E^c;L^1(0,T))} < \vartheta$$

for all ε such that $|\varepsilon| < \varepsilon_1(\vartheta, n)$. It remains to observe that

$$\begin{aligned} &\left\| (f(u(\omega,t) + \varepsilon y(\omega,t)) - f(u(\omega,t)))/\varepsilon - Df(u(\omega,t))y(\omega,t) \right\| \\ &\quad \le \sup_{z\in B_n(H)} \left\| (f(u(\omega,t) + \varepsilon z) - f(u(\omega,t)))/\varepsilon - Df(u(\omega,t))z \right\| \end{aligned}$$

for a.a. $(\omega, t) \in E^c \times [0, T]$ to get that $I_1(E^c) < \vartheta$ for all ε such that $|\varepsilon| < \varepsilon_1(\vartheta, n)$. Since n depends only on R, we conclude that there exists $\varepsilon_1 = \varepsilon_1(\vartheta, R)$ such that $I_1 < 2\vartheta$ for all $|\varepsilon| < \varepsilon_1(\vartheta, R)$.

Let us now consider the term I_2: the argument is similar to the one just carried out, so we provide slightly less detail. We have to show that I_2 converges to 0 uniformly with respect to $y \in B_R(\mathbb{S}^q)$. For any measurable $E \subset \Omega$, one has, with obvious meaning of the notation,

$$I_2 \le I_2(E) + I_2(E^c),$$

where, by the Lipschitz continuity of B,

$$I_2(E)^p \le (2C_B)^p T^{p/2}\, \mathbb{E}\, 1_E\, (y_T^*)^p.$$

Choosing E as before, using the uniform integrability of the family Y combined with the Markov inequality, we infer that for any $\vartheta > 0$ there exists $n > 0$ such that $I_2(E) < \vartheta$. The Fréchet differentiability of B implies that, for any $x \in H$,

$$\lim_{\varepsilon\to 0} \sup_{z\in B_n(H)} \left\| (B(x+\varepsilon z) - B(x))/\varepsilon - DB(x)z \right\|_{\mathscr{L}^2(K,H)} = 0$$

in $E^c \times [0, T]$, where, by the Lipschitz continuity of B,

$$\sup_{z\in B_n(H)} \left\| (B(u(\omega,t) + \varepsilon z) - B(u(\omega,t)))/\varepsilon - DB(u(\omega,t))z \right\|_{\mathscr{L}^2(K,H)} \lesssim 2nC_B,$$

for a.a. $(\omega, t) \in E^c \times [0, T]$. Hence, the dominated convergence theorem yields

$$\begin{aligned} \lim_{\varepsilon\to 0} \Bigg\| \sup_{z\in B_n(H)} \Big\| &(B(u(\omega,t) + \varepsilon z) - B(u(\omega,t)))/\varepsilon \\ &- DB(u(\omega,t))z \Big\|_{\mathscr{L}^2(K,H)} \Bigg\|_{L^p(E^c;L^2(0,T))} = 0, \end{aligned}$$

that is, for any $\vartheta > 0$ there exists ε_2 depending only on ϑ and n such that

$$\left\| \sup_{z\in B_n(H)} \left\| (B(u+\varepsilon z) - B(u))/\varepsilon - DB(u)z \right\|_{\mathscr{L}^2(K,H)} \right\|_{L^p(E^c;L^2(0,T))} < \vartheta$$

for all ε such that $|\varepsilon| < \varepsilon_2(\vartheta, n)$, from which also $I_2(E^c) < \vartheta$ for all ε such that $|\varepsilon| < \varepsilon_2(\vartheta, n)$. Hence, there exists $\varepsilon_2 = \varepsilon_2(\vartheta, R)$ such that $I_2 < 2\vartheta$ for all ε with $|\varepsilon| < \varepsilon_2(\vartheta, R)$.

The convergence to zero of I_3 as $\varepsilon \to 0$, uniformly with respect to $y \in B_R(\mathbb{S}^q)$, while still similar to the above arguments, is slightly more delicate as random measures are involved. As already shown in the proof of Theorem 4.3, one has, recalling that Fréchet differentiability implies Gâteaux differentiability,

$$\left\|\varepsilon^{-1}\big(G(u_- + \varepsilon y_-) - G(u_-)\big) - DG(u_-)y_-\right\|_{\mathbb{G}^p} \longrightarrow 0$$

as $\varepsilon \to 0$. We need to show that the convergence holds uniformly over y bounded in $\mathbb{S}^q$. Let $R > 0$ and $y \in B_R(\mathbb{S}^q)$. For any measurable $E \in \mathscr{F}$, the Lipschitz continuity assumptions on G and (A5$_p$) imply, setting $G_1 := G_2 := G$ if $p \geq 2$, that

$$\begin{aligned}
&\left\|1_E\big(\varepsilon^{-1}\big(G(v + \varepsilon w) - G(v)\big) - DG(v)w\big)\right\|_{\mathbb{G}^p}^p \\
&\quad \leq \mathbb{E}\, 1_E \int \left\|\varepsilon^{-1}\big(G_1(u_- + \varepsilon y_-) - G_1(u_-)\big) - DG_1(u_-)y_-\right\|^p \mathrm{d}\nu \\
&\qquad + \mathbb{E}\, 1_E \left(\int \left\|\varepsilon^{-1}\big(G_2(u_- + \varepsilon y_-) - G_2(u_-)\big) - DG_2(u_-)y_-\right\|^2 \mathrm{d}\nu\right)^{p/2} \\
&\quad \lesssim_p \kappa(T)^p\, \mathbb{E}\, 1_E (y_T^*)^p.
\end{aligned}$$

As the set $\{(y_T^*)^p : y \in B_R(\mathbb{S}^q)\}$ is bounded in $L^{q/p}(\Omega)$, hence uniformly integrable, for any $\vartheta > 0$ there exists $n > 0$ (by Markov's inequality) such that, choosing $E := \{y^* > n\}$ as before, we have

$$\kappa(T)^p\, \mathbb{E}\, 1_E (y_T^*)^p < \vartheta.$$

On E^c one has, possibly outside a set of $\mathbb{P}$-measure zero, for both $j = 1$ and $j = 2$,

$$\begin{aligned}
&\left\|\varepsilon^{-1}\big(G_j(u_- + \varepsilon y_-) - G_j(u_-)\big) - DG_j(u_-)y_-\right\| \\
&\quad \leq \sup_{\xi \in B_n(H)} \left\|\varepsilon^{-1}\big(G_j(u_- + \varepsilon\xi) - G_j(u_-)\big) - DG_j(u_-)\xi\right\|
\end{aligned}$$

where the right-hand side converges to zero by the characterization of Fréchet differentiability of Lemma 2.1 and is bounded by $2ng_j$ for all $(t, z) \in [0, T] \times Z$. Since $g_1 \in L^p(\nu)$ and $g_2 \in L^2(\nu)$ $\mathbb{P}$-a.s. in E^c, the dominated convergence theorem and (A5$_p$) yield

$$\left\|1_{E^c}\big(\varepsilon^{-1}\big(G(u_- + \varepsilon y_-) - G(u_-)\big) - DG(u_-)y_-\big)\right\|_{\mathbb{G}^p} \longrightarrow 0$$

as $\varepsilon \to 0$, uniformly with respect to $y \in B_R(\mathbb{S}^q)$. Proceeding exactly as in the case of I_1, we conclude that there exists $\varepsilon_3 = \varepsilon_3(\vartheta, R)$ such that $I_3 < 2\vartheta$ for all $|\varepsilon| < \varepsilon_3$.

We have thus shown that $\varepsilon^{-1}(u_\varepsilon - u - \varepsilon y) \to 0$ in $\mathbb{S}^p(0, T)$, uniformly over h in any bounded subset of $L^q(\Omega; H)$, as claimed. □

6. Fréchet differentiability of higher order

In this section we show that the n-th-order Fréchet differentiability of the coefficients of (1.1), in a suitable sense, implies the n-th-order Fréchet differentiability of the solution map. We shall work under the following assumptions that are stated in terms of the parameter $n \in \mathbb{N}$, $n \geq 2$:

(F^n) The maps $f(\omega, t, \cdot)$ and $B(\omega, t, \cdot)$ are n times Fréchet-differentiable for all $(\omega, t) \in \Omega \times [0, T]$, and the maps $G(\omega, t, z, \cdot)$, $G_i(\omega, t, z, \cdot)$, $i = 1, 2$, are n times Fréchet differentiable for all $(\omega, t, z) \in \Omega \times [0, T] \times Z$. Moreover, there exists a constant $m \geq 0$ such that, for every $j = 2, \ldots, n$,

$$\left\| D^j f(\omega, t, x) \right\|_{\mathscr{L}_j(H;H)} + \left\| D^j B(\omega, t, x) \right\|_{\mathscr{L}_j(H;\mathscr{L}^2(K,H))} \lesssim 1 + \|x\|^m$$

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

$$\left\| D^j G(\omega, t, z, x) \right\|_{\mathscr{L}_j(H;H)} \lesssim g(\omega, t, z)\big(1 + \|x\|^m\big),$$
$$\left\| D^j G_i(\omega, t, z, x) \right\|_{\mathscr{L}_j(H;H)} \lesssim g_i(\omega, t, z)\big(1 + \|x\|^m\big), \qquad i = 1, 2.$$

We also stipulate that (F^1) is simply hypothesis (F) of the previous section. It would be possible to replace the functions g, g_1 and g_2 with different ones, thus reaching a bit more generality, but it does not seem to be worth the (mostly notational) effort.

Example 6.1. Let us give an explicit example where assumption (F^n) is satisfied with a suitable choice of $m > 0$ and not for $m = 0$. We shall consider $B = G = 0$ for simplicity and concentrate only on f: typical examples for B and G can be produced following the same argument. Let $H = L^2(D)$, where $D \subset \mathbb{R}^d$ is a smooth bounded domain, and consider the function

$$\gamma : \mathbb{R} \to \mathbb{R}, \qquad \gamma(r) := \int_0^r \sin(s^2)\, \mathrm{d}s, \quad r \in \mathbb{R}.$$

It is not difficult to check that $\gamma \in C^\infty(\mathbb{R})$, γ is Lipschitz-continuous (hence $\gamma' \in C_b(\mathbb{R})$), and

$$|\gamma^{(j)}(r)| \lesssim 1 + |r|^{j-1} \qquad \forall\, r \in \mathbb{R}, \quad \forall\, j \in \mathbb{N}, \quad j \geq 1.$$

However, the derivatives $\gamma^{(j)}$ are *not* bounded in $\mathbb{R}$ for any $j \geq 2$. Furthermore, let us fix $\mathcal{L} \in \mathscr{L}(H, L^\infty(D))$, and define the operator

$$f : H \to H, \qquad f(u) := \gamma(\mathcal{L}u), \quad u \in H.$$

Clearly, f is well defined, Lipschitz-continuous and linearly bounded, so that (A2) is satisfied. Moreover, using the fact that $\mathcal{L} \in \mathscr{L}(H, L^\infty(D))$ it a standard matter to check that f is Fréchet-differentiable, and its derivative is given by

$$Df : H \to \mathscr{L}(H, H), \qquad Df(u)h = \gamma'(\mathcal{L}u)\mathcal{L}h, \quad u, h \in H,$$

so that also assumption (F) is satisfied. Note in particular that the first derivative Df is also bounded in H thanks to the Lipschitz continuity of f. Furthermore, using the fact that $\mathcal{L} \in \mathscr{L}(H, L^\infty(D))$ a direct computation shows that for every $j \in \mathbb{N}$, with $j > 1$, f is Fréchet-differentiable j times and

$$D^j f(u) : H \to \mathscr{L}_j(H^j; H), \qquad D^j f(u)(h_1, \dots, h_j) = \gamma^{(j)}(\mathcal{L}u)\mathcal{L}h_1 \dots \mathcal{L}h_j,$$
$$u, h_1, \dots, h_j \in H.$$

For every $j > 1$, by the Hölder inequality and the properties of γ and $\mathcal{L}$ we have that

$$\left\| D^j f(u) \right\|_{\mathscr{L}_j(H^j;H)} \lesssim 1 + \left\| \gamma^{(j)}(\mathcal{L}u) \right\|_{L^\infty(D)} \lesssim_\gamma 1 + \left\| \mathcal{L}u \right\|_{L^\infty(D)}^{j-1} \lesssim_{\mathcal{L}} 1 + \left\| u \right\|_H^{j-1},$$

so that assumption (F^n) is satisfied for every n with the choice $m = n - 1$. However, note that the higher-order derivatives of f are *not* bounded in H because of the choice of the function γ: hence, coefficients f in this form cannot be treated using available results in literature (as, for example, [15]). On the other hand, these are nonetheless included in our analysis.

In the following we shall write, for compactness of notation, $\mathbb{L}^q$ in place of $L^q(\Omega; H)$. If u (identified with the solution map $u_0 \mapsto u : \mathbb{L}^p \to \mathbb{S}^p$, which is well defined if assumption ($\mathrm{A5}_p$) holds) is n times Fréchet differentiable along $\mathbb{L}^{q_1}, \dots, \mathbb{L}^{q_n}$, we have

$$D^n u(u_0) \in \mathscr{L}_n\left(\mathbb{L}^{q_1}, \dots, \mathbb{L}^{q_n}; \mathbb{S}^p\right).$$

Under the assumptions of Theorem 5.1, u is once Fréchet differentiable and $v := Du(u_0)$ satisfies the equation

$$v = S(\cdot)I + S * Df(u)v + S \diamond DB(u)v + S \diamond_\mu DG(u_-)v_-,$$

where I is the identity map. This equation has to be interpreted in the sense that, for any $h \in \mathbb{L}^q$, $q > p$, setting $y := [Du(u_0)]h$, one has

$$y = S(\cdot)h + S * Df(u)y + S \diamond DB(u)y + S \diamond_\mu DG(u_-)y_-.$$

Note that by Lemma 4.1 this equation admits a unique solution $y \in \mathbb{S}^p$ also for $h \in \mathbb{L}^p$, and that $h \mapsto y \in \mathscr{L}(\mathbb{L}^p, \mathbb{S}^p)$. However, if h belongs only to $\mathbb{L}^p$, we can no longer claim that $h \mapsto y$ is the Fréchet derivative of $u_0 \mapsto u$, as Theorem 5.1 does not necessarily apply.

We are now going to introduce a system of equations, indexed by $n \geq 2$, that are formally expected to be satisfied by $D^j u(u_0)$, $j = 1, \dots, n$, if they exist. For any $n \geq 2$, the equation for $u^{(n)}$ can be written as

$$\begin{aligned} \mathrm{d}u^{(n)} + Au^{(n)}\,\mathrm{d}t = \left(\Psi_n + Df(u)u^{(n)}\right)\mathrm{d}t + \left(\Phi_n + DB(u)u^{(n)}\right)\mathrm{d}W \\ + \int_Z \left(\Theta_n + DG(u_-)u_-^{(n)}\right) d\bar{\mu}, \qquad u^{(n)}(0) = 0, \end{aligned} \tag{6.1}$$

where Ψ_n, Φ_n and Θ_n are the formal n-th Fréchet derivatives of $f(u)$, $B(u)$ and $G(u_-)$, respectively, excluding the terms involving the (formal) derivative of u of order n. More precisely, assume that E_1, E_2 and E_3 are Banach spaces and $\phi : E_1 \to E_2$, $F : E_2 \to E_3$ are n times Fréchet-differentiable. The chain rule implies that there exists a function $\tilde{\Phi}_n^F$ such that

$$D^n[F(\phi)] = \tilde{\Phi}_n^F\big(D\phi, \dots, D^{n-1}\phi\big) + DF(\phi)D^n\phi.$$

We set $\Phi_n := \tilde{\Phi}_n^B\big(u^{(1)}, u^{(2)}, \dots, u^{(n-1)}\big)$. The definition of Ψ_n and Θ_n is, *mutatis mutandis*, identical.

The concept of solution for Eq. (6.1) is intended as in the case of the first-order derivative equation, i.e., in the sense of testing against arbitrary directions. More precisely, we shall say that

$$u^{(n)} \in \mathscr{L}_n\big(\mathbb{L}^{q_1}, \dots, \mathbb{L}^{q_n}; \mathbb{S}^p\big), \qquad p, q_1, \dots, q_n \geq 1,$$

is a solution to (6.1) if, for any

$$(h_1, \dots, h_n) \in \mathbb{L}^{q_1} \times \cdots \times \mathbb{L}^{q_n},$$

the process $u^{(n)}(h_1, \dots, h_n) \in \mathbb{S}^p$ satisfies

$$\begin{aligned} u^{(n)}(h_1, \dots, h_n) = {} & S * \Psi_n(h_1, \dots, h_n) + S * Df(u)u^{(n)}(h_1, \dots, h_n) \\ & + S \diamond \Phi_n(h_1, \dots, h_n) + S \diamond DB(u)u^{(n)}(h_1, \dots, h_n) \\ & + S \diamond_\mu \Theta_n(h_1, \dots, h_n) + S \diamond_\mu DG(u_-)u_-^{(n)}(h_1, \dots, h_n). \end{aligned}$$

Let us show some properties of the coefficients Ψ_n, Φ_n and Θ_n. We are going to use some algebraic properties of the "representing" map $\tilde{\Phi}_n^F$. In particular, although a (kind of) explicit expression for $\tilde{\Phi}_n^F$ can be written in terms of a variant of the Faà di Bruno formula (as it was done, for example, in [2]), for our purposes it suffices to know that $\tilde{\Phi}_n^F$ is a sum of terms of the form

$$D^j F(\phi)\big(D^{\alpha_1}\phi, \dots, D^{\alpha_j}\phi\big),$$

with $j \in \{2, \dots, n\}$, $\alpha_1 + \cdots + \alpha_j = n$, $\alpha_i \geq 1$ for all $i \in \{1, \dots, j\}$. Moreover, since $D^n[F(\phi)]$ is an n-linear map on E_1^n with values in E_3 (with E_1^n being the cartesian product of E_1 by itself n times), one has that, for any $(h_1, \dots, h_n) \in E_1^n$, $D^n[F(\phi)](h_1, \dots, h_n)$ is a sum of terms of the form

$$D^j F(\phi)\big(D^{\alpha_1}\phi(h_{\sigma(1)}, \dots, h_{\sigma(\alpha_1)}), \dots, D^{\alpha_j}\phi(h_{\sigma(A_j+1)}, \dots, h_{\sigma(n)})\big),$$

where $A_j := \alpha_1 + \cdots + \alpha_{j-1}$, and σ is an element of the permutation group of $\{1, \dots, n\}$. We shall also need the following identities, that we write already in the specific form needed later, although they are obviously a consequence of the definition

of $\tilde{\Phi}_n^F$:

$$\begin{aligned} D\Psi_n &= \Psi_{n+1} - D^2 f(u)(u', u^{(n)}), \\ D\Phi_n &= \Phi_{n+1} - D^2 B(u)(u', u^{(n)}), \\ D\Theta_n &= \Theta_{n+1} - D^2 G(u_-)(u'_-, u_-^{(n)}), \end{aligned} \tag{6.2}$$

where we have written, as customary, u' in place of $u^{(1)}$. We are going to write, for the convenience of the reader, the first three formal derivatives of $B(u)$ and the expressions for Φ_n (the corresponding calculations for $f(u)$, $G(u_-)$, Ψ_n, and Θ_n are entirely analogous). One has

$$\begin{aligned} D[B(u)] &= DB(u)u', \\ D^2[B(u)] &= D^2B(u)\big(u', u'\big) + DB(u)u^{(2)}, \\ D^3[B(u)] &= D^3B(u)\big(u', u', u'\big) + D^2B(u)\big(u^{(2)}, u'\big) + D^2B(u)\big(u', u^{(2)}\big) \\ &\quad + D^2B(u)\big(u', u^{(2)}\big) + DB(u)u^{(3)} \\ &= D^3B(u)\big(u', u', u'\big) + 3D^2B(u)\big(u^{(2)}, u'\big) + DB(u)u^{(3)}, \\ \Phi_1 &= 0, \\ \Phi_2 &= D^2B(u)\big(u', u'\big), \\ \Phi_3 &= D^3B(u)\big(u', u', u'\big) + 3D^2B(u)\big(u^{(2)}, u'\big), \end{aligned}$$

where we have used Schwarz's theorem on the symmetry of higher-order continuous Fréchet derivatives.

The first result that we present concerns the existence and uniqueness of solutions to Eq. (6.1) in the sense specified above. More precisely, we show in the next proposition that Eq. (6.1) admits a unique solution $u^{(n)}$, belonging to $\mathscr{L}_n\big(\mathbb{L}^{p_1}, \dots, \mathbb{L}^{p_n}; \mathbb{S}^p\big)$. Note that to study differentiability we shall restrict to the case $p_1 = \cdots = p_n$ (see Remark 6.3 below). However, since well-posedness for linear stochastic equations for multilinear maps such as (6.1) could be interesting in its own right, we shall provide a general result considering arbitrary $p_1, \dots, p_n$.

Proposition 6.2. *Let $n \geq 1$ and p, p_0, $p_1, \dots, p_n \geq 1$ be such that $u_0 \in \mathbb{L}^p \cap \mathbb{L}^{mp_0} = \mathbb{L}^{p \vee mp_0}$ and*

$$\frac{n-1}{p_0} + \frac{1}{p_1} + \cdots + \frac{1}{p_n} \leq \frac{1}{p}. \tag{6.3}$$

Assume that

(i) *hypothesis* (F^n) *is satisfied;*
(ii) *hypothesis* (A5$_r$) *holds for all $r \in [p, \max_{i \geq 1} p_i] \cup \{mp_0\}$.*

Then, (6.1) *admits a unique solution*

$$u^{(n)} \in \mathscr{L}_n\big(\mathbb{L}^{p_1}, \dots, \mathbb{L}^{p_n}; \mathbb{S}^p\big).$$

Proof. First of all, let us explain why $u^{(n)}$, if it exists, must be n-linear (in the algebraic sense). Since $u' = Du$ is indeed a linear map, we can use induction as follows: assuming that $u^{(j)}$ is j-linear for all $j < k$, with $k \in \{2, \dots, n\}$, we are going to show that $u^{(k)}$ is k-linear. The inductive assumption and the functional form of Ψ_k, Φ_k, and Θ_k imply that they are k-linear. Considering the equation

$$\begin{aligned} v = {} & S * \Psi_k(h_1, \dots, h_k) + S * Df(u)v \\ & + S \diamond \Phi_k(h_1, \dots, h_k) + S \diamond DB(u)v \\ & + S \diamond_\mu \Theta_k(h_1, \dots, h_k) + S \diamond_\mu DG(u_-)v_-, \end{aligned}$$

assuming that a solution exists for every $(h_1, \dots, h_k) \in \mathbb{L}^{q_1} \times \cdots \times \mathbb{L}^{q_k}$, $q_1, \dots, q_k \geq 1$, it suffices to show that the map $(h_1, \dots, h_k) \mapsto v$ is k-linear, which is immediate.

Let us focus now on existence. We are going to reason by induction on the order of (formal) derivation $k \in \{1, \dots, n\}$. The claim is certainly true for $k = 1$: Theorem 4.3 implies, thanks for assumption (ii), that $u' \in \mathscr{L}(\mathbb{L}^r, \mathbb{S}^r)$ for every $r \in [p, \max_{i \geq 1} p_i]$, hence also $u' \in \mathscr{L}(\mathbb{L}^s, \mathbb{S}^r)$ for every $s \geq r$, as then $\mathbb{L}^s$ is contractively embedded in $\mathbb{L}^r$.

Let us now assume the claim is true for all $j \leq k \in \{1, \dots, n-1\}$, and consider $h_j \in \mathbb{L}^{p_j}$ with $p_j \geq p$, for $j = 1, \dots, k+1$, such that

$$\frac{k}{p_0} + \frac{1}{p_1} + \cdots + \frac{1}{p_{k+1}} \leq \frac{1}{p}.$$

In order to control the $\mathbb{S}^p$ norm of $u^{(k+1)}(h_1, \dots, h_{k+1})$ it is enough to estimate

$$\begin{aligned} & \big\| \Psi_{k+1}(h_1, \dots, h_{k+1}) \big\|_{L^p(\Omega; L^1(0,T;H))}, \\ & \big\| \Phi_{k+1}(h_1, \dots, h_{k+1}) \big\|_{L^p(\Omega; L^2(0,T; \mathscr{L}^2(K,H)))}, \\ & \big\| \Theta_{k+1}(h_1, \dots, h_{k+1}) \big\|_{\mathbb{G}^p}. \end{aligned}$$

In fact, recalling that $Df(u)$, $DB(u)$ and $DG(u)$ are bounded linear operators (in the same sense as in the proofs of Theorems 4.3 and 5.1), one has, for any $[t_0, t_1] \subseteq [0, T]$, omitting the indication of the arguments (h_j) for simplicity of notation,

$$\begin{aligned} \big\| u^{(k+1)} \big\|_{\mathbb{S}^p(t_0,t_1)} \leq {} & \big\| u^{(k+1)}(t_0) \big\|_{\mathbb{L}^p} + \big\| S * \Psi_{k+1} \big\|_{\mathbb{S}^p(t_0,t_1)} + \big\| S * Df(u)u^{(k+1)} \big\|_{\mathbb{S}^p(t_0,t_1)} \\ & + \big\| S \diamond \Phi_{k+1} \big\|_{\mathbb{S}^p(t_0,t_1)} + \big\| S \diamond DB(u)u^{(k+1)} \big\|_{\mathbb{S}^p(t_0,t_1)} \\ & + \big\| S \diamond_\mu \Theta_{k+1} \big\|_{\mathbb{S}^p(t_0,t_1)} + \big\| S \diamond_\mu DG(u)u^{(k+1)} \big\|_{\mathbb{S}^p(t_0,t_1)} \\ \lesssim {} & \big\| \Psi_{k+1} \big\|_{L^p(\Omega; L^1(t_0,t_1;H))} \\ & + \big\| \Phi_{k+1} \big\|_{L^p(\Omega; L^2(t_0,t_1; \mathscr{L}^2(K,H)))} + \big\| \Theta_{k+1} \big\|_{\mathbb{G}^p(t_0,t_1)} \\ & + \big\| u^{(k+1)}(t_0) \big\|_{\mathbb{L}^p} + \kappa(t_1 - t_0) \big\| u^{(k+1)} \big\|_{\mathbb{S}^p(t_0,t_1)}, \end{aligned}$$

where the implicit constant does not depend on $t_1 - t_0$ (and also not on k). We proceed now as in the proof of Lemma 4.2: choosing $T_0 > 0$ sufficiently small and partitioning

$[0, T]$ in intervals of length not exceeding T_0, it follows from $u^{(k+1)}(0) = 0$ that

$$\|u^{(k+1)}\|_{\mathbb{S}^p} \lesssim \|\Psi_{k+1}\|_{L^p(\Omega;L^1(0,T;H))} + \|\Phi_{k+1}\|_{L^p(\Omega;L^2(0,T;\mathscr{L}^2(K,H)))} + \|\Theta_{k+1}\|_{\mathbb{G}^p}, \tag{6.4}$$

as claimed. Let us consider the second term on the right-hand side of the previous inequality (the first one can be handled in a completely similar way). As already seen, the generic term in $\Phi_{k+1}(h_1, \dots, h_{k+1})$ is of the form

$$D^j B(u)\big(u^{(\alpha_1)}(h_{\sigma(1)}, \dots, h_{\sigma(\alpha_1)}), \dots, u^{(\alpha_j)}(h_{\sigma(\beta+1)}, \dots, h_{\sigma(k)})\big),$$

where $j \in \{2, \dots, k+1\}$, $\alpha_1 + \cdots + \alpha_j = k+1$, $\beta := \alpha_1 + \cdots + \alpha_{j-1}$, and σ is an element of the permutation group of $\{1, \dots, k+1\}$. Since $j \geq 2$ implies

$$1 + (\alpha_1 - 1) + \cdots + (\alpha_j - 1) = \alpha_1 + \cdots + \alpha_j + 1 - j = k + 2 - j \leq k,$$

one has

$$\begin{aligned}
&\frac{1}{p_0} + \frac{\alpha_1 - 1}{p_0} + \frac{1}{p_{\sigma(1)}} + \cdots + \frac{1}{p_{\sigma(\alpha_1)}} \\
&\quad + \frac{\alpha_2 - 1}{p_0} + \frac{1}{p_{\sigma(\alpha_1+1)}} + \cdots + \frac{1}{p_{\sigma(\alpha_1+\alpha_2)}} \\
&\quad \vdots \\
&\quad + \frac{\alpha_j - 1}{p_0} + \frac{1}{p_{\sigma(\alpha_1+\cdots+\alpha_{j-1}+1)}} + \cdots + \frac{1}{p_{\sigma(k+1)}} \\
&\qquad \leq \frac{n}{p_0} + \frac{1}{p_1} + \cdots + \frac{1}{p_{k+1}} = \frac{1}{p},
\end{aligned}$$

so that setting

$$\begin{aligned}
\frac{1}{\tilde{p}_1} &:= \frac{\alpha_1 - 1}{p_0} + \frac{1}{p_{\sigma(1)}} + \cdots + \frac{1}{p_{\sigma(\alpha_1)}}, \\
\frac{1}{\tilde{p}_2} &:= \frac{\alpha_2 - 1}{p_0} + \frac{1}{p_{\sigma(\alpha_1+1)}} + \cdots + \frac{1}{p_{\sigma(\alpha_1+\alpha_2)}}, \\
&\vdots \\
\frac{1}{\tilde{p}_j} &:= \frac{\alpha_j - 1}{p_0} + \frac{1}{p_{\sigma(\alpha_1+\cdots+\alpha_{j-1}+1)}} + \cdots + \frac{1}{p_{\sigma(k+1)}},
\end{aligned}$$

it holds

$$\frac{1}{p_0} + \frac{1}{\tilde{p}_1} + \cdots + \frac{1}{\tilde{p}_j} \leq \frac{1}{p}.$$

Assumption (F^n) now implies

$$\begin{aligned}
&\|D^j B(u)\big(u^{(\alpha_1)}(h_{\sigma(1)}, \dots, h_{\sigma(\alpha_1)}), \dots, u^{(\alpha_j)}(h_{\sigma(\beta+1)}, \dots, h_{\sigma(k)})\big)\|_{\mathscr{L}^2(K,H)} \\
&\quad \lesssim \big(1 + \|u\|^m\big)\|u^{(\alpha_1)}(h_{\sigma(1)}, \dots, h_{\sigma(\alpha_1)})\|_{\mathscr{L}^2(K,H)} \cdots \\
&\qquad \cdots \|u^{(\alpha_j)}(h_{\sigma(\beta+1)}, \dots, h_{\sigma(k)})\|_{\mathscr{L}^2(K,H)},
\end{aligned}$$

which yields, thanks to the estimate $\|\cdot\|_{L^2(0,T)} \le T^{1/2}\|\cdot\|_{L^\infty(0,T)}$,

$$\begin{aligned}&\big\|D^jB(u)\big(u^{(\alpha_1)}(h_{\sigma(1)},\dots,h_{\sigma(\alpha_1)}),\dots,u^{(\alpha_j)}(h_{\sigma(\beta+1)},\dots,h_{\sigma(k)})\big)\big\|_{L^2(0,T;\mathscr{L}^2(K,H))}\\&\quad\lesssim \big(1+u^{*m}\big)\big(u^{(\alpha_1)}(h_{\sigma(1)},\dots,h_{\sigma(\alpha_1)})\big)^*\ \cdots\ \big(u^{(\alpha_j)}(h_{\sigma(\beta+1)},\dots,h_{\sigma(k)})\big)^*,\end{aligned}$$

where the implicit constant depends also on T. Here and in the following we write, for simplicity of notation, $\phi^* := \phi_T^*$ for any càdlàg function ϕ. Hölder's inequality yields

$$\begin{aligned}&\big\|D^jB(u)\big(u^{(\alpha_1)}(h_{\sigma(1)},\dots,h_{\sigma(\alpha_1)}),\dots,u^{(\alpha_j)}(h_{\sigma(\beta+1)},\dots,h_{\sigma(k)})\big)\big\|_{L^p(\Omega;L^2(0,T;\mathscr{L}^2(K,H)))}\\&\quad\lesssim \big(1+\big\|u^{*m}\big\|_{\mathbb{L}^{p_0}}\big)\big\|u^{(\alpha_1)}(h_{\sigma(1)},\dots,h_{\sigma(\alpha_1)})\big\|_{\mathbb{S}^{\tilde p_1}}\ \cdots\ \big\|u^{(\alpha_j)}(h_{\sigma(\beta+1)},\dots,h_{\sigma(k+1)})\big\|_{\mathbb{S}^{\tilde p_j}},\end{aligned}$$

where, as before, $\beta := \alpha_1+\cdots+\alpha_{j-1}$. It follows by the definition of $\tilde p_1,\dots,\tilde p_j$ and the inductive assumption that

$$\begin{aligned}u^{(\alpha_1)} &\in \mathscr{L}_{\alpha_1}\big(\mathbb{L}^{p_{\sigma(1)}},\dots,\mathbb{L}^{p_{\sigma(\alpha_1)}};\mathbb{S}^{\tilde p_1}\big),\\&\ \ \vdots\\u^{(\alpha_j)} &\in \mathscr{L}_{\alpha_j}\big(\mathbb{L}^{p_{\sigma(\beta+1)}},\dots,\mathbb{L}^{p_{\sigma(k+1)}};\mathbb{S}^{\tilde p_j}\big),\end{aligned}$$

hence, recalling that $\big\|u^{*m}\big\|_{\mathbb{L}^{p0}} = \big\|u\big\|^m_{\mathbb{S}^{mp0}} \lesssim 1+\big\|u_0\big\|^m_{\mathbb{L}^{mp0}}$ by Theorem 3.2,

$$\begin{aligned}&\big\|D^jB(u)\big(u^{(\alpha_1)}(h_{\sigma(1)},\dots,h_{\sigma(\alpha_1)}),\dots,u^{(\alpha_j)}(h_{\sigma(\beta+1)},\dots,h_{\sigma(k)})\big)\big\|_{L^p(\Omega;L^2(0,T;\mathscr{L}^2(K,H)))}\\&\quad\lesssim \big(1+\big\|u_0\big\|^m_{\mathbb{L}^{mp_0}}\big)\big\|h_{\sigma(1)}\big\|_{\mathbb{L}^{p_{\sigma(1)}}}\ \cdots\ \big\|h_{\sigma(n+1)}\big\|_{\mathbb{L}^{p_{\sigma(k+1)}}}\\&\quad\lesssim \big(1+\big\|u_0\big\|^m_{\mathbb{L}^{mp_0}}\big)\big\|h_1\big\|_{\mathbb{L}^{p_1}}\ \cdots\ \big\|h_{k+1}\big\|_{\mathbb{L}^{p_{k+1}}}.\end{aligned}$$

Estimating the $\mathbb{G}^p$ norm of Θ_{k+1} is similar: using the same notation used thus far, the generic term in $\Theta_{k+1}(h_1,\dots,h_{k+1})$ is of the type

$$D^jG(u_-)\big(u_-^{(\alpha_1)}(h_{\sigma(1)},\dots,h_{\sigma(\alpha_1)}),\dots,u_-^{(\alpha_j)}(h_{\sigma(\beta+1)},\dots,h_{\sigma(k)})\big),$$

and hypothesis (F^n) implies

$$\begin{aligned}&\big\|D^jG_i(u_-)\big(u_-^{(\alpha_1)}(h_{\sigma(1)},\dots,h_{\sigma(\alpha_1)}),\dots,u_-^{(\alpha_j)}(h_{\sigma(\beta+1)},\dots,h_{\sigma(k)})\big)\big\|\\&\quad\lesssim g_i\big(1+\big\|u\big\|^m\big)\big\|u^{(\alpha_1)}(h_{\sigma(1)},\dots,h_{\sigma(\alpha_1)})\big\|\ \cdots\\&\qquad\cdots\ \big\|u^{(\alpha_j)}(h_{\sigma(\beta+1)},\dots,h_{\sigma(k)})\big\|\end{aligned}$$

for all $(t,z)\in[0,T]\times Z$, $\mathbb{P}$-a.s., for both $i=1$ and $i=2$ (we can identify again g_1 and g_2 with g depending on the value of p, and similarly for G_1 and G_2). This yields, after standard computations already detailed more than once,

$$\begin{aligned}&\big\|D^jG(u_-)\big(u_-^{(\alpha_1)}(h_{\sigma(1)},\dots,h_{\sigma(\alpha_1)}),\dots,u_-^{(\alpha_j)}(h_{\sigma(\beta+1)},\dots,h_{\sigma(k)})\big)\big\|_{\mathbb{G}^p}\\&\quad\lesssim \kappa(T)\big(1+\big\|u^{*m}\big\|_{\mathbb{L}^{p_0}}\big)\big\|u^{(\alpha_1)}(h_{\sigma(1)},\dots,h_{\sigma(\alpha_1)})\big\|_{\mathbb{S}^{\tilde p_1}}\ \cdots\\&\qquad\cdots\ \big\|u^{(\alpha_j)}(h_{\sigma(\beta+1)},\dots,h_{\sigma(k)})\big\|_{\mathbb{S}^{\tilde p_j}},\end{aligned}$$

It hence follows by the inductive assumption, as before, that

$$\begin{aligned}&\left\| D^j G(u)\big(u^{(\alpha_1)}(h_{\sigma(1)},\dots,h_{\sigma(\alpha_1)}),\dots,u^{(\alpha_j)}(h_{\sigma(\beta+1)},\dots,h_{\sigma(k)})\big)\right\|_{\mathbb{G}^p}\\ &\quad\lesssim \big(1+\|u_0\|^m_{\mathbb{L}^{mp_0}}\big)\|h_1\|_{\mathbb{L}^{p_1}}\ \cdots\ \|h_{k+1}\|_{\mathbb{L}^{p_{k+1}}}.\end{aligned}$$

Since $p_1,\dots,p_{k+1}$ were arbitrary, we have proved that $k/p_0+\sum_{j=1}^{k+1}1/p_j\le 1/p$ implies

$$u^{(k+1)}\in\mathscr{L}_{k+1}(\mathbb{L}^{p_1},\dots,\mathbb{L}^{p_{k+1}};\mathbb{S}^p),$$

thus completing the induction argument by arbitrariness of k. □

Remark 6.3. If $p_1=\cdots=p_n=q$, condition (6.3) becomes

$$\frac{n-1}{p_0}+\frac{n}{q}\le\frac{1}{p},$$

which implies $q\ge np$ and $p_0\ge(n-1)p$, hence $p_0\ge p$ if $n\ge 2$. In particular, if $q=np$, then $p_0=+\infty$, i.e., u_0 must be bounded almost surely. If $q>np$, then p_0 will also be finite, and strictly larger than p if $n\ge 2$. Furthermore, if $q>(n+nm-m)p$, then $u_0\in\mathbb{L}^q$ implies $u^{(n)}\in\mathscr{L}_n(\mathbb{L}^q;\mathbb{S}^p)$. In fact, for this to be true it suffices that $\mathbb{L}^q\subseteq\mathbb{L}^{mp_0\vee p}$, which is equivalent to $q\ge mp_0\vee p$. But since $q\ge np\ge p$, we can simply choose $q=mp_0$, which yields, excluding the case $p_0=+\infty$,

$$\frac{(n-1)m}{q}+\frac{n}{q}<\frac{1}{p},$$

or, equivalently, $q>(n+nm-m)p$.

We repeat, however, that even under these conditions we cannot yet claim that $u^{(n)}$ identifies the n-th Fréchet derivative of u. In fact, we shall prove that $D^n u$ satisfies the equation for $u^{(n)}$ when "tested" on $(\mathbb{L}^q)^n$, with q satisfying a strictly stronger constraint than just $q>(n+mn-m)p$.

Before considering Fréchet differentiability of n-th order, we need some preparations. The following two lemmata are used to apply the theorem on the Fréchet differentiability of the composition of two Fréchet-differentiable functions.

By the assumptions (A2), (A3) and (A4), it follows immediately that the superposition operators associated with f, B and G on $\mathbb{S}^p$, i.e., $\phi\mapsto f(\phi)$, $B(\phi)$, $G(\phi_-)$, can be considered as maps, denoted by the same symbols for simplicity,

$$\begin{aligned}&f:\mathbb{S}^p\longrightarrow\mathbb{S}^p,\\ &B:\mathbb{S}^p\longrightarrow L^p(\Omega;L^2(0,T;\mathscr{L}^2(K,H))),\\ &G:\mathbb{S}^p\longrightarrow\mathbb{G}^p.\end{aligned}$$

Lemma 6.4. *Let $p\ge 1$, $r>0$, $q\ge 1$, and $n\in\mathbb{N}$ satisfy*

$$\frac{1}{r}+\frac{n}{q}<\frac{1}{p},\qquad\frac{n+m}{q}<\frac{1}{p}.$$

If hypothesis (F^n) *is satisfied, then* f, B *and* G *are* n *times Fréchet-differentiable in* $\mathbb{S}^{mr} \cap \mathbb{S}^p = \mathbb{S}^{mr \vee p}$ *along* $\mathbb{S}^q$, *with*

$$\begin{aligned} D^j f &: \mathbb{S}^{mr\vee p} \longrightarrow \mathscr{L}_j(\mathbb{S}^q; \mathbb{S}^p), \\ D^j B &: \mathbb{S}^{mr\vee p} \longrightarrow \mathscr{L}_j\big(\mathbb{S}^q; L^p(\Omega; L^2(0,T; \mathscr{L}^2(K,H)))\big), \\ D^j G &: \mathbb{S}^{mr\vee p} \longrightarrow \mathscr{L}_j(\mathbb{S}^q; \mathbb{G}^p) \end{aligned}$$

for all $j \in \{1, \dots, n\}$.

Proof. We proceed by induction on j, and we treat only the third term, as all other cases are analogous (in fact slightly simpler). If $j = 1$, the proof is exactly the same as the corresponding one of Theorem 5.1. In particular, one has

$$DG(\omega, t, z, \cdot) : H \to \mathscr{L}(H, H) \qquad \forall (\omega, t, z) \in \Omega \times [0, T] \times Z,$$

hence, given $v \in \mathbb{S}^p$ and $w \in \mathbb{S}^q$ with $q > 1 \cdot p = p$,

$$\big\| \varepsilon^{-1}\big(G(v_- + \varepsilon w_-) - G(v_-)\big) - DG(v_-)w_-\big\|_{\mathbb{G}^p} \longrightarrow 0$$

as $\varepsilon \to 0$, uniformly over w belonging to bounded subsets of $\mathbb{S}^q$. Assuming now that the statement is true for $j \in \{1, \dots, n-1\}$, let us show that it also holds for $j+1$. By the inductive hypothesis we thus have

$$D^j G : \mathbb{S}^{mr} \cap \mathbb{S}^p \to \mathscr{L}_j(\mathbb{S}^q; \mathbb{G}^p).$$

Let $u \in \mathbb{S}^p$ and $v_1, \dots, v_{j+1} \in \mathbb{S}^q$. The $(j+1)$-th Fréchet derivatives

$$D^{j+1} G(\omega, t, z, \cdot),\ D^{j+1} G_i(\omega, t, z, \cdot) : H \to \mathscr{L}_{j+1}(H; H), \qquad i = 1, 2,$$

exists for all $(\omega, t, z) \in \Omega \times [0, T] \times Z$, hence, setting $\mathbf{v}_k := (v_1, \dots, v_k)$, $k = 1, \dots, n$, one has, as $\varepsilon \to 0$,

$$\left\| \frac{1}{\varepsilon}\big(D^j G(u_- + \varepsilon (v_{j+1})_-)\mathbf{v}_{j-} - D^j G(u_-)\mathbf{v}_{j-}\big) - D^{j+1} G(u_-)(\mathbf{v}_{j+1})_- \right\| \longrightarrow 0 \tag{6.5}$$

for all $(\omega, t) \in [0, T] \times Z$, $\mathbb{P}$-a.s., uniformly with respect to v_{j+1} in bounded sets of H. For any $h \in H$, the fundamental theorem of calculus yields

$$\begin{aligned} &\big\langle D^j G(u_- + \varepsilon (v_{j+1})_-)\mathbf{v}_{j-} - D^j G(u_-)\mathbf{v}_{j-}, h\big\rangle \\ &\quad = \int_0^\varepsilon \big\langle D^{j+1} G(u_- + s(v_{j+1})_-)(\mathbf{v}_{j+1})_-, h\big\rangle \, \mathrm{d}s, \end{aligned}$$

hence, since h is arbitrary,

$$\begin{aligned}
&\left\| \frac{1}{\varepsilon}\big(D^j G(u_- + \varepsilon(v_{j+1})_- \mathbf{v}_{j-} - D^j G(u_-)\mathbf{v}_{j-}\big) \right\| \\
&\quad = \left\| \frac{1}{\varepsilon} \int_0^\varepsilon D^{j+1} G(u_- + s(v_{j+1})_-)(\mathbf{v}_{j+1})_- \, \mathrm{d}s \right\| \\
&\quad \lesssim (g_1 + g_2)\big(1 + \|u_-\|^m + \|(v_{j+1})_-\|^m\big)\|v_{1-}\| \cdots \|v_{j-}\| \, \|(v_{j+1})_-\| \\
&\quad \le (g_1 + g_2)\bigg(\prod_{k=1}^{j+1} v_k^* + u^{*m} \prod_{k=1}^{j+1} v_k^* + v_{j+1}^{*(m+1)} \prod_{k=1}^{j} v_k^*\bigg)
\end{aligned}$$

for any $|\varepsilon| \le 1$, where, as already done before, $g_1 := g_2 := g/2$ if $p \ge 2$. The left-hand side of (6.5) is thus dominated for all $(t, z) \in [0, T] \times Z$, $\mathbb{P}$-a.s., modulo a constant, by the same expression appearing on the right-hand side of the previous inequality. This implies

$$\begin{aligned}
&\left\| \frac{1}{\varepsilon}\big(D^j G(u_- + \varepsilon(v_{j+1})_-)\mathbf{v}_{j-} - D^j G(u_-)\mathbf{v}_{j-}\big) - D^{j+1} G(u_-)(\mathbf{v}_{j+1})_- \right\|_{\mathbb{G}^p} \\
&\quad \lesssim \kappa(T)\bigg(\left\| \prod_{k=1}^{j+1} v_k^* \right\|_{L^p(\Omega)} + \left\| u^{*m} \prod_{k=1}^{j+1} v_k^* \right\|_{L^p(\Omega)} + \left\| v_{j+1}^{*(m+1)} \prod_{k=1}^{j} v_k^* \right\|_{L^p(\Omega)} \bigg)
\end{aligned}$$

where, by Hölder's inequality,

$$\begin{aligned}
\left\| \prod_{k=1}^{j+1} v_k^* \right\|_{L^p(\Omega)} &\le \prod_{k=1}^{j+1} \|v_k\|_{\mathbb{S}^q} < \infty, \\
\left\| u^{*m} \prod_{k=1}^{j+1} v_k^* \right\|_{L^p(\Omega)} &\le \|u\|_{\mathbb{S}^{mr}}^m \prod_{k=1}^{j+1} \|v_k\|_{\mathbb{S}^q} < \infty, \\
\left\| v_{j+1}^{*(m+1)} \prod_{k=1}^{j} v_k^* \right\|_{L^p(\Omega)} &\le \|v_{j+1}\|_{\mathbb{S}^q}^{(m+1)} \prod_{k=1}^{j} \|v_k\|_{\mathbb{S}^q} < \infty.
\end{aligned}$$

In fact, these three inequalities follow from

$$\frac{j+1}{q} < \frac{1}{p}, \qquad \frac{1}{r} + \frac{j+1}{q} < \frac{1}{p}, \qquad \frac{m+1}{q} + \frac{j}{q} < \frac{1}{p},$$

respectively, all of which are immediate consequences of the assumptions. The dominated convergence theorem thus yields

$$\left\| \frac{1}{\varepsilon}\big(D^j G(u_- + \varepsilon(v_{j+1})_-)\mathbf{v}_{j-} - D^j G(u_-)\mathbf{v}_{j-}\big) - D^{j+1} G(u_-)(\mathbf{v}_{j+1})_- \right\|_{\mathbb{G}^p} \longrightarrow 0$$

as $\varepsilon \to 0$.

It remains to show that the convergence is uniform with respect to $v_1, \ldots, v_{j+1}$ bounded in $\mathbb{S}^q$. To this end, we proceed as in the case $j = 1$: for every measurable

$E \in \mathscr{F}$, the computations just carried out yield

$$\begin{aligned}\Big\| 1_E \Big(\frac{1}{\varepsilon} \big(D^k G(u_- + \varepsilon (v_{k+1})_-) \mathbf{v}_{k-} - D^k G(u_-) \mathbf{v}_{k-} \big) - D^{k+1} G(u_-)(\mathbf{v}_{k+1})_- \Big) \Big\|_{\mathbb{G}^p} \\ \lesssim \mathbb{E}\, 1_E \big(v_1^* \cdots v_{k+1}^* \big)^p + \mathbb{E}\, 1_E \big(u^{*m} v_1^* \cdots v_{k+1}^* \big)^p \\ + \mathbb{E}\, 1_E \big(v_1^* \cdots v_k^* v_{k+1}^{*(m+1)} \big)^p,\end{aligned}$$

where the implicit constant depends on $\kappa(T)$. Since $v_1, \dots, v_{k+1}$ are bounded in $\mathbb{S}^q$ and $k+1 \le n$, the product $v_1^* \cdots v_{k+1}^*$ is bounded in $\mathbb{L}^{q/n}$. Therefore, as $q/n > p$ by assumption, it follows that $\big(v_1^* \cdots v_{k+1}^* \big)^p$ is uniformly integrable. Similarly, defining s by

$$\frac{1}{s} := \frac{1}{r} + \frac{n}{q} < \frac{1}{p},$$

Hölder's inequality yields

$$\big\| u^{*m} v_1^* \cdots v_{j+1}^* \big\|_{L^s(\Omega)} \le \|u\|_{\mathbb{S}^{mr}}^m \|v_1\|_{\mathbb{S}^q} \cdots \|v_{j+1}\|_{\mathbb{S}^q},$$

where the right-hand side is finite by assumption. Since $s > p$, $\big(u^{*m} v_1^* \cdots v_{k+1}^* \big)^p$ is uniformly integrable. Finally, defining ℓ by

$$\frac{1}{\ell} := \frac{n-1}{q} + \frac{m+1}{q} = \frac{m+n}{q} < \frac{1}{p},$$

Hölder's inequality yields, recalling that $j \le n-1$,

$$\begin{aligned}\big\| v_1^* \cdots v_j^* v_{j+1}^{*(m+1)} \big\|_{L^\ell(\Omega)} &\le \|v_1\|_{\mathbb{S}^q} \cdots \|v_j\|_{\mathbb{S}^q} \big\| v_{j+1}^{*(m+1)} \big\|_{L^{q/(m+1)}(\Omega)} \\ &= \|v_1\|_{\mathbb{S}^q} \cdots \|v_j\|_{\mathbb{S}^q} \|v_{j+1}\|_{\mathbb{S}^q}^{m+1} < \infty,\end{aligned}$$

hence $\big(v_1^* \cdots v_j^* v_{j+1}^{*(m+1)} \big)^p$ is also uniformly integrable. One can now choose the set E and proceed exactly as in the proof of Theorem 5.1 for the case $j = 1$ to conclude. □

The previous lemma implies, in particular, that

$$\begin{aligned} f &: \mathbb{S}^q \longrightarrow \mathbb{S}^p, \\ B &: \mathbb{S}^q \longrightarrow L^p(\Omega; L^2(0,T; \mathscr{L}^2(K, H))), \\ G &: \mathbb{S}^q \longrightarrow \mathbb{G}^p \end{aligned}$$

are n times Fréchet-differentiable for every $q > (m+n)p$. Indeed, for any such q, one has $\frac{m}{q} + \frac{n}{q} = \frac{m+n}{q} < \frac{1}{p}$, implying in particular that $\frac{1}{p} - \frac{n}{q} \in (0,1)$. Setting now $\frac{1}{r} := (\frac{m}{q}) \vee \frac{1}{2}(\frac{1}{p} - \frac{n}{q})$, one has $r > 1$, $\frac{1}{r} + \frac{n}{q} < \frac{1}{p}$, and $\mathbb{S}^q \subseteq \mathbb{S}^{p \vee mr}$.

In fact, if $\frac{m}{q} > \frac{1}{2}(\frac{1}{p} - \frac{n}{q})$ one has $\frac{1}{r} = \frac{m}{q}$, from which $\frac{1}{r} + \frac{n}{q} = \frac{m}{q} + \frac{n}{q} < \frac{1}{p}$ and $q = mr > p$, hence $\mathbb{S}^q \subset \mathbb{S}^{p \vee mr}$. If $\frac{m}{q} \le \frac{1}{2}(\frac{1}{p} - \frac{n}{q})$, one has $\frac{1}{r} = \frac{1}{2}(\frac{1}{p} - \frac{n}{q}) < \frac{1}{p} - \frac{n}{q}$, from which $\frac{1}{r} + \frac{n}{q} < \frac{1}{p}$, and $q \ge mr$, hence $\mathbb{S}^q \subseteq \mathbb{S}^{p \vee mr}$. The assertion follows then from Lemma 6.4.

We can now state the main result of this section, as well as of the whole paper.

Theorem 6.5. *Let $n \geq 1$,*

$$q > \frac{(m+n)!}{(m+1)!}p.$$

Assume (F^n) *and* $(\mathrm{A5}_r)$ *for all* $r \in [p, q]$. *Then, the solution map* $(u_0 \mapsto u) : \mathbb{L}^q \to \mathbb{S}^p$ *is n times Fréchet-differentiable.*

Moreover, $D^n u(u_0) \in \mathscr{L}_n(\mathbb{L}^q; \mathbb{S}^p)$ *is the unique mild solution* $u^{(n)}$ *to*

$$du^{(n)} + Au^{(n)}\,dt = \left(\Psi_n + Df(u)u^{(n)}\right)dt + \left(\Phi_n + DB(u)u^{(n)}\right)dW + \int_Z \left(\Theta_n + DG(u_-)u_-^{(n)}\right)d\bar{\mu}.$$

Note that this equation is nothing else than (6.1) and must be interpreted as the latter, i.e., in the sense of testing against an n-tuple of vectors in $\mathbb{L}^q$. Moreover, the initial condition of the equation is the identity map if $n = 1$, and zero if $n \geq 2$.

Proof. We shall assume, for simplicity, that $f = B = 0$, as the argument in the general case $f \neq 0$, $B \neq 0$ is entirely analogous. We are going to argue by induction on $\ell \in \{1, \dots, n\}$. The statement is true for $\ell = 1$ by Theorem 5.1. Now we assume that the statement is true for all $j \leq \ell - 1$, $\ell \in \{2, \dots, n\}$, and we prove it for ℓ. Let $k \in \mathbb{L}^q$, with $q > \frac{(m+\ell)!}{(m+1)!}p = (m+\ell)\dots(m+2)p$. Thanks to Proposition 6.2 and the remarks following its proof, the equation

$$\mathrm{d}u^{(\ell)} + Au^{(\ell)}\,\mathrm{d}t = \int_Z \left(\Theta_\ell + DG(u_-)u_-^{(\ell)}\right)\mathrm{d}\bar{\mu}, \qquad u^{(\ell)}(0) = 0,$$

admits a unique mild solution $u^{(\ell)} \in \mathscr{L}_\ell(\mathbb{L}^q; \mathbb{S}^p)$, because

$$q > (m+\ell)\cdots(m+2)p \geq (\ell + m\ell - m)p.$$

We are going to show that $u^{(\ell)} = D^\ell u(u_0)$ in $\mathscr{L}_\ell(\mathbb{L}^q; \mathbb{S}^p)$. Let $k \in \mathbb{L}^q$: for brevity, we shall use the notation $u^{(\ell)}(u_0)k := u^{(\ell)}(u_0)(k, \cdot, \dots, \cdot) \in \mathscr{L}_{\ell-1}(\mathbb{L}^p, \mathbb{S}^p)$ and $\Theta_\ell(u_0)k := \Theta_\ell(u_0)(k, \cdot, \dots, \cdot) \in \mathscr{L}_{\ell-1}(\mathbb{L}^p, \mathbb{S}^p)$. One has

$$\begin{aligned}
&\frac{u^{(\ell-1)}(u_0 + \varepsilon k) - u^{(\ell-1)}(u_0)}{\varepsilon} - u^{(\ell)}(u_0)k \\
&\quad = S \diamond_\mu \left(\frac{\Theta_{\ell-1}(u_0 + \varepsilon k) - \Theta_{\ell-1}(u_0)}{\varepsilon} - \Theta_\ell(u_0)k\right) \\
&\qquad + S \diamond_\mu \left(\frac{DG(u_-(u_0 + \varepsilon k))u_-^{(\ell-1)}(u_0 + \varepsilon k) - DG(u_-(u_0))u_-^{(\ell-1)}(u_0)}{\varepsilon}\right. \\
&\qquad \left. - DG(u_-(u_0))u_-^{(\ell)}(u_0)k\right),
\end{aligned} \tag{6.6}$$

where, by the inductive hypothesis, $u^{(\ell-1)}(u_0) = D^{\ell-1}u(u_0)$ and $u^{(\ell-1)}(u_0 + \varepsilon k) = D^{\ell-1}u(u_0 + \varepsilon k)$ in $\mathscr{L}_{\ell-1}(\mathbb{L}^q; \mathbb{S}^p)$. We need to prove that the left-hand side of (6.6)

converges to zero as $\varepsilon \to 0$ in $\mathscr{L}_{\ell-1}(\mathbb{L}^q, \mathbb{S}^p)$ uniformly over k belonging to bounded sets of $\mathbb{L}^q$. Thanks to (6.2), one has

$$\begin{aligned}&\frac{\Theta_{\ell-1}(u_0+\varepsilon k)-\Theta_{\ell-1}(u_0)}{\varepsilon}-\Theta_\ell(u_0)k\\&\quad=\frac{\Theta_{\ell-1}(u_0+\varepsilon k)-\Theta_{\ell-1}(u_0)}{\varepsilon}-D\Theta_{\ell-1}(u_0)k-D^2G(u_-)(u'_-,u_-^{(\ell-1)})k,\end{aligned}$$

and we claim that

$$\frac{\Theta_{\ell-1}(u_0+\varepsilon k)-\Theta_{\ell-1}(u_0)}{\varepsilon}-D\Theta_{\ell-1}(u_0)k\to 0 \tag{6.7}$$

in $\mathscr{L}_{\ell-1}(\mathbb{L}^q;\mathbb{G}^p)$ as $\varepsilon\to 0$, uniformly over k belonging to bounded subsets of $\mathbb{L}^q$. In fact, all terms in $\Theta_{\ell-1}$ are of the form

$$D^jG(u_-(u_0))\big(D^{\alpha_1}u_-(u_0),\dots,D^{\alpha_j}u_-(u_0)\big), \tag{6.8}$$

with $j\le\ell-1$ and $\alpha_i\ge 1$, $\sum\alpha_i=\ell-1$. Now, let $r>(\ell+m)p$ be such that $(u_0\mapsto u):\mathbb{L}^r\to\mathbb{S}^r$ (which is possible because $q>(\ell+m)p$). Then, $G:\mathbb{S}^r\to\mathbb{G}^p$ is n times Fréchet-differentiable by Lemma 6.4. Moreover, by the inductive hypothesis applied to $(u_0\mapsto u)\in(\mathbb{L}^r\to\mathbb{S}^r)$, we have that $u_0\mapsto u$ is $\ell-1$ times Fréchet-differentiable along $\mathbb{L}^q$ if

$$q>\frac{(m+\ell-1)!}{(m+1)!}r>\frac{(m+\ell-1)!}{(m+1)!}(m+\ell)p=\frac{(m+\ell)!}{(m+1)!}p.$$

Therefore, if q satisfies this condition, each term of the form (6.8) is Fréchet-differentiable along $\mathbb{L}^q$ by the theorem on the Fréchet differentiability of composite functions (see, for example, [1, Prop. 1.4]). Hence, (6.7) is indeed true, and the expression within parentheses in the first term on the right-hand side of (6.6) converges to

$$-D^2G(u_-(u_0))\big(Du_-(u_0),D^{\ell-1}u_-(u_0)\big)k$$

in $\mathscr{L}_{\ell-1}(\mathbb{L}^q;\mathbb{G}^p)$ as $\varepsilon\to 0$, uniformly over k belonging to bounded subsets of $\mathbb{L}^q$. Let us now consider the second term on the right-hand side of (6.6). One has, recalling that $v_j=D^ju$ for every $j\le\ell-1$ by inductive hypothesis,

$$\begin{aligned}&\frac{DG(u_-(u_0+\varepsilon k))u_-^{(\ell-1)}(u_0+\varepsilon k)-DG(u_-(u_0))u_-^{(\ell-1)}(u_0)}{\varepsilon}\\&\quad-DG(u_-(u_0))u_-^{(\ell)}(u_0)k\\&=DG(u_-(u_0))\left(\frac{D^{\ell-1}u_-(u_0+\varepsilon k)-D^{\ell-1}u_-(u_0)}{\varepsilon}-u_-^{(\ell)}(u_0)k\right)\\&\quad+\frac{DG(u_-(u_0+\varepsilon k))-DG(u_-(u_0))}{\varepsilon}D^{\ell-1}u_-(u_0+\varepsilon k),\end{aligned}$$

where the second term on the right-hand side converges to

$$D^2G(u_-(u_0))\big(Du_-(u_0), D^{\ell-1}u_-(u_0)\big)k$$

in $\mathscr{L}_{\ell-1}(\mathbb{L}^q;\mathbb{G}^p)$ as $\varepsilon\to 0$, uniformly over k bounded in $\mathbb{L}^q$, because again everything depends only on derivatives of order at most $\ell-1$ and we can apply the usual criteria on Fréchet differentiability of multilinear maps and composite functions. Note that this term cancels out with the corresponding one obtained previously. Going back then to (6.6), testing by an arbitrary element $(k_2,\dots,k_\ell)\in(\mathbb{L}^q)^{\ell-1}$, and using Lemma 2.4, we infer that

$$\begin{aligned}
&\left\|\left(\frac{u^{(\ell-1)}(u_0+\varepsilon k)(k_2,\dots,k_\ell)-u^{(\ell-1)}(u_0)(k_2,\dots,k_\ell)}{\varepsilon}-u^{(\ell)}(u_0)(k,k_2,\dots,k_\ell)\right)\right\|_{\mathbb{S}^p}\\
&\lesssim\left\|\left(\frac{\Theta_{\ell-1}(u_0+\varepsilon k)(k_2,\dots,k_\ell)-\Theta_{\ell-1}(u_0)(k_2,\dots,k_\ell)}{\varepsilon}-\Theta_\ell(u_0)(k,k_2,\dots,k_\ell)\right)\right.\\
&\quad\left.+\frac{DG(u_-(u_0+\varepsilon k))-DG(u_-(u_0))}{\varepsilon}D^{\ell-1}u_-(u_0+\varepsilon k)(k_2,\dots,k_\ell)\right\|_{\mathbb{G}^p}\\
&\quad+\left\|DG(u_-(u_0))\left(\frac{D^{\ell-1}u_-(u_0+\varepsilon k)-D^{\ell-1}u_-(u_0)}{\varepsilon}(k_2,\dots,k_\ell)-u_-^{(\ell)}(u_0)(k,k_2,\dots,k_\ell)\right)\right\|_{\mathbb{G}^p}.
\end{aligned}$$

Taking supremum over $(k_2,\dots,k_\ell)$ bounded in $(\mathbb{L}^q)^{\ell-1}$ and using the Lipschitz continuity of G, we infer that, for every $T_0\in(0,T]$,

$$\begin{aligned}
&\left\|\left(\frac{u^{(\ell-1)}(u_0+\varepsilon k)-u^{(\ell-1)}(u_0)}{\varepsilon}-u^{(\ell)}(u_0)k\right)\right\|_{\mathscr{L}_{\ell-1}(\mathbb{L}^q;\mathbb{S}^p(0,T_0))}\\
&\lesssim\left\|\left(\frac{\Theta_{\ell-1}(u_0+\varepsilon k)-\Theta_{\ell-1}(u_0)}{\varepsilon}-\Theta_\ell(u_0)k\right)\right.\\
&\quad\left.+\frac{DG(u_-(u_0+\varepsilon k))-DG(u_-(u_0))}{\varepsilon}D^{\ell-1}u_-(u_0+\varepsilon k)\right\|_{\mathscr{L}_{\ell-1}(\mathbb{L}^q;\mathbb{G}^p(0,T_0))}\\
&\quad+\kappa(T_0)\left\|\frac{D^{\ell-1}u_-(u_0+\varepsilon k)-D^{\ell-1}u_-(u_0)}{\varepsilon}-u_-^{(\ell)}(u_0)k\right\|_{\mathscr{L}_{\ell-1}(\mathbb{L}^q;\mathbb{G}^p(0,T_0))}.
\end{aligned}$$

By the continuity of κ we can choose T_0 sufficiently small such that, after rearranging the terms,

$$\begin{aligned}
&\left\|\left(\frac{u^{(\ell-1)}(u_0+\varepsilon k)-u^{(\ell-1)}(u_0)}{\varepsilon}-u^{(\ell)}(u_0)k\right)\right\|_{\mathscr{L}_{\ell-1}(\mathbb{L}^q;\mathbb{S}^p(0,T_0))}\\
&\lesssim\left\|\left(\frac{\Theta_{\ell-1}(u_0+\varepsilon k)-\Theta_{\ell-1}(u_0)}{\varepsilon}-\Theta_\ell(u_0)k\right)\right.\\
&\quad\left.+\frac{DG(u_-(u_0+\varepsilon k))-DG(u_-(u_0))}{\varepsilon}D^{\ell-1}u_-(u_0+\varepsilon k)\right\|_{\mathscr{L}_{\ell-1}(\mathbb{L}^q;\mathbb{G}^p(0,T_0))}.
\end{aligned}$$

Using the same argument leading to (6.4) in the proof of Proposition 6.2, a classical patching argument yields then

$$\begin{aligned}
&\left\| \left(\frac{u^{(\ell-1)}(u_0+\varepsilon k) - u^{(\ell-1)}(u_0)}{\varepsilon} - u^{(\ell)}(u_0)k \right) \right\|_{\mathscr{L}_{\ell-1}(\mathbb{L}^q;\mathbb{S}^p)} \\
&\lesssim \left\| \left(\frac{\Theta_{\ell-1}(u_0+\varepsilon k) - \Theta_{\ell-1}(u_0)}{\varepsilon} - \Theta_\ell(u_0)k \right) \right. \\
&\quad \left. + \frac{DG(u_-(u_0+\varepsilon k)) - DG(u_-(u_0))}{\varepsilon} D^{\ell-1}u_-(u_0+\varepsilon k) \right\|_{\mathscr{L}_{\ell-1}(\mathbb{L}^q;\mathbb{G}^p)}
\end{aligned}$$

on the whole time interval $[0,T]$. Taking into account the remarks made above we have that

$$\begin{aligned}
&\lim_{\varepsilon\to 0^+} \sup_{\|k\|_{\mathbb{L}^q}\le 1} \left\| \left(\frac{u^{(\ell-1)}(u_0+\varepsilon k) - u^{(\ell-1)}(u_0)}{\varepsilon} - u^{(\ell)}(u_0)k \right) \right\|_{\mathscr{L}_{\ell-1}(\mathbb{L}^q;\mathbb{S}^p)} \\
&\lesssim \lim_{\varepsilon\to 0^+} \sup_{\|k\|_{\mathbb{L}^q}\le 1} \left\| \left(\frac{\Theta_{\ell-1}(u_0+\varepsilon k) - \Theta_{\ell-1}(u_0)}{\varepsilon} - \Theta_\ell(u_0)k \right) \right. \\
&\quad \left. + \frac{DG(u_-(u_0+\varepsilon k)) - DG(u_-(u_0))}{\varepsilon} D^{\ell-1}u_-(u_0+\varepsilon k) \right\|_{\mathscr{L}_{\ell-1}(\mathbb{L}^q;\mathbb{G}^p)} \\
&= \left\| -D^2G(u_-(u_0))\big(Du_-(u_0), D^{\ell-1}u_-(u_0)\big)k \right. \\
&\quad \left. + D^2G(u_-(u_0))\big(Du_-(u_0), D^{\ell-1}u_-(u_0)\big)k \right\|_{\mathscr{L}_{\ell-1}(\mathbb{L}^q;\mathbb{G}^p)} \\
&= 0.
\end{aligned}$$

We conclude that the left-hand side of converges to zero in $\mathscr{L}_{\ell-1}(\mathbb{L}^q;\mathbb{S}^p)$ as $\varepsilon\to 0$, uniformly with respect to k belonging to any bounded subset of $\mathbb{L}^q$, as required. □

Acknowledgements

Open access funding provided by University of Vienna. A large part of the work for this paper was done during several visits of the first-named author to the Interdisziplinäres Zentrum für Komplexe Systeme (IZKS) at the University of Bonn, Germany, and a visit to the University of Vienna, Austria. The warm hospitality of his hosts (S. Albeverio and U. Stefanelli, respectively) and the good working conditions are gratefully acknowledged. The second-named author is funded by Vienna Science and Technology Fund (WWTF) through Project MA14-009. The authors are indebted to G. Luise for contributing to a preliminary draft.

REFERENCES

[1] A. Ambrosetti and G. Prodi, *A primer of nonlinear analysis*, Cambridge University Press, Cambridge, 1995. (96a:58019)

[2] A. Andersson, A. Jentzen, R. Kurniawan, and T. Welti, *On the differentiability of solutions of stochastic evolution equations with respect to their initial values*, Nonlinear Anal. **162** (2017), 128–161.

[3] V. I. Averbukh and O. G. Smolyanov, *Differentiation theory in linear topological spaces*, Uspekhi Mat. Nauk **22** (1967), no. 6 (138), 201–260.

[4] V. I. Bogachev, N. V. Krylov, M. Röckner, and S. V. Shaposhnikov, *Fokker-Planck-Kolmogorov equations*, American Mathematical Society, Providence, RI, 2015.

[5] V. I. Bogachev and O. G. Smolyanov, *Topological vector spaces and their applications*, Springer, Cham, 2017.

[6] S. Cerrai, *Second order PDE's in finite and infinite dimension*, Lecture Notes in Mathematics, vol. 1762, Springer-Verlag, Berlin, 2001.

[7] G. Da Prato, *Kolmogorov equations for stochastic PDEs*, Birkhäuser Verlag, Basel, 2004.

[8] G. Da Prato, S. Kwapień, and J. Zabczyk, *Regularity of solutions of linear stochastic equations in Hilbert spaces*, Stochastics **23** (1987), no. 1, 1–23.

[9] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, second ed., Cambridge University Press, Cambridge, 2014.

[10] S. Dirksen and I. Yaroslavtsev, *L^q-valued Burkholder–Rosenthal inequalities and sharp estimates for stochastic integrals*, arXiv:1707.00109.

[11] J. Jacod and A. N. Shiryaev, *Limit theorems for stochastic processes*, second ed., Springer-Verlag, Berlin, 2003.

[12] N. V. Krylov, *Introduction to the theory of diffusion processes*, American Mathematical Society, Providence, RI, 1995.

[13] C. Marinelli, *On well-posedness of semilinear stochastic evolution equations on L_p spaces*, SIAM J. Math. Anal. **50** (2018), no. 2, 2111–2143.

[14] C. Marinelli, S. Dirksen, and I. Yaroslavtsev, *Stochastic evolution equations with jumps on L^p spaces*, in preparation.

[15] C. Marinelli, C. Prévôt, and M. Röckner, *Regular dependence on initial data for stochastic evolution equations with multiplicative Poisson noise*, J. Funct. Anal. **258** (2010), no. 2, 616–649.

[16] C. Marinelli and M. Röckner, *On maximal inequalities for purely discontinuous martingales in infinite dimensions*, Séminaire de Probabilités XLVI, Lecture Notes in Math., vol. 2123, Springer, Cham, 2014, pp. 293–315.

[17] C. Marinelli and L. Scarpa, *Well-posedness of monotone semilinear SPDEs with semimartingale noise*, arXiv:1805.07562.

[18] C. Marinelli and L. Scarpa, *Ergodicity and Kolmogorov equations for dissipative SPDEs with singular drift: a variational approach*, Potential Anal. (in press).

[19] C. Marinelli and G. Ziglio, *Ergodicity for nonlinear stochastic evolution equations with multiplicative Poisson noise*, Dyn. Partial Differ. Equ. **7** (2010), no. 1, 1–23.

[20] W. Stannat, *Stochastic partial differential equations: Kolmogorov operators and invariant measures*, Jahresber. Dtsch. Math.-Ver. **113** (2011), no. 2, 81–109.

Carlo Marinelli
Department of Mathematics
University College London
Gower Street
London WC1E 6BT
UK

Luca Scarpa
Fakultät für Mathematik
Universität Wien
Oskar-Morgenstern-Platz 1
1090 Wien
Austria
E-mail: luca.scarpa@univie.ac.at