AN INFINITE-DIMENSIONAL APPROACH TO PATH-DEPENDENT KOLMOGOROV EQUATIONS

BY FRANCO FLANDOLI AND GIOVANNI ZANCO

Università di Pisa

In this paper, a Banach space framework is introduced in order to deal with finite-dimensional path-dependent stochastic differential equations. A version of Kolmogorov backward equation is formulated and solved both in the space of L^p paths and in the space of continuous paths using the associated stochastic differential equation, thus establishing a relation between path-dependent SDEs and PDEs in analogy with the classical case. Finally, it is shown how to establish a connection between such Kolmogorov equation and the analogue finite-dimensional equation that can be formulated in terms of the path-dependent derivatives recently introduced by Dupire, Cont and Fournié.

1. Introduction. In the recent literature, a growing interest for path-dependent stochastic equations has arisen, due both to their mathematical interest and to their possible applications in finance.

The path-dependent SDEs considered here will be of the form

$$(1) \qquad \begin{cases} \mathrm{d}X(t) = b_t(X_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W(t), & \text{for } t \in [t_0, T], \\ X_{t_0} = \gamma_{t_0}, \end{cases}$$

where $\{W(t)\}_{t\geq 0}$ is a Brownian motion in $\mathbb{R}^d$, σ is a diagonalizable $d \times d$ matrix, the solution $X(t)$ at time t takes values in $\mathbb{R}^d$, the notation X_t stands for the path of the solution on the interval $[0, t]$, b_t is, for each $t \in [0, T]$, a map from a suitable space of paths to $\mathbb{R}^d$, γ_{t_0} is a given path on $[0, t_0]$.

After the insightful ideas proposed by Dupire (2009) and Cont and Fournié (2010a, 2010b, 2013), who introduced a new concept of derivative and developed a path-dependent Itô formula which exhibits a first connection between SDEs and PDEs in the path-dependent situation, some effort was made into generalizing some classical concept to this setting, like forward–backward systems and viscosity solutions [see Peng and Wang (2011), Tang and Zhang (2013), Ekren et al. (2014), Ekren, Touzi and Zhang (2016a, 2016b), Cosso (2012)]. Also, depending on the approach, there are some similarities with investigations about delay equations; see, for instance, Federico, Goldys and Gozzi (2010), Gozzi and Marinelli (2006), Fuhrman, Masiero and Tessitore (2010). Some of these works formulate a

---

path-dependent Kolmogorov equation associated to the path-dependent SDE (1). Several issues about such Kolmogorov equation are of interest. The purpose of our work is to prove *existence of classical C^2 solutions* and to develop a Banach space framework suitable for this problem. To this aim, we follow the classical method based on the probabilistic representation formula in terms of solutions to the SDE, which however, as explained in detail below, requires a new nontrivial analysis in our framework.

1.1. *Notation.* We will use the following notation throughout the paper (in addition to those introduced above): T will stand for a fixed finite time-horizon; $X_t(r)$ will stand again for the value of X at r, $r \leq t$. Stochastic processes will be denoted with upper-case letters, while Greek lower-case letters will be used for deterministic paths, most of the times seen as points in some paths space. As long as no stochastics are involved, one can always think of a path γ as defined on the whole interval $[0, T]$ and read γ_t as its restriction to $[0, t]$.

By $C([a, b]; \mathbb{R}^d)$ and $D([a, b]; \mathbb{R}^d)$ we will denote, respectively, the space of continuous and càdlàg functions from the real interval $[a, b]$ into $\mathbb{R}^d$; $D([a, b); \mathbb{R}^d)$ will denote the set of càdlàg functions that have finite left limit also for $t \to b$.

1.2. *Main results.* A *path-dependent nonanticipative function* is a family of functions $b = \{b_t\}_{t \in [0,T]}$, each one being defined on $D([0, t]; \mathbb{R}^d)$ with values in $\mathbb{R}^d$ and being measurable with respect to the canonical σ-field on $D([0, t]; \mathbb{R}^d)$. Some possible examples of path-dependent functions are the following:

(i) for $g : [0, T] \times [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ smooth, consider the function

$$b_t(\gamma_t) = \int_0^t g(t, s, \gamma(t), \gamma(s)) \, ds;$$

(ii) for $0 = t_0 \leq t_1 \leq t_2 \leq \cdots \leq t_n \leq T$ fixed consider the function

$$b_t(\gamma_t) = h_{i(t)}(\gamma(t), \gamma(t_1), \ldots, \gamma(t_{i(t)})),$$

where for each $t \in [0, T]$ the index $i(t) \in \{0, \ldots, n\}$ is such that $t_{i(t)} \leq t < t_{i(t)+1}$ and, for each $j \in \{0, \ldots, n\}$, $h_j : \mathbb{R}^{d \times (j+1)} \to \mathbb{R}^d$ is a given function with suitable properties;

(iii) for $\delta \in (0, T)$ and $q : \mathbb{R}^{2d} \to \mathbb{R}^d$ smooth, consider the function

$$b_t(\gamma_t) = q(\gamma(t), \gamma(t - \delta));$$

(iv) in dimension $d = 1$ consider the function

$$b_t(\gamma_t) = \sup_{s \in [0,t]} \gamma(s).$$

In order to formulate the path-dependent SDE (1) as an SDE in Banach spaces, we consider it as a couple (endpoint, path) in some infinite-dimensional space, as it is usually done for delay equations, and reformulate consequently equation (1) as the infinite-dimensional abstract SDE

$$\text{(2)} \quad dY(t) = AY(t)\,dt + B(t, Y(t))\,dt + \Sigma\,d\beta(t) \qquad \text{for } t \in [t_0, T], Y(t_0) = y$$

(understood in mild sense) where A is the derivative operator, B is a sufficiently smooth (in Fréchet sense) nonlinear operator with range in $\mathbb{R}^d \times \{0\}$ and β is a finite-dimensional Brownian motion (Section 2.1).

We associate to it the backward Kolmogorov equation in integral form with terminal condition Φ

$$\text{(3)} \qquad u(t, y) - \Phi(y) = \int_t^T \langle Du(s, y), Ay + B(s, y) \rangle\,ds$$

$$+ \frac{1}{2} \int_t^T \sum_{j=1}^d \sigma_j^2 D^2 u(s, y)(e_j, e_j)\,ds$$

and the related concept of solution (Section 3).

Our main result, under suitable regularity assumptions on B and Φ, as explained in Section 5 is the following (see Theorem 5.4 for the precise statement):

THEOREM. *The function*

$$u(s, y) = \mathbb{E}[\Phi(Y^{s,y}(T))],$$

where $Y^{s,y}(t)$ solves equation (2), *is of class C^2 with respect to y and solves the backward Kolmogorov equation.*

Since under our assumptions all the integrands appearing in (3) are in L^∞, a posteriori the function u is Lipschitz in t and hence, by Rademacher's theorem, differentiable almost everywhere with respect to t. Therefore, for almost every t it satisfies Kolmogorov backward equation in its differential form:

$$\begin{cases} \dfrac{\partial u}{\partial t}(t, y) + \langle Du(t, y), Ay + B(t, y) \rangle + \dfrac{1}{2} \sum_{j=1}^d \sigma_j^2 D^2 u(t, y)(e_j, e_j) = 0, \\ u(T, \cdot) = \Phi. \end{cases}$$

We moreover show (Section 6) that all usual examples satisfy the regularity requirements of the previous theorem. Finally, we provide some links between our results and the path-dependent calculus developed by Cont and Fournié (Section 7). In doing so, the main result we have (again under some regularity assumptions compatible with those of the previous theorem) is the following:

THEOREM. *The function*

$$v_s(\gamma_s) = \mathbb{E}\big[f\big(X^{\gamma_t}(T)\big)\big],$$

where $X^{\gamma_s}(t)$ is the solution to equation (1), *solves the path-dependent backward Kolmogorov equation*

(4)
$$\begin{cases} \mathscr{D}_t v(\gamma_t) + b_t(\gamma_t) \cdot \mathscr{D} v_t(\gamma_t) + \dfrac{1}{2} \sum_{j=1}^{d} \sigma_j^2 \mathscr{D}_i^2 v_t(\gamma_t) = 0, \\ v_T(\gamma_T) = f(\gamma_T), \end{cases}$$

in which the derivatives are understood as horizontal and vertical derivatives as defined by Cont and Fournié (2013).

1.3. *Some ideas about the proofs.* We intend here to find regular solutions to the Kolmogorov equation, by analogy with the classical theory. To this aim, the space of L^2 paths would appear to be the easiest setting to work in; unfortunately there are no significant example of path-dependent functions, not even integral functions, that satisfy the natural condition of having uniformly continuous second Fréchet derivative in L^2; this is discussed in detail in Section 6. To include a wider class of functions, one would want to formulate and solve equations (2) and (3) in the space of continuous paths, that in our framework is the space

$$\widehat{\mathcal{C}} := \left\{ y = \begin{pmatrix} x \\ \varphi \end{pmatrix} \in \mathbb{R}^d \times C\big([-T, 0); \mathbb{R}^d\big) \text{ s.t. } x = \lim_{s \uparrow 0} \varphi(s) \right\}.$$

This leads to two issues: first, the operator B (our abstract realization of the functional b) takes values in a space larger than $\widehat{\mathcal{C}}$, thus we have to consider paths with a single jump-discontinuity at the final time $t = 0$. But then the semigroup generated by A shifts such discontinuity so that we have to deal with paths with a single discontinuity at an arbitrary time t. The need to work with a linear space and possibly with a Banach space structure suggests the choice of

$$\mathcal{D} := \mathbb{R}^d \times D\big([-T, 0); \mathbb{R}^d\big)$$

with the uniform norm as the ambient space for our equations.

The second issue comes along when we try to establish the link between the SDE and the PDE. As in the classical theory, we need to work with some Itô-type formula. We decide not to use some version of the Itô formula in Banach spaces due to the difficulties one encounters in defining a concept of quadratic variation there [see, e.g., Di Girolami and Russo (2010, 2012, 2014), Di Girolami, Fabbri and Russo (2014)], although we intend to address this problem in our future works; we proceed therefore using a Taylor expansion, but we are not able to control the second-order terms in spaces endowed with the uniform norm.

Therefore, we adopt the following strategy: we go back to an L^p setting with $p \geq 2$ (recovering in this way at least examples like integral functionals) and we

develop rigorously the relation between the SDE and the PDE in this framework (Section 4). We then introduce an approximation procedure to extend our results to the space of continuous paths (Section 5). This step requires us to introduce an additional assumption that remarks again the deep effort that is needed in order to obtain a satisfactory general theory already in the easiest case of regular coefficients.

2. The stochastic equation.

2.1. *Framework.* From now onward, fix a time horizon $0 < T < \infty$ and a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0,T]}, \mathbb{P})$. We introduce the following spaces:

$$\mathcal{C} := \mathbb{R}^d \times \left\{ \varphi \in C_b([-T, 0); \mathbb{R}^d) : \exists \lim_{s \uparrow 0} \varphi(s) \right\},$$

$$\widehat{\mathcal{C}} := \left\{ y = \begin{pmatrix} x \\ \varphi \end{pmatrix} \in \mathcal{C} \text{ s.t. } x = \lim_{s \uparrow 0} \varphi(s) \right\},$$

$$\mathcal{D} := \mathbb{R}^d \times D_b([-T, 0); \mathbb{R}^d),$$

$$\mathcal{D}_t := \left\{ y = \begin{pmatrix} x \\ \varphi \end{pmatrix} \in \mathcal{D} \text{ s.t. } \varphi \text{ is discontinuous at most in the only point } t \right\},$$

$$\mathcal{L}^p := \mathbb{R}^d \times L^p(-T, 0; \mathbb{R}^d), \qquad p \geq 2.$$

All of them apart from $\mathcal{L}^p$ are Banach spaces with respect to the norm $\| \binom{x}{\varphi} \|^2 = |x|^2 + \|\varphi\|_\infty^2$, while $\mathcal{L}^p$ is a Banach space with respect to the norm $\| \binom{x}{\varphi} \|^2 = |x|^2 + \|\varphi\|_p^2$; the space $\mathcal{D}$ turns out to be not separable with respect to this norm but this will not undermine our method.

With these norms, we have the natural relations

$$\widehat{\mathcal{C}} \subset \mathcal{C} \subset \mathcal{D} \subset \mathcal{L}^p$$

with continuous embeddings. We remark that $\widehat{\mathcal{C}}$, $\mathcal{C}$ and $\mathcal{D}$ are dense in $\mathcal{L}^p$ while neither $\widehat{\mathcal{C}}$ nor $\mathcal{C}$ are dense in $\mathcal{D}$. The choice for the interval $[-T, 0]$ is made in accordance with most of the classical literature on delay equations.

Notice that the space $\widehat{\mathcal{C}}$ has not the structure of a product space; notice also that it is isomorphic to the space $C([-T, 0]; \mathbb{R}^d)$.

As said above, we consider a family $b = \{b_t\}_{t \in [0,T]}$ of functions

$$b_t : D([0, t]; \mathbb{R}^d) \to \mathbb{R}^d$$

adapted to the canonical filtration and we formulate the path-dependent stochastic differential equation

(5)
$$\begin{cases} dX(t) = b_t(X_t) \, dt + \sigma \, dW(t), & \text{for } t \in [t_0, T], \\ X_{t_0} = \gamma_{t_0}, \end{cases}$$

where σ is a diagonalizable $d \times d$ matrix and W is a d-dimensional Brownian motion. b can also be seen as an $\mathbb{R}^d$-valued function on the space $D = \bigcup_t D([0, t]; \mathbb{R}^d)$.

To reformulate the path-dependent SDE (5) in our framework, we need to introduce two linear bounded operators: for every $t \in [0, T]$ define the *restriction operator*

(6)
$$M_t \colon D\big([-T, 0); \mathbb{R}^d\big) \longrightarrow D\big([0, t); \mathbb{R}^d\big),$$
$$M_t(\varphi)(s) = \varphi(s - t), \qquad s \in [0, t),$$

and the *backward extension* operator

(7)
$$L_t \colon D\big([0, t); \mathbb{R}^d\big) \longrightarrow D\big([-T, 0); \mathbb{R}^d\big),$$
$$L_t(\gamma)(s) = \gamma(0)\mathbb{1}_{[-T, -t)}(s) + \gamma(t + s)\mathbb{1}_{[-t, 0)}(s), \qquad s \in [-T, 0).$$

Since the extension in the definition of L_t is arbitrary, one has that

(8)
$$M_t L_t \gamma = \gamma$$

while in general

$$L_t M_t \varphi \neq \varphi.$$

Note also that both L_t and M_t map continuous functions into continuous functions. Set moreover

(9)
$$\widetilde{M}_t \begin{pmatrix} x \\ \varphi \end{pmatrix}(s) = \begin{cases} M_t \varphi(s), & s \in [0, t), \\ x, & s = t. \end{cases}$$

Now given a functional b as in (5) one can define a function $\hat{b}$ on $[0, T] \times \mathcal{D}$ setting

(10)
$$\hat{b}\left(t, \begin{pmatrix} x \\ \varphi \end{pmatrix}\right) = \hat{b}(t, x, \varphi) := b_t\left(\widetilde{M}_t \begin{pmatrix} x \\ \varphi \end{pmatrix}\right);$$

conversely if $\hat{b}$ is given one can obtain a functional b on D setting

(11)
$$b_t(\gamma) := \hat{b}\big(t, \gamma(t), L_t \gamma\big).$$

The idea is simply to shift and extend (or restrict) the path in order to pass from one formulation to another.

For instance, the functional of example (i) in Section 1 would define a function $\hat{b}$ on $[0, T] \times \mathcal{D}$ given by

(12)
$$\hat{b}\left(t, \begin{pmatrix} x \\ \varphi \end{pmatrix}\right) = \int_0^t g(t, s, x, \varphi(s - t))\, ds.$$

We consider again the path-dependent SDE (5) with the initial condition given now by a path ψ on $[-T + t_0, t_0]$ and its terminal value $x = \psi(t_0)$,

(13)
$$\begin{cases} dX(s) = b_s(X_s)\, ds + \sigma\, dW(s), & \text{for } s \in [t_0, T], \\ X(t_0) = x = \psi(t_0), \\ X(s) = \psi(s), & \text{for } s \in [-T + t_0, t_0). \end{cases}$$

Recall that by X_s we denote the path of X starting from 0 up to time s, not a portion of the path of X of length T, which would be anyway well defined in this setting. If X solves (13) (in some space), for $t \in [t_0, T]$ we set

$$Y(t) = \begin{pmatrix} X(t) \\ \{X(t+s)\}_{s \in [-T,0]} \end{pmatrix}$$

and then differentiate with respect to t formally obtaining

(14) $$\frac{dY(t)}{dt} = \begin{pmatrix} \dot{X}(t) \\ \{\dot{X}(t+s)\}_s \end{pmatrix} = \begin{pmatrix} 0 \\ \{\dot{X}(t+s)\}_s \end{pmatrix} + \begin{pmatrix} b_t(X_t) \\ 0 \end{pmatrix} + \begin{pmatrix} \sigma \dot{W}(t) \\ 0 \end{pmatrix}.$$

It is therefore natural to define the operators

(15) $$A \begin{pmatrix} x \\ \varphi \end{pmatrix} := \begin{pmatrix} 0 \\ \dot{\varphi} \end{pmatrix},$$

(16) $$B \left(t, \begin{pmatrix} x \\ \varphi \end{pmatrix} \right) := \begin{pmatrix} \hat{b} \left(t, \begin{pmatrix} x \\ \varphi \end{pmatrix} \right) \\ 0 \end{pmatrix}$$

and

(17) $$\Sigma \begin{pmatrix} x \\ \varphi \end{pmatrix} := \begin{pmatrix} \sigma x \\ 0 \end{pmatrix}$$

and to formulate the infinite-dimensional SDE

(14′) $$dY(t) = AY(t)\, dt + B(t, Y(t))\, dt + \Sigma\, d\beta(t), \qquad t \in [t_0, T],$$

where β is given by

(18) $$\beta(t) = \begin{pmatrix} W(t) \\ 0 \end{pmatrix},$$

with some initial condition $Y(t_0) = y$.

Solutions of this SDE will always be understood to be *mild solutions*, that is, we want to solve

(14″) $$Y(t) = e^{(t-t_0)A} y + \int_{t_0}^t e^{(t-s)A} B(s, Y(s))\, ds + \int_{t_0}^t e^{(t-s)A} \Sigma\, d\beta(s).$$

It is not difficult to show that if Y solves (14′) then its first coordinate $X(t)$ solves the original SDE (13).

2.2. *Some properties of the convolution integrals.* The operator A has different domains depending on the space that we work in; we set

$$\mathrm{Dom}(A) = \left\{ \begin{pmatrix} x \\ \varphi \end{pmatrix} \in \mathcal{L}^p : \varphi \in W^{1,p}(-T, 0; \mathbb{R}^d), \varphi(0) = x \right\},$$

$$\mathrm{Dom}(A_{\widehat{\mathcal{C}}}) = \left\{ \begin{pmatrix} x \\ \varphi \end{pmatrix} \in \widehat{\mathcal{C}} : \varphi \in C^1([-T, 0); \mathbb{R}^d) \right\};$$

one can think to define A on $\mathcal{L}^p$ and then consider its restriction to $\mathcal{D}$ or to $\widehat{\mathcal{C}}$, as the notation above emphasizes.

It is well known [see Theorem 4.4.2 in Bensoussan et al. (1992)] that A is the infinitesimal generator of a strongly continuous semigroup both in $\mathcal{L}^p$ and in $\widehat{\mathcal{C}}$; it is easy to check that it still generates a semigroup in $\mathcal{D}$ which is not uniformly continuous. Indeed we have that

$$(19) \qquad e^{tA}\begin{pmatrix} x \\ \varphi \end{pmatrix} = \begin{pmatrix} x \\ \{\varphi(\xi + t)\mathbb{1}_{[-T,-t)}(\xi) + x\mathbb{1}_{[-t,0]}(\xi)\}_{\xi \in [-T,0]} \end{pmatrix}.$$

This formula comes from the trivial delay equation

$$\begin{cases} \dfrac{\mathrm{d}x(t)}{\mathrm{d}t} = 0, & t \geq 0, \\ x(0) = x, & x(\xi) = \varphi(\xi) \text{ for } \xi \in [-T, 0]; \end{cases}$$

its solution, for $t \geq 0$, is simply $x(t) = x$. If we introduce the pair

$$y(t) := \begin{pmatrix} x(t) \\ x_{|[t-T,t]} \end{pmatrix}$$

then

$$y(t) = e^{tA}\begin{pmatrix} x \\ \varphi \end{pmatrix}.$$

However, it still holds that

$$(20) \qquad \|e^{tA}\|_{L(\mathcal{D},\mathcal{D})} \leq C \qquad \text{for } t \in [0, T]$$

with C not depending on t. Moreover, it is evident from (19) that e^{tA} maps $\mathcal{L}^p$ into $\mathcal{L}^p$, $\mathcal{D}$ into $\mathcal{D}$ and $\widehat{\mathcal{C}}$ into $\widehat{\mathcal{C}}$, but it maps $\mathcal{C}$ into $\mathcal{D}_{-t}$ because an element of $\mathcal{C}$ is essentially a continuous function with a unique discontinuity at its endpoint, and the semigroup just shifts that discontinuity. In particular this happens for elements of $\mathbb{R}^d \times \{0\}$.

Consider the stochastic convolution

$$Z^{t_0}(t) := \int_{t_0}^{t} e^{(t-s)A} \Sigma \, \mathrm{d}\beta(s) = \int_{t_0}^{t} e^{(t-s)A}\begin{pmatrix} \sigma \, \mathrm{d}W(s) \\ 0 \end{pmatrix}, \qquad t \geq t_0.$$

It is not obvious to investigate Z^{t_0} by infinite-dimensional stochastic integration theory, due to the difficult nature of the Banach space $\mathcal{D}$. However, we may study its properties thanks to the following explicit formulas. From now on, we work in a set $\Omega_0 \subseteq \Omega$ of full probability on which W has continuous trajectories. For $\omega \in \Omega_0$ fixed, for any $x \in \mathbb{R}^d$ we have

$$e^{(t-s)A} \Sigma \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} \sigma x \\ \{\sigma x \mathbb{1}_{[-(t-s),0]}(\xi)\}_{\xi \in [-T,0]} \end{pmatrix}$$

hence

$$Z^{t_0}(t) = \begin{pmatrix} \displaystyle\int_{t_0}^{t} \sigma \, dW(s) \\ \displaystyle\int_{t_0}^{t} \mathbb{1}_{[-(t-s),0]}(\cdot)\sigma \, dW(s) \end{pmatrix}$$

(21)

$$= \begin{pmatrix} \sigma\big(W(t) - W(t_0)\big) \\ \sigma\big(W((t+\cdot) \vee t_0) - W(t_0)\big) \end{pmatrix}$$

because

$$\int_{t_0}^{t} \mathbb{1}_{[-(t-s),0]}(\xi)\sigma \, dW(s) = \int_{t_0}^{t} \mathbb{1}_{[0,t+\xi]}(s)\sigma \, dW(s).$$

From the previous formula, we see that $Z^{t_0}(t) \in \widehat{\mathcal{C}}$, hence $Z^{t_0}(t) \in \mathcal{L}^p$.

We have

$$\big\|Z^{t_0}(t)\big\|_{\widehat{\mathcal{C}}} = 2 \sup_{\xi\in[-T,0]} \big|\sigma\big(W((t+\xi) \vee t_0) - W(t_0)\big)\big|$$

hence [using the fact that $r \mapsto W(t_0+r) - W(t_0)$ is a Brownian motion and applying Doob's inequality]

$$\mathbb{E}\big[\big\|Z^{t_0}(t)\big\|_{\widehat{\mathcal{C}}}^4\big] \leq 2^4 \mathbb{E}\Big[\sup_{s\in[0,t-t_0]} \big|\sigma W(s)\big|^4\Big]$$

(22)

$$\leq C'\mathbb{E}\big[\big|W(t-t_0)\big|^4\big] \leq C''(t-t_0)^2,$$

where C' and C'' are suitable constants. Consequently, the same property holds in $\mathcal{L}^p$ (possibly with a different constant) by continuity of the embedding $\widehat{\mathcal{C}} \subset \mathcal{L}^p$. Moreover, from (21) we obtain that for ω fixed

$$\big\|Z^{t_0}(t) - Z^{t_0}(s)\big\|_{\widehat{\mathcal{C}}}$$

$$= C\Big(\big|W(t) - W(s)\big| + \sup_{\xi\in[-T,0]} \big|W((t+\xi) \vee t_0) - W((s+\xi) \vee t_0)\big|\Big).$$

Observe that (we suppose $s < t$ for simplicity)

$$W((t+\xi) \vee t_0) - W((s+\xi) \vee t_0)$$

$$= \begin{cases} 0, & \xi \in [-T, t_0-t], \\ W(t+\xi) - W(t_0), & \xi \in [t_0-t, t_0-s], \\ W(t+\xi) - W(s+\xi), & \xi \in [t_0-s, 0] \end{cases}$$

and

$$\sup_{\xi\in[t_0-t,t_0-s]} \big|W(t+\xi) - W(t_0)\big| = \sup_{\eta\in[t_0,t_0+(t-s)]} \big|W(\eta) - W(t_0)\big|,$$

therefore, Z^{t_0} is a continuous process in $\widehat{\mathcal{C}}$, since any fixed trajectory of W is uniformly continuous. The same property holds then in $\mathcal{L}^p$ again by continuity

of the embedding $\overset{\frown}{\mathcal{C}} \subset \mathcal{L}^p$. We can argue in a similar way for $F^{t_0} : [t_0, T] \times L^\infty([t_0, T]; \mathcal{D}) \to \mathcal{D}$,

$$F^{t_0}(t, \theta) = \int_{t_0}^t e^{(t-s)A} B(s, \theta(s)) \, ds.$$

From (16), using (19) one deduces that

$$e^{(t-s)A} B(s, \theta(s)) = \begin{pmatrix} b_s(\widetilde{M}_s \theta(s)) \\ b_s(\widetilde{M}_s \theta(s)) \mathbb{1}_{[-t+s]}(\xi) \end{pmatrix}$$

and, therefore,

$$\int_{t_0}^t e^{(t-s)A} B(s, \theta(s)) \, ds = \begin{pmatrix} \int_{t_0}^t b_s(\widetilde{M}_s \theta(s)) \, ds \\ \left\{ \int_{t_0}^{t+\xi} b_s(\widetilde{M}_s \theta(s)) \, ds \right\}_\xi \end{pmatrix}$$

which shows that $F^{t_0}(t, \theta)$ always belongs to $\overset{\frown}{\mathcal{C}}$. Writing

$$Y^{t_0, y}(t) = e^{(t-t_0)A} y + F^{t_0}(t, Y^{t_0, y}) + Z^{t_0}(t)$$

we see immediately that, for any $t \in [t_0, T]$, $Y^{t_0, y}(t) \in \mathcal{D}$ if $y \in \mathcal{D}$ and $Y^{t_0, y}(t) \in \overset{\frown}{\mathcal{C}}$ if $y \in \overset{\frown}{\mathcal{C}}$. This will be crucial in the sequel.

2.3. *Existence, uniqueness and differentiability of solutions to the SDE.* We state and prove here some abstract results about existence and differentiability of solutions to the stochastic equation

$$(14') \qquad dY(t) = AY(t) \, dt + B(t, Y(t)) \, dt + \Sigma \, d\beta(t), \qquad Y(t_0) = y,$$

with respect to the initial data. By abstract, we mean that we consider a general B not necessarily defined through a given b as in previous sections. Also A can be thought here to be a generic infinitesimal generator of a semigroup which is strongly continuous in $\mathcal{L}^p$ and satisfies (20) in $\mathcal{D}$. Although all these theorems are analogous to well-known results for stochastic equations in Hilbert spaces [see, e.g., Da Prato and Zabczyk (1992)], we give here complete and exact proofs due to the lack of them in the literature for the case of time-dependent coefficients in Banach spaces, which is the one of interest here.

We are interested in solving the SDE in $\mathcal{L}^p$ and in $\mathcal{D}$; since almost all the proofs can be carried out in the same way for each of the spaces we consider and since we do not need any particular property of these spaces themselves, we state all our results in this section in a general Banach space E, stressing out possible distinctions that could arise from different choices of E. In the following, we will identify $L(E, L(E, E))$ with $L(E, E; E)$ (the space of bilinear forms on E) in the usual way.

We will make the following assumption.

ASSUMPTION 2.1.

$$B \in L^{\infty}\big(0, T; C_b^{2,\alpha}(E, E)\big)$$

for some $\alpha \in (0, 1)$, where we have denoted by $C_b^{2,\alpha}(E, E)$ the space of twice Fréchet differentiable functions φ from E to E, bounded with their differentials of first and second order, such that $x \mapsto D^2\varphi(x)$ is α-Hölder continuous from E to $L(E, E; E)$. The L^{∞} property in time means that the differentials are measurable in (t, x) and both the function, the two differentials and the Hölder norms are bounded in time. Under these conditions, B, DB, D^2B are globally uniformly continuous on E [with values in E, $L(E, E)$, $L(E, E; E)$], respectively, and with a uniform in time modulus of continuity.

THEOREM 2.2. *Equation* (14$'$) *can be solved in a mild sense path by path*: *for any $y \in E$, any $t_0 \in [0, T]$ and every $\omega \in \Omega_0$ there exists a unique function* $[t_0, T] \ni t \to Y^{t_0, y}(t, \omega) \in E$ *which satisfies identity* (14$''$)

$$(14'') \qquad \begin{aligned} Y^{t_0, y}(t, \omega) &= e^{(t-t_0)A}y + \int_{t_0}^{t} e^{(t-s)A} B\big(s, Y^{t_0, y}(s, \omega)\big) \, ds \\ &\quad + \int_{t_0}^{t} e^{(t-s)A} \Sigma \, d\beta(s, \omega). \end{aligned}$$

Such a function is continuous if $E = \mathcal{L}^p$, it is only in L^{∞} if $E = \mathcal{D}$.

PROOF. Thanks to the Lipschitz property of B the proof follows through a standard argument based on the contraction mapping principle. The lack of continuity in $\mathcal{D}$ is due to the fact that the semigroup e^{tA} is not strongly continuous in $\mathcal{D}$. □

THEOREM 2.3. *For every $\omega \in \Omega_0$, for all $t_0 \in [0, T]$ and $t \in [t_0, T]$ the map* $y \mapsto Y^{t_0, y}(t, \omega)$ *is twice Fréchet differentiable and the map $y \mapsto D^2 Y^{t_0, y}(t, \omega)$ is* α-Hölder continuous from E to $L(E, E; E)$. Moreover, if $E = \mathcal{L}^p$, for any fixed t *and y the map $s \mapsto Y^{s, y}(t, \omega)$ is continuous. If $E = \mathcal{D}$, the same conclusion holds* *only for any fixed $y \in \overset{\frown}{\mathcal{C}}$.*

PROOF. Due to its length the proof is postponed to the Appendix. □

THEOREM 2.4. *If the solution $Y^{t_0, y}(t)$ is continuous as a function of t with values in E, then it has the Markov property.*

PROOF. This follows immediately from Theorem 9.15 on Da Prato and Zabczyk (1992). Notice that there the authors require a different set of hypothesis

which, however, are needed only for proving existence and uniqueness of solutions and not in the actual proof of the result. It therefore applies to our situation as well.

$\square$

In Section 4, we will need the notion of modulus of continuity for the second Fréchet derivative of a map from E into E, together with some of its properties; we summarize what we will need in the following general remark.

REMARK 2.5.   Given a map $R : E \to L(E, E; \mathbb{R})$, we define its modulus of continuity

$$\mathfrak{w}(R, r) = \sup_{\|y - y'\|_E \leq r} \|R(y) - R(y')\|_{L(E,E;\mathbb{R})}.$$

Let $v : E \to \mathbb{R}$ be a function with two Fréchet derivatives at each point, uniformly continuous on bounded sets. Then there exists a function $r_v : E^2 \to \mathbb{R}$ such that

$$v(x) - v(x_0) = \langle Dv(x_0), x - x_0 \rangle + \tfrac{1}{2} D^2 v(x_0)(x - x_0, x - x_0) + \tfrac{1}{2} r_v(x, x_0),$$

$$\left| r_v(x, x_0) \right| \leq \mathfrak{w}(D^2 v, \|x - x_0\|_E) \|x - x_0\|_E^2$$

for every $x, x_0 \in E$. Indeed,

$$v(x) - v(x_0) = \langle Dv(x_0), x - x_0 \rangle + \tfrac{1}{2} D^2 v(\xi_{x,x_0})(x - x_0, x - x_0),$$

where ξ_{v,x,x_0} is an intermediate point between x_0 and x, and thus

$$
\begin{aligned}
\left| r_v(x, x_0) \right| &= \left| (D^2 v(\xi_{v,x,x_0}) - D^2 v(x_0))(x - x_0, x - x_0) \right| \\
&\leq \left\| D^2 v(\xi_{v,x,x_0}) - D^2 v(x_0) \right\|_{L(E,E;\mathbb{R})} \|x - x_0\|_E^2 \\
&\leq \mathfrak{w}(D^2 v, \|x - x_0\|_E) \|x - x_0\|_E^2.
\end{aligned}
$$

If $D^2 v$ is α-Hölder continuous, namely

$$\left\| D^2 v(y) - D^2 v(y') \right\|_{L(E,E;\mathbb{R})} \leq M \|y - y'\|_E^\alpha$$

then

$$\mathfrak{w}(D^2 v, \|x - x_0\|_E) \leq M \|x - x_0\|_E^\alpha$$

and thus

$$\left| r_v(x, x_0) \right| \leq M \|x - x_0\|_E^{2+\alpha}.$$

3. The Kolmogorov equation.   In this and the following two sections, we introduce and solve the backward Kolmogorov equation in our infinite-dimensional setting. The relation between the results we shall show and the finite-dimensional path-dependent SDE we started from will be investigated in Section 7.

Suppose for a moment we are working in a standard Hilbert-space setting, that is, in a space $\mathcal{H} = \mathbb{R}^d \times H$ where H is a Hilbert space. Then [see again Da Prato and Zabczyk (1992)] the backward Kolmogorov equation, for the unknown $u :$ $[0, T] \times \mathcal{H} \to \mathbb{R}$, is

$$(23) \quad \begin{cases} \dfrac{\partial u}{\partial t}(t, y) + \dfrac{1}{2}\operatorname{Tr}\big(\Sigma^* \Sigma D^2 u(t, y)\big) + \langle Du(t, y), Ay + B(t, y)\rangle = 0, \\ u(T, \cdot) = \Phi, \end{cases}$$

where Φ is a given terminal condition and Du, $D^2 u$ represent the first and second Fréchet differentials with respect to the variable y. Its solution, under suitable hypothesis on A, B, Σ and Φ, is given by

$$(24) \qquad\qquad u(t, y) = \mathbb{E}\big[\Phi\big(Y^{t,y}(T)\big)\big],$$

where $Y^{t,y}(t)$ solves the associated SDE

$$dY(s) = \big[AY(s) + B\big(s, Y(s)\big)\big]\,ds + \Sigma\,d\beta(s),$$

$(14'\text{ bis})$

$$s \in [t, T], Y(t) = y$$

in $\mathcal{H}$. In our framework, where the spaces are only Banach spaces, we have to give a precise meaning to the Kolmogorov equation and prove its relation above with the SDE.

As outlined in the Introduction, we would like to solve it on the space $\overset{\frown}{\mathcal{C}}$, but since $B(t, y)$ belongs to $\mathbb{R}^d \times \{0\} \not\subseteq \overset{\frown}{\mathcal{C}}$, in order to give meaning to the term $\langle Du(t, y), B(t, y)\rangle$ we need $Du(t, y)$ to be a functional defined at least on $\mathcal{C}$, which necessarily implies u to be defined on $[0, T] \times \mathcal{C}$. Therefore, we should solve (in mild sense) the SDE for $y \in \mathcal{C}$ and this implies that $Y^{t,y}(s) \in \mathcal{D}_{-t+s}$ for $s \neq t$; this in turn requires Φ to be defined at least on $\bigcup_{s \in [t,T]} \mathcal{D}_{-t+s}$ in order for a function of the form (24) to be well defined. However, the space $\bigcup \mathcal{D}_s$ is not a linear space, thus it turns out that it is more convenient, also for exploiting a Banach space structure, to formulate everything in $\mathcal{D}$, that is,

$$u : [0, T] \times \mathcal{D} \to \mathbb{R}.$$

Therefore, we interpret $\langle \cdot, \cdot \rangle$ in this setting as the duality pairing between $\mathcal{D}'$ and $\mathcal{D}$.

For the trace term, if we denote by $e_1, \ldots, e_d$ an orthonormal basis of $\mathbb{R}^d$ where σ diagonalizes, that is, $\sigma e_j = \sigma_j e_j$ for some real σ_j (in any of the spaces considered up to now), we could complete it to an orthonormal system $\{e_n\}$ in $\mathcal{H}$ obtaining that

$$\operatorname{Tr}\big(\Sigma^* \Sigma D^2 u(t, y)\big) = \sum_j \sigma_j^2 \langle D^2 u(t, y) e_j, e_j \rangle;$$

hence, by analogy, also when working in $\mathcal{D}$ we interpret the trace term as

$$(25) \qquad\qquad \operatorname{Tr}\big(\Sigma^* \Sigma D^2 u(t, y)\big) = \sum_{j=1}^{d} \sigma_j^2 D^2 u(t, y)(e_j, e_j).$$

Moreover, we consider Kolmogorov equation in its integrated form with respect to time, that is, given a (sufficiently regular; see below) real function Φ on $\mathcal{D}$ we seek for a solution of the PDE:

$$
\begin{aligned}
u(t, y) - \Phi(y) = & \int_t^T \langle Du(s, y), Ay + B(s, y) \rangle \, ds \\
& + \frac{1}{2} \int_t^T \sum_{j=1}^d \sigma_j^2 D^2 u(s, y)(e_j, e_j) \, ds.
\end{aligned}
$$

(26)

Here, one can see one of the difficulties in working with Banach spaces: the second-order term in the equation comes from the quadratic variation of the solution of the SDE, but in such spaces there is no general way of defining a quadratic variation [although, as mentioned at the beginning, a general theory of quadratic variation is currently being developed by F. Russo and collaborators; see the works Di Girolami and Russo (2014), Di Girolami, Fabbri and Russo (2014) and the references therein].

Although we will seek for such a u, when dealing with the equation we will always choose y to be in $\mathrm{Dom}(A_{\widehat{\mathcal{C}}})$, to let all the terms appearing there be well defined.

All these observations lead to our definition of solution to (26); first, we say that a functional u on $[0, T] \times \mathcal{D}$ belongs to

$$
L^\infty\big(0, T; C_b^{2,\alpha}(\mathcal{D}, \mathbb{R})\big)
$$

if it is twice Fréchet differentiable on $\mathcal{D}$, u, Du and $D^2 u$ are bounded, the map $x \mapsto D^2 u(x)$ is α-Hölder continuous from $\mathcal{D}$ to $L(\mathcal{D}, \mathcal{D}; \mathcal{D})$ (the space of bilinear forms on $\mathcal{D}$), the differentials are measurable in (t, x) and the function, the two differentials and the Hölder norms are bounded in time.

DEFINITION 3.1. Given $\Phi \in C_b^{2,\alpha}(\mathcal{D}, \mathbb{R})$, we say that $u : [0, T] \times \mathcal{D} \to \mathbb{R}$ is a classical solution of the Kolmogorov equation with terminal condition Φ if

$$
u \in L^\infty\big(0, T; C_b^{2,\alpha}(\mathcal{D}, \mathbb{R})\big) \cap C\big([0, T] \times \widehat{\mathcal{C}}, \mathbb{R}\big),
$$

$u(\cdot, y)$ is Lipschitz for any $y \in \mathrm{Dom}(A_{\widehat{\mathcal{C}}})$ and satisfies identity (26) for every $t \in [0, T]$ and $y \in \mathrm{Dom}(A_{\widehat{\mathcal{C}}})$, with the duality terms understood with respect to the topology of $\mathcal{D}$.

It will be clear in Section 5 that the restriction $y \in \mathrm{Dom}(A_{\widehat{\mathcal{C}}})$ is necessary and that it would not be possible to obtain the same result choosing y in some larger space.

Our aim is to show that, in analogy with the classical case, the function

$$
u(t, y) = \mathbb{E}\big[\Phi\big(Y^{t,y}(T)\big)\big]
$$

solves equation (26).

However, we are not able to prove this result directly, due essentially to the lack of an appropriate Itô-type formula for our setting. Therefore, we will proceed as follows: first, we are going to show how to prove such a result in $\mathcal{L}^p$, then we will show that if the problem is formulated in $\mathcal{D}$ it is possible to approximate it with a sequence of $\mathcal{L}^p$ problems; the solutions to such approximating problems will be finally shown to converge to a function that solves the Komogorov backward PDE in the sense of Definition 3.1.

All the above discussion about the meaning of Kolmogorov equation applies verbatim to the space $\mathcal{L}^p$. A solution in $\mathcal{L}^p$ is defined in a straightforward way as follows.

DEFINITION 3.2.   Given $\Phi \in C_b^{2,\alpha}(\mathcal{L}^p, \mathbb{R})$, we say that $u : [0, T] \times \mathcal{L}^p \to \mathbb{R}$ is a solution of the Kolmogorov equation in $\mathcal{L}^p$ with terminal condition Φ if

$$u \in L^\infty(0, T; C_b^{2,\alpha}(\mathcal{L}^p, \mathbb{R})) \cap C([0, T] \times \mathcal{L}^p, \mathbb{R})$$

$u(\cdot, y)$ is Lipschitz for any $y \in \mathrm{Dom}(A)$ and satisfies identity (26) for every $t \in [0, T]$ and $y \in \mathrm{Dom}(A)$, with the duality terms understood with respect to the topology of $\mathcal{L}^p$.

4. Solution in $\mathcal{L}^p$.   The choice to work in a general $\mathcal{L}^p$ space instead of working with the Hilbert space $\mathcal{L}^2$ could seem unjustified at first sight. As long as solving Kolmogorov equation in $\mathcal{L}^p$ is only a step toward solving it in $\mathcal{D}$ through approximations it would be enough to develop the theory in $\mathcal{L}^2$, where the results needed are well known. Nevertheless, we give and prove here this more general statement for $\mathcal{L}^p$ spaces for some reasons. First, the proof shows a method to obtain this kind of result without actually using a Itô-type formula, but only a Taylor expansion; the difference is tiny but it allows to work in spaces where there is no Itô formula to apply. Second, the proof points out where a direct argument of this kind (which is essentially the classical scheme for these results) fails. Last, also the easiest examples do not behave well in $\mathcal{L}^2$ but they can be regular enough in some $\mathcal{L}^p$ instead (see Examples 6.1 and 6.2 hereinafter). Therefore, proving the result in $\mathcal{L}^p$ is already enough to deal with some examples, without the need to go further in the development of the theory.

If B satisfies Assumption 2.1 with $E = \mathcal{L}^p$, Theorems 2.2, 2.3 and 2.4 yield that the SDE

$$dY(s) = [AY(s) + B(s, Y(s))]\,ds + \Sigma\,d\beta(s),$$

(14′ bis)

$$s \in [t, T], Y(t) = y$$

admits a unique mild solution $Y^{t_0,y}(t)$ in $\mathcal{L}^p$ which is continuous in time, $C_b^{2,\alpha}$ with respect to y and has the Markov property.

THEOREM 4.1. *Let* $\Phi : \mathcal{L}^p \to \mathbb{R}$ *be in* $C_b^{2,\alpha}$ *and let Assumption* 2.1 *hold in* $\mathcal{L}^p$. *Then the function*

$$u(t, y) := \mathbb{E}[\Phi(Y^{t,y}(T))], \qquad (t, y) \in [0, T] \times \mathcal{L}^p,$$

is a solution of the Kolmogorov equation in $\mathcal{L}^p$ *with terminal condition* Φ.

PROOF. Throughout this proof $\|\cdot\|$ will denote the norm in $\mathcal{L}^p$ and $\langle\cdot,\cdot\rangle$ will denote duality between $\mathcal{L}^p$ and $\mathcal{L}^{p'}$, where $\frac{1}{p} + \frac{1}{p'} = 1$.

The function u has the regularity properties required by the definition of solution: boundedness in time is straightforward, while the fact that Φ belongs to $C_b^{2,\alpha}(\mathcal{L}^p; \mathbb{R}^d)$ and the regularity properties of Y with respect to the initial data stated in Theorem 2.3 imply, by composition and the dominated convergence theorem, that u is continuous on $[0, T] \times \mathcal{L}^p$ and $u(t, \cdot)$ is in $C_b^{2,\alpha}(\mathcal{L}^p; \mathbb{R}^d)$ for every $t \in [0, T]$; the Lipschitz property in time is a consequence of being a solution of an integral equation where all the terms are bounded. We have thus to show that it satisfies equation (26). Recall that we choose y in the domain of A.

Step 1. Fix $t_0 \in [0, T]$. From Markov property, for any $t_1 > t_0$ in $[0, T]$, we have

$$u(t_0, y) = \mathbb{E}[u(t_1, Y^{t_0,y}(t_1))]$$

because

$$\begin{aligned}
\mathbb{E}[\Phi(Y^{t_0,y}(T))] &= \mathbb{E}[\mathbb{E}[\Phi(Y^{t_0,y}(T))|Y^{t_0,y}(t_1)]] \\
&= \mathbb{E}[\mathbb{E}[\Phi(Y^{t_1,w}(T))]_{w=Y^{t_0,y}(t_1)}] = \mathbb{E}[u(t_1, Y^{t_0,y}(t_1))].
\end{aligned}$$

From Taylor formula applied to the function $y \mapsto u(t, y)$, we have

$$\begin{aligned}
u(t_1, &Y^{t_0,y}(t_1)) - u(t_1, e^{(t_1-t_0)A}y) \\
&= \langle Du(t_1, e^{(t_1-t_0)A}y), Y^{t_0,y}(t_1) - e^{(t_1-t_0)A}y\rangle \\
&\quad + \tfrac{1}{2}D^2u(t_1, e^{(t_1-t_0)A}y)(Y^{t_0,y}(t_1) - e^{(t_1-t_0)A}y, Y^{t_0,y}(t_1) - e^{(t_1-t_0)A}y) \\
&\quad + \tfrac{1}{2}r_{u(t_1,\cdot)}(Y^{t_0,y}(t_1), e^{(t_1-t_0)A}y),
\end{aligned}$$

where

$$\begin{aligned}
|r_{u(t_1,\cdot)}&(Y^{t_0,y}(t_1), e^{(t_1-t_0)A}y)| \\
&\leq \mathfrak{w}(D^2u(t_1, \cdot), \|Y^{t_0,y}(t_1) - e^{(t_1-t_0)A}y\|)\|Y^{t_0,y}(t_1) - e^{(t_1-t_0)A}y\|^2
\end{aligned}$$

(for the definitions of r and $\mathfrak{w}$ see Remark 2.5). Due to the $C_b^{2,\alpha}(\mathcal{L}^p, \mathbb{R})$-property, uniform in time, we have

$$|r_{u(t_1,\cdot)}(Y^{t_0,y}(t_1), e^{(t_1-t_0)A}y)| \leq M\|Y^{t_0,y}(t_1) - e^{(t_1-t_0)A}y\|^{2+\alpha}.$$

Recall that

$$Y^{t_0, y}(t_1) - e^{(t_1 - t_0)A} y = F^{t_0}(t_1, Y^{t_0, y}) + Z^{t_0}(t_1),$$

$$F^{t_0}(t_1, Y^{t_0, y}) = \int_{t_0}^{t_1} e^{(t_1 - s)A} B(s, Y^{t_0, y}(s)) \, ds$$

and

$$\mathbb{E}[Z^{t_0}(t_1)] = 0,$$

$$\mathbb{E}[\|Z^{t_0}(t_1)\|^4] \leq C_Z^4 (t_1 - t_0)^2,$$

$$\|F^{t_0}(t_1, Y^{t_0, y})\| \leq C \|B\|_{\infty} (t_1 - t_0),$$

where $\|B\|_{\infty} = \sup_t \sup_y \|B(t, y)\|$.

Hence, recalling $u(t_0, y) = \mathbb{E}[u(t_1, Y^{t_0, y}(t_1))]$,

$$u(t_0, y) - u(t_1, e^{(t_1 - t_0)A} y)$$

$$= \langle Du(t_1, e^{(t_1 - t_0)A} y), \mathbb{E}[F^{t_0}(t_1, Y^{t_0, y})]\rangle$$

$$+ \tfrac{1}{2} \mathbb{E}[D^2 u(t_1, e^{(t_1 - t_0)A} y)(F^{t_0}(t_1, Y^{t_0, y}) + Z^{t_0}(t_1),$$

$$F^{t_0}(t_1, Y^{t_0, y}) + Z^{t_0}(t_1))]$$

$$+ \tfrac{1}{2} \mathbb{E}[r_{u(t_1, \cdot)}(Y^{t_0, y}(t_1), e^{(t_1 - t_0)A} y)].$$

Step 2. Now let us explain the strategy. Given $t \in [0, T]$, taken a sequence of partitions π_n of $[t, T]$, of the form $t = t_1^n \leq \cdots \leq t_{k_n + 1}^n = T$ of $[t, T]$ with $|\pi_n| \to 0$, we take $t_0 = t_i^n$ and $t_1 = t_{i+1}^n$ in the previous identity and sum over the partition π_n to get

$$u(t, y) - \Phi(y) + I_n^{(1)} = I_n^{(2)} + I_n^{(3)} + I_n^{(4)},$$

where

$$I_n^{(1)} := \sum_{i=1}^{k_n} (u(t_{i+1}^n, y) - u(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A} y)),$$

$$I_n^{(2)} := \sum_{i=1}^{k_n} \langle Du(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A} y), \mathbb{E}[F^{t_i^n}(t_{i+1}^n, Y^{t_i^n, y})]\rangle,$$

$$I_n^{(3)} := \frac{1}{2} \sum_{i=1}^{k_n} \mathbb{E}[D^2 u(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A} y)$$

$$\times (F^{t_i^n}(t_{i+1}^n, Y^{t_i^n, y}) + Z^{t_i^n}(t_{i+1}^n), F^{t_i^n}(t_{i+1}^n, Y^{t_i^n, y}) + Z^{t_i^n}(t_{i+1}^n))],$$

$$I_n^{(4)} := \frac{1}{2} \sum_{i=1}^{k_n} \mathbb{E}[r_{u(t_{i+1}^n, \cdot)}(Y^{t_i^n, y}(t_{i+1}^n), e^{(t_{i+1}^n - t_i^n)A} y)].$$

We want to show that:

2660 F. FLANDOLI AND G. ZANCO

(I) $\lim_{n\to\infty} I_n^{(1)} = -\int_t^T \langle Du(s,y), Ay\rangle\,ds$ if $y\in\mathrm{Dom}(A)$,

(II) $\lim_{n\to\infty} I_n^{(2)} = \int_t^T \langle Du(s,y), B(s,y)\rangle\,ds$,

(III) $\lim_{n\to\infty} I_n^{(3)} = \frac{1}{2}\int_t^T \sum_{j=1}^d \sigma_j^2 D^2 u(s,y)(e_j,e_j)\,ds$,

(IV) $\lim_{n\to\infty} I_n^{(4)} = 0$.

Step 3. We have, for $y\in\mathrm{Dom}(A)$ (in this case $\frac{d}{dt}e^{tA}y = Ae^{tA}y$)

$$\sum_i^{k_n} u(t_{i+1}^n, y) - u(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A}y)$$

$$= -\sum_i^{k_n} \int_0^{t_{i+1}^n - t_i^n} \langle Du(t_{i+1}^n, e^{sA}y), Ae^{sA}y\rangle\,ds$$

$$= -\sum_i^{k_n} \int_{t_i^n}^{t_{i+1}^n} \langle Du(t_{i+1}^n, e^{(s-t_i^n)A}y), Ae^{(s-t_i^n)A}y\rangle\,ds$$

$$= -\int_t^T \sum_i^{k_n} \langle Du(t_{i+1}^n, e^{(s-t_i^n)A}y), Ae^{(s-t_i^n)A}y\rangle \mathbb{1}_{[t_i^n, t_{i+1}^n]}(s)\,ds.$$

The semigroup e^{tA} is strongly continuous in $\mathcal{L}^p$ therefore it converges to the identity as t goes to 0; hence, since y is fixed, taking the limit in n yields (I) applying the dominated convergence theorem.

Step 4. By standard properties of the Bochner integral, we have

$$\sum_{i=1}^{k_n} \Big\langle Du(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A}y), \mathbb{E}\int_{t_i^n}^{t_{i+1}^n} e^{(t_{i+1}^n - s)A}B(s, Y^{t_i^n, y}(s))\,ds\Big\rangle$$

$$= \sum_{i=1}^{k_n} \mathbb{E}\int_{t_i^n}^{t_{i+1}^n} \langle Du(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A}y), e^{(t_{i+1}^n - s)A}B(s, Y^{t_i^n, y}(s))\rangle\,ds$$

$$= \mathbb{E}\int_t^T \sum_{i=1}^{k_n} \langle Du(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A}y), e^{(t_{i+1}^n - s)A}B(s, Y^{t_i^n, y}(s))\rangle \mathbb{1}_{[t_i^n, t_{i+1}^n]}(s)\,ds;$$

now arguing as in the previous step it's easy to prove that this quantity converges to

$$\int_t^T \langle Du(s,y), B(s,y)\rangle\,ds.$$

Step 5. First, split each of the addends appearing in $I_n^{(3)}$ as follows:

$$D^2 u(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A}y)$$
$$\times \big(F^{t_i^n}(t_{i+1}^n, Y^{t_i^n, y}) + Z^{t_i^n}(t_{i+1}^n), F^{t_i^n}(t_{i+1}^n, Y^{t_i^n, y}) + Z^{t_i^n}(t_{i+1}^n)\big)$$

$$= D^2u\big(t^n_{i+1}, e^{(t^n_{i+1}-t^n_i)A}y\big)\big(F^{t^n_i}\big(t^n_{i+1}, Y^{t^n_i,y}\big), F^{t^n_i}\big(t^n_{i+1}, Y^{t^n_i,y}\big)\big)$$
$$+ D^2u\big(t^n_{i+1}, e^{(t^n_{i+1}-t^n_i)A}y\big)\big(F^{t^n_i}\big(t^n_{i+1}, Y^{t^n_i,y}\big), Z^{t^n_i}\big(t^n_{i+1}\big)\big)$$
$$+ D^2u\big(t^n_{i+1}, e^{(t^n_{i+1}-t^n_i)A}y\big)\big(Z^{t^n_i}\big(t^n_{i+1}\big), F^{t^n_i}\big(t^n_{i+1}, Y^{t^n_i,y}\big)\big)$$
$$+ D^2u\big(t^n_{i+1}, e^{(t^n_{i+1}-t^n_i)A}y\big)\big(Z^{t^n_i}\big(t^n_{i+1}\big), Z^{t^n_i}\big(t^n_{i+1}\big)\big).$$

Let us give the main estimates. We have

$$\big|\mathbb{E}\big[D^2u\big(t, e^{(t-t_0)A}y\big)\big(F^{t_0}(t, Y^{t_0,y}), F^{t_0}(t, Y^{t_0,y})\big)\big]\big|$$
$$\leq \|D^2u\|_\infty \mathbb{E}\big[\|F^{t_0}(t, Y^{t_0,y})\|^2\big]$$
$$\leq \|D^2u\|_\infty C^2\|B\|^2_\infty(t-t_0)^2$$

and

$$\big|\mathbb{E}\big[D^2u\big(t, e^{(t-t_0)A}y\big)\big(F^{t_0}(t, Y^{t_0,y}), Z^{t_0}(t)\big)\big]\big|$$
$$\leq \|D^2u\|_\infty \mathbb{E}\big[\|F^{t_0}(t, Y^{t_0,y})\|^2\big]^{1/2}\mathbb{E}\big[\|Z^{t_0}(t)\|^2\big]^{1/2}$$
$$\leq \|D^2u\|_\infty C \cdot C_Z\|B\|_\infty(t-t_0)^{3/2},$$

where we have set

$$\|D^2u\|_\infty = \sup_t \sup_y \|D^2u(t, y)\|_{L(E,E;E)},$$

hence the first three terms give no contribution when summing up over i, because they are estimated by a power of $t_{i+1} - t_i$ greater than 1.

Therefore, it remains to show that the term

$$(27) \qquad \sum_{i=1}^{k_n} \mathbb{E}\big[D^2u\big(t^n_{i+1}, e^{(t^n_{i+1}-t^n_i)A}y\big)\big(Z^{t^n_i}\big(t^n_{i+1}\big), Z^{t^n_i}\big(t^n_{i+1}\big)\big)\big]$$

converges to

$$\int_{t_0}^t \sigma^2 D^2u(s, y)(e, e)\,\mathrm{d}s.$$

To this aim, we recall that

$$Z^{t^n_i}\big(t^n_{i+1}\big) = \int_{t^n_i}^{t^n_{i+1}} e^{(t^n_{i+1}-r)A}\begin{pmatrix} \sigma\,\mathrm{d}W(r) \\ 0 \end{pmatrix}$$
$$= \begin{pmatrix} \sigma\big(W(t^n_{i+1}) - W(t^n_i)\big) \\ \sigma\big(W((t^n_{i+1} + \cdot) \vee t^n_i) - W(t^n_i)\big) \end{pmatrix}$$
$$=: \begin{pmatrix} Z^i_0 \\ Z^i_1 \end{pmatrix}.$$

We split again (27) into

$$\sum_{i=1}^{k_n} \mathbb{E}\bigg[D^2 u\big(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A} y\big)\bigg(\begin{pmatrix} Z_0^i \\ 0 \end{pmatrix}, \begin{pmatrix} Z_0^i \\ 0 \end{pmatrix} \bigg)$$

$$+ D^2 u\big(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A} y\big)\bigg(\begin{pmatrix} Z_0^i \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ Z_1^i \end{pmatrix} \bigg)$$

$$+ D^2 u\big(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A} y\big)\bigg(\begin{pmatrix} 0 \\ Z_1^i \end{pmatrix}, \begin{pmatrix} Z_0^i \\ 0 \end{pmatrix} \bigg)$$

$$+ D^2 u\big(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A} y\big)\bigg(\begin{pmatrix} 0 \\ Z_1^i \end{pmatrix}, \begin{pmatrix} 0 \\ Z_1^i \end{pmatrix} \bigg)\bigg].$$

For the first term we have, using Itô isometry in $\mathbb{R}^d$, that

$$\sum_{i=1}^{k_n} \mathbb{E}\bigg[D^2 u\big(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A} y\big)\bigg(\begin{pmatrix} Z_0^i \\ 0 \end{pmatrix}, \begin{pmatrix} Z_0^i \\ 0 \end{pmatrix} \bigg)\bigg]$$

$$= \sum_{j=1}^{d} \sigma_j^2 \sum_{i=1}^{k_n} D^2 u\big(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A} y\big)(e_j, e_j)(t_{i+1}^n - t_i^n)$$

and the right-hand side in this equation converges to

$$\sum_{j=1}^{d} \sigma_j^2 \int_{t_0}^{t} D^2 u(s, y)(e_j, e_j)\, ds$$

thanks to the strong continuity of e^{tA}.

For the second term, we can write (here $\|\sigma\| = \max_j |\sigma_j|$)

$$(28) \quad \mathbb{E}\bigg| D^2 u\big(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A} y\big)\bigg(\begin{pmatrix} Z_0^i \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ Z_1^i \end{pmatrix} \bigg)\bigg|$$

$$\leq \|\sigma\| \|D^2 u\|_\infty \mathbb{E}\big[|W(t_{i+1}^n) - W(t_i^n)| \|W((t_{i+1}^n + \cdot) \vee t_i^n) - W(t_i^n)\|_{L^p} \big]$$

$$\leq \|\sigma\| \|D^2 u\|_\infty \mathbb{E}\bigg[|W(t_{i+1}^n) - W(t_i^n)| \bigg(\int_0^{t_{i+1}^n - t_i^n} |W(r)|^p\, dr \bigg)^{1/p} \bigg]$$

$$\leq \|\sigma\| \|D^2 u\|_\infty \big(\mathbb{E}|W(t_{i+1}^n) - W(t_i^n)|^2 \big)^{1/2}$$

$$\times \bigg(\mathbb{E}\bigg[\bigg(\int_0^{t_{i+1}^n - t_i^n} |W(r)|^p\, dr \bigg)^{2/p} \bigg] \bigg)^{1/2}$$

$$\leq \|\sigma\| \|D^2 u\|_\infty (t_{i+1}^n - t_i^n)^{1/2} (t_{i+1}^n - t_i^n)^{1/p}$$

$$\times \bigg(\mathbb{E}\bigg[\bigg(\sup_{[0, t_{i+1}^n - t_i^n]} (|W(r)|^p) \bigg)^{2/p} \bigg] \bigg)^{1/2}$$

$$(29) \quad \leq \|\sigma\| \|D^2 u\|_\infty (t_{i+1}^n - t_i^n)^{1+1/p},$$

using Itô isometry and Burkholder–Davis–Gundy inequality, thus it converges to zero when summing over i and letting n go to ∞.

The third term can be shown to go to zero in the exact same way and by the same estimates as above, one obtains that

$$\mathbb{E}\left|D^2 u\big(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A} y\big)\left(\begin{pmatrix} 0 \\ Z_1^i \end{pmatrix}, \begin{pmatrix} 0 \\ Z_1^i \end{pmatrix}\right)\right| \leq (t_{i+1}^n - t_i^n)^{1+2/p},$$

hence it follows that also this term gives no contribution when passing to the limit.

Step 6. Since

$$\left|r_{u(t,\cdot)}\big(Y^{t_0,y}(t), e^{(t-t_0)A} y\big)\right| \leq M \|Y^{t_0,y}(t) - e^{(t-t_0)A} y\|^{2+\alpha}$$

we have that

$$\begin{aligned}
&\left|\mathbb{E}\big[r_{u(t,\cdot)}\big(Y^{t_0,y}(t), e^{(t-t_0)A_E} y\big)\big]\right| \\
&\quad \leq M \mathbb{E}\big[\|Y^{t_0,y}(t) - e^{(t-t_0)A} y\|^{2+\alpha}\big] \\
&\quad \leq K\big(\mathbb{E}\big[\|F^{t_0}(t, Y^{t_0,y})\|^4\big]^{(2+\alpha)/4} + \mathbb{E}\big[\|Z^{t_0}(t)\|^4\big]^{(2+\alpha)/4}\big) \\
&\quad \leq \widetilde{K}(t-t_0)^{1+\alpha/2}
\end{aligned}$$

and from this one proves that $\lim_{n\to\infty} I_n^{(4)} = 0$. $\quad\square$

REMARK 4.2. The point in which the above argument fails when working directly in $\mathcal{D}$ is item (III) of step 2. Indeed step 5, which is the proof of the convergence in (III), cannot be carried out when working with the sup-norm: if we start again from (28) using the norm of $\mathcal{D}$ we would end up with the estimate

$$(29') \qquad \mathbb{E}\left|D^2 u\left(t_{i+1}^n, e^{(t_{i+1}^n - t_i^n)A} y\right)\left(\begin{pmatrix} Z_0^i \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ Z_1^i \end{pmatrix}\right)\right| \leq \|D^2 u\|_\infty (t_{i+1}^n - t_i^n)$$

which is not enough to obtain the convergence to 0 that we need.

5. Solution in $\overset{\frown}{\mathcal{C}}$. We now show how to use $\mathcal{L}^p$ approximations in order to obtain classical solutions of Kolmogorov equations in the sense of Definition 3.1. As before, we will assume that B satisfied Assumption 2.1 for $E = \mathcal{D}$, that is,

$$B \in L^\infty\big(0, T; C_b^{2,\alpha}(\mathcal{D}, \mathcal{D})\big)$$

for some $\alpha \in (0, 1)$. Suppose we have a sequence $\{J_n\}$ of linear continuous operators from $L^p(-T, 0; \mathbb{R}^d)$ into $C([-T, 0]; \mathbb{R}^d)$ such that $J_n \varphi \overset{n\to\infty}{\longrightarrow} \varphi$ uniformly for any $\varphi \in C([-T, 0]; \mathbb{R}^d)$. By the Banach–Steinhaus theorem, we have that $\sup_n \|J_n\|_{L(C([-T,0];\mathbb{R}^d);C([-T,0];\mathbb{R}^d))} < \infty$; however, we need a slightly stronger property, namely that $\|J_n f\|_\infty \leq C_J \|f\|_\infty$ for all f with at most one jump, uniformly in n. Then we can define the sequence of operators

$$(30) \qquad B_n : [0, T] \times \mathcal{L}^p \to \mathbb{R}^d \times \{0\},$$

$$B_n(t, y) = B_n\left(t, \begin{pmatrix} x \\ \varphi \end{pmatrix}\right) = B_n(t, x, \varphi) := B(t, x, J_n \varphi).$$

We will often write $J_n \binom{x}{\varphi}$ for $\binom{x}{J_n \varphi}$ in the sequel.

It can be easily proved that if B satisfies Assumption 2.1 in $\mathcal{D}$ then for every n the operator B_n satisfies Assumption 2.1 both in $\mathcal{D}$ and in $\mathcal{L}^p$. Thus, if we consider the approximated SDE

$$(31) \quad dY_n(t) = AY_n(t)\,dt + B_n\big(t, Y_n(t)\big)\,dt + \Sigma\,d\tilde{\beta}(t), \qquad Y_n(s) = y \in \mathcal{L}^p$$

by Theorem 2.2 it admits a unique path by path mild solution $Y_n^{s,y}$ such that, thanks to Theorem 2.3, the map $t \mapsto Y_n^{s,y}(t)$ is in $C_b^{2,\alpha}$. Suppose also we are given a terminal condition $\Phi : \mathcal{D} \to \mathbb{R}$ for the backward Kolmogorov equation (26) associated to the original problem with B; approximations Φ_n can be defined in the exact same way. We have then a sequence of approximated backward Kolmogorov equations in $\mathcal{L}^p$, namely

$$
\begin{aligned}
(32) \quad u_n(t, y) - \Phi(y) &= \int_t^T \big\langle Du_n(s, y), Ay + B_n(s, y)\big\rangle\,ds \\
&\quad + \frac{1}{2}\int_t^T \sum_{j=1}^d \sigma_j^2 D^2 u_n(s, y)(e_j, e_j)\,ds
\end{aligned}
$$

with terminal condition $u_n(T, \cdot) = \Phi_n$. Theorem 4.1 yields in fact that for each n the function

$$(33) \qquad u_n(s, y) = \mathbb{E}\big[\Phi_n\big(Y_n^{s,y}(T)\big)\big]$$

is a solution to equation (32) in $\mathcal{L}^p$. If we choose the initial condition y in the space $\widehat{\mathcal{C}}$ then $Y_n^{s,y}(t) \in \widehat{\mathcal{C}}$ as well for every n and every $t \in [s, T]$.

An example of a sequence $\{J_n\}$ satisfying the required properties can be constructed as follows: for any $\varepsilon \in (0, \frac{T}{2})$ define a function $\tau_\varepsilon : [-T, 0] \to [-T, 0]$ as

$$
\tau_\varepsilon(x) = \begin{cases} -T + \varepsilon, & \text{if } x \in [-T, -T + \varepsilon], \\ x, & \text{if } x \in [-T + \varepsilon, -\varepsilon], \\ -\varepsilon, & \text{if } x \in [-\varepsilon, 0]. \end{cases}
$$

Then choose any $C^\infty(\mathbb{R}; \mathbb{R})$ function ρ such that $\|\rho\|_1 = 1$, $0 \le \rho \le 1$ and $\operatorname{supp}(\rho) \subseteq [-1, 1]$ and define a sequence $\{\rho_n\}$ of mollifiers by $\rho_n(x) := n\rho(nx)$. Finally set, for any $\varphi \in L^1(-T, 0; \mathbb{R}^d)$

$$(34) \qquad J_n\varphi(x) := \int_{-T}^0 \rho_n\big(\tau_{1/n}(x) - y\big)\varphi(y)\,dy.$$

We will need one further assumption, together with the required properties for J_n that we write again for future reference.

DEFINITION 5.1. Let F be a Banach space, $R : \mathcal{D} \to F$ a twice Fréchet differentiable function and $\Gamma \subseteq \mathcal{D}$. We say that R has *one-jump-continuous Fréchet*

differentials of first and second order on Γ if there exists a sequence of linear continuous operators $J_n : L^p(-T, 0; \mathbb{R}^d) \to C([-T, 0]; \mathbb{R}^d)$ such that $J_n \varphi \overset{n \to \infty}{\longrightarrow} \varphi$ uniformly for any $\varphi \in C([-T, 0]; \mathbb{R}^d)$, $\sup_n \|J_n \varphi\|_\infty \leq C_J \|\varphi\|_\infty$ for every φ that has at most one jump and is continuous elsewhere and such that for every $y \in \Gamma$ and for almost every $a \in [-T, 0]$ the following hold:

$$DR(y)J_n \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix} \longrightarrow DR(y) \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix},$$

$$D^2 R(y) \left(J_n \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix} - \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix}, \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix} \right) \longrightarrow 0,$$

$$D^2 R(y) \left(\begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix}, J_n \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix} - \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix} \right) \longrightarrow 0,$$

$$D^2 R(y) \left(J_n \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix} - \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix}, J_n \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix} - \begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix} \right) \longrightarrow 0,$$

where we adopt the convention that $\begin{pmatrix} 1 \\ \mathbb{1}_{[a,0)} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ when $a = 0$.

We will call a sequence $\{J_n\}$ as above a *smoothing sequence*.

ASSUMPTION 5.2. For any $r \in [0, T]$, $B(r, \cdot)$ and Φ have one-jump-continuous Fréchet differentials of first and second order on $\widehat{\mathcal{C}}$ and the smoothing sequence of B does not depend on r.

REMARK 5.3. Assumption 5.2 implies that the same set of properties holds if we substitute $\begin{pmatrix} 1 \\ \mathbb{1}_{[a,0]} \end{pmatrix}$ with any element $q = \begin{pmatrix} \psi(0) \\ \psi \end{pmatrix} \in \mathcal{D}_{-a}$, that is, it has at most one jump and no other discontinuities; this happens by linearity, because any such ψ is the sum of a continuous function and an indicator function.

We state and prove now the main result in this work.

THEOREM 5.4. *Let* $\Phi \in C_b^{2,\alpha}(\mathcal{D}, \mathbb{R})$ *be given and let Assumption* 2.1 *hold for* $E = \mathcal{D}$. *Under Assumption* 5.2, *the function* $u : [0, T] \times \mathcal{D} \to \mathbb{R}$ *given by*

$$(35) \qquad u(t, y) = \mathbb{E}\big[\Phi\big(Y^{t,y}(T)\big)\big],$$

where $Y^{t,y}$ *is the solution to equation* (14' bis) *in* $\mathcal{D}$, *is a classical solution of the Kolmogorov equation with terminal condition* Φ, *that is, for every* $(t, y) \in [0, T] \times \mathrm{Dom}(A_{\widehat{\mathcal{C}}})$ *it satisfies identity*

$$
\begin{aligned}
(26) \qquad u(t, y) - \Phi(y) &= \int_t^T \langle Du(s, y), Ay + B(s, y)\rangle \, ds \\
&\quad + \frac{1}{2} \int_t^T \sum_{j=1}^d \sigma_j^2 D^2 u(s, y)(e_j, e_j) \, ds.
\end{aligned}
$$

PROOF. Throughout this proof, $\|\cdot\|$ will denote the norm of $\mathcal{D}$. Sometimes we will write $\|y\|_{\widehat{\mathcal{C}}}$ to stress the fact that y belongs to $\widehat{\mathcal{C}}$. The duality $\langle \cdot, \cdot \rangle$ will be always intended between $\mathcal{D}'$ and $\mathcal{D}$. We suppose here for simplicity that we can choose the same sequence $\{J_n\}$ for B and Φ in Assumption 5.2; this does not turn in a loss of generality and the proof can be carried on in the same way also when the two smoothing sequences are different.

Using that smoothing sequence define B_n, Φ_n, Y_n and u_n as above. The proof will be divided into some steps that will prove the following: for $y \in \text{Dom}(A_{\widehat{\mathcal{C}}})$:

$\diamond$ $Y_n^{s,y}(t) \to Y^{s,y}(t)$ in $\widehat{\mathcal{C}}$ for every t uniformly in ω;
$\diamond$ $u_n(s, y) \to u(s, y) = \mathbb{E}[\Phi(Y^{s,y}(T))]$ for every s pointwise in y;
$\diamond$ equation (32) converges to equation (26) for any $t \in [0, T]$.

Step 1. Fix $\omega \in \Omega_0$. We first need to compute

$$\|Y_n^{s,y}(t) - Y^{s,y}(t)\|_{\widehat{\mathcal{C}}}$$

$$= \left\| \int_s^t e^{(t-r)A} B_n(r, Y_n^{s,y}(r)) \, dr - \int_s^t e^{(t-r)A} B(r, Y^{s,y}(r)) \, dr \right\|_{\widehat{\mathcal{C}}}$$

$$(36) \qquad \leq \left\| \int_s^t e^{(t-r)A} B_n(r, Y^{s,y}(r)) \, dr - \int_s^t e^{(t-r)A} B(r, Y^{s,y}(r)) \, dr \right\|_{\widehat{\mathcal{C}}}$$

$$(37) \qquad + \left\| \int_s^t e^{(t-r)A} B_n(r, Y_n^{s,y}(r)) \, dr - \int_s^t e^{(t-r)A} B_n(r, Y^{s,y}(r)) \, dr \right\|_{\widehat{\mathcal{C}}}.$$

For the term (36), recall that

$$e^{(t-r)A} B_n(r, Y^{s,y}(r)) = e^{(t-r)A} B(r, J_n Y^{s,y}(r))$$

and that, thanks to the properties of J_n,

$$J_n Y^{s,y}(r) \xrightarrow{n \to \infty} Y^{s,y}(r)$$

in $\widehat{\mathcal{C}}$, hence by continuity of B

$$(38) \qquad\qquad B(r, J_n Y^{s,y}(r)) \longrightarrow B(r, Y^{s,y}(r))$$

pointwise as functions of r. Since B is uniformly bounded in $r \in [s, t]$, by the dominated convergence theorem

$$\lim_{n \to \infty} \int_s^t e^{(t-r)A} B_n(r, Y^{s,y}(r)) \, dr = \int_s^t e^{(t-r)A} B(r, Y^{s,y}(r)) \, dr;$$

hence for any $\varepsilon > 0$

$$(39) \qquad \left\| \int_s^t e^{(t-r)A} B_n(r, Y^{s,y}(r)) \, dr - \int_s^t e^{(t-r)A} B(r, Y^{s,y}(r)) \, dr \right\|_{\widehat{\mathcal{C}}} < \varepsilon$$

for n big enough. Consider now (37):

$$\left\| \int_s^t e^{(t-r)A} B_n\big(r, Y_n^{s,y}(r)\big) \, dr - \int_s^t e^{(t-r)A} B_n\big(r, Y^{s,y}(r)\big) \, dr \right\|_{\widehat{\mathcal{C}}}$$

$$\leq C \int_s^t \big\| B\big(r, J_n Y_n^{s,y}(r)\big) - B\big(r, J_n Y^{s,y}(r)\big) \big\| \, dr$$

$$\leq C \int_s^t K_B \big\| Y_n^{s,y}(r) - Y^{s,y}(r) \big\| \, dr$$

because, for any $\psi \in C([-T, 0]; \mathbb{R}^d)$, $\|J_n \psi\|_\infty \leq C_J \|\psi\|_\infty$ and, therefore, $\|J_n y\| \leq C_J \|y\|$. Hence, this and (39) yield, by Gronwall's lemma,

$$\big\| Y_n^{s,y}(t) - Y^{s,y}(t) \big\|_{\widehat{\mathcal{C}}} \leq \varepsilon e^{TCK_B}$$

for any $\varepsilon > 0$ and n big enough. This implies that $Y_n^{s,y}(t)$ converges to $Y^{s,y}(t)$ in $\widehat{\mathcal{C}}$ for any t.

Step 2. It is now easy to deduce that $u_n(s, y)$ converges to $u(s, y)$ for any s, $y \in \widehat{\mathcal{C}}$. In fact,

$$\big| u_n(s, y) - u(s, y) \big|$$
$$\leq \mathbb{E} \big| \Phi_n\big(Y_n^{s,y}(T)\big) - \Phi_n\big(Y^{s,y}(T)\big) \big| + \mathbb{E} \big| \Phi_n\big(Y^{s,y}(T)\big) - \Phi\big(Y^{s,y}(T)\big) \big|$$

and for almost every ω

$$\big| \Phi_n\big(Y_n^{s,y}(T)\big) - \Phi_n\big(Y^{s,y}(T)\big) \big| \leq K_\Phi \big\| Y_n^{s,y}(T) - Y^{s,y}(Y) \big\|$$

and

$$\big| \Phi_n\big(Y^{s,y}(T)\big) - \Phi\big(Y^{s,y}(T)\big) \big| \leq K_\Phi \big\| J_n Y^{s,y}(T) - Y^{s,y}(T) \big\|,$$

both of which are arbitrarily small for n large enough; now since B is bounded and we assumed that $\mathbb{E}\|Z\|^4$ is finite, we can apply again the dominated convergence theorem (integrating in the variable ω) to conclude this argument.

Step 3. We now approach the convergence of the term

$$\langle Du_n(s, y), B_n(s, y) \rangle;$$

it is enough to consider a generic sequence $\tilde{g}_n \to \tilde{g}$ in $\mathbb{R}^d$, to which we associate the corresponding sequence $g_n = \binom{\tilde{g}_n}{0} \to g = \binom{\tilde{g}}{0}$ in $\mathcal{C} \subset \mathcal{D}$. From (33) and (35), we have that for $h \in \mathcal{D}$

(40) $$\langle Du_n(s, y), h \rangle = \mathbb{E}\big[\langle D\Phi_n\big(Y_n^{s,y}(T)\big), DY_n^{s,y}(T)h \rangle\big]$$

and

(41) $$\langle Du(s, y), h \rangle = \mathbb{E}\big[\langle D\Phi_n\big(Y^{s,y}(T)\big), DY^{s,y}(T)h \rangle\big].$$

We remark here that the duality $_{\mathcal{D}'}\langle Du_n(s, y), g_n\rangle_{\mathcal{D}}$ is well defined and equals $_{\mathcal{L}^{p'}}\langle Du_n(s, y), g_n\rangle_{\mathcal{L}^p}$; a simple proof of this fact is the following: u_n is Fréchet differentiable both on $\mathcal{D}$ and on $\mathcal{L}^p$ and its Gâteaux derivatives along the directions in $\mathcal{D}$ are of course the same in $\mathcal{D}$ and in $\mathcal{L}^p$, therefore, also the Fréchet derivatives must be equal. Now

$$
\begin{aligned}
&\left|\langle Du_n, g_n\rangle - \langle Du, g\rangle\right| \\
&\quad = \left|\langle Du_n, g_n - g\rangle + \langle Du_n - Du, g\rangle\right| \\
&\quad \leq \left|\langle Du_n - Du, g\rangle\right| + \left|\langle Du_n, g_n - g\rangle\right| \\
&\quad \leq \mathbb{E}\left|\langle D\Phi_n(Y_n^{s,y}(T)), DY_n^{s,y}(T)g\rangle - \langle D\Phi(Y^{s,y}(T)), DY^{s,y}(T)g\rangle\right| \\
&\qquad + \mathbb{E}\left|\langle D\Phi_n(Y_n^{s,y}(T)), DY_n^{s,y}(T)(g_n - g)\rangle\right| \\
&\quad = \mathbb{E}|A| + \mathbb{E}|B|.
\end{aligned}
$$

We show that this last expression goes to 0 as $n \to \infty$. We start from B. It is easily shown that

$$
D\Phi_n(\hat{y}) = D\Phi(J_n \hat{y})J_n
$$

for any $\hat{y} \in \mathcal{D}$. $D\Phi$ is bounded by assumption, whereas by the required properties of J_n

$$
\|J_n DY_n^{s,y}(T)c\| \leq C_J \|DY_n^{s,y}(T)c\|
$$

for any $c \in \mathcal{C}$. Since the $\|DY_n\|$'s are uniformly bounded by a constant depending only on e^{tA} and on DB (see the proof of Theorem 2.3 in the Appendix), we have that the Du_n's are uniformly bounded on $\mathcal{C}$ as well and, therefore, $\mathbb{E}|B| \to 0$ as $g_n \to g$.

The term A requires some work: from now on fix $\omega \in \Omega_0$ and write (suppressing indexes s, y, ω and T)

$$
\begin{aligned}
A &= \langle D\Phi_n(Y_n), DY_ng\rangle - \langle D\Phi(Y), DYg\rangle \\
&= \langle D\Phi_n(Y_n), (DY_n - DY)g\rangle + \langle D\Phi_n(Y_n) - D\Phi(Y), DYg\rangle = A_1 + A_2, \\
A_2 &= \langle D\Phi_n(Y_n) - D\Phi_n(Y), DYg\rangle + \langle D\Phi_n(Y) - D\Phi(Y), DYg\rangle = A_{21} + A_{22}.
\end{aligned}
$$

Since the Lipschitz constants of $D\Phi_n$ are uniformly bounded in $\widehat{\mathcal{C}}$, we have that

$$
\begin{aligned}
|A_{21}| &\leq \|D\Phi_n(Y_n) - D\Phi_n(Y)\|_{\mathcal{D}'} \|DYg\|_{\mathcal{D}} \\
&\leq K_1 \|Y_n - Y\| \|DYg\|
\end{aligned}
$$

and the last line goes to zero as n goes to infinity. For A_{22}, write

$$
\begin{aligned}
|A_{22}| &= \left|\langle D\Phi(J_nY)J_n, DYg\rangle - \langle D\Phi(Y), DYg\rangle\right| \\
&\leq \left|\langle D\Phi(J_nY)J_n, DYg\rangle - \langle D\Phi(Y)J_n, DYg\rangle\right| \\
&\quad + \left|\langle D\Phi(Y)J_n, DYg\rangle - \langle D\Phi(Y), DYg\rangle\right| \\
&\leq K_{D\Phi}\|J_nY - Y\| \|DYg\| + \left|\langle D\Phi(Y)J_n, DYg\rangle - \langle D\Phi(Y), DYg\rangle\right|;
\end{aligned}
$$

the first term goes to zero by properties of J_n, the second one thanks to Assumption 5.2: this is because from the defining equation for DY one easily sees that for any $\binom{g}{0} \in \mathcal{C}$ the second component of DYg has a unique discontinuity point, and our assumption is made exactly in order to be able to control the convergence of these terms. Now we consider A_1:

$$
\begin{aligned}
DY_n^{s,y}(T)g &- DY^{s,y}(T)g \\
&= \int_s^T e^{(T-r)A} DB_n\big(r, Y_n^{s,y}(r)\big)\big[DY_n^{s,y}(r) - DY^{s,y}(r)\big]g\,dr \\
&\quad + \int_s^T e^{(T-r)A}\big[DB_n\big(r, Y_n^{s,y}(r)\big) - DB\big(r, Y^{s,y}(r)\big)\big]DY^{s,y}(r)g\,dr \\
&= A_{11} + A_{12}
\end{aligned}
\tag{42}
$$

and A_{12} can be written as

$$
\begin{aligned}
A_{12} &= \int_s^T e^{(T-r)A}\big[DB_n\big(r, Y_n^{s,y}(r)\big) - DB_n\big(r, Y^{s,y}(r)\big)\big]DY^{s,y}(r)g\,dr \\
&\quad + \int_s^T e^{(T-r)A}\big[DB_n\big(r, Y^{s,y}(r)\big) - DB\big(r, Y^{s,y}(r)\big)\big]DY^{s,y}(r)g\,dr \\
&= A_{121} + A_{122}
\end{aligned}
$$

whence

$$
\begin{aligned}
\|A_{121}\| &\le C \int_s^T \big\|DY^{s,y}(r)g\big\|\big\|DB\big(r, J_n Y_n^{s,y}(r)\big) - DB\big(r, J_n Y^{s,y}(r)\big)\big\|\,dr \\
&\le C \int_s^T \big\|DY^{s,y}(r)g\big\|\big\|D^2 B(r, \cdot)\big\|\big\|J_n Y_n^{s,y}(r) - J_n Y^{s,y}(r)\big\| \\
&\le C \cdot C_J \int_s^T \big\|DY^{s,y}(r)g\big\|\big\|D^2 B(r, \cdot)\big\|\big\|Y_n^{s,y}(r) - Y^{s,y}(r)\big\|\,dr
\end{aligned}
$$

that goes to zero; for A_{122}

$$
\begin{aligned}
\big\|\big[DB_n\big(r, Y^{s,y}(r)\big) &- DB\big(r, Y^{s,y}(r)\big)\big]DYg\big\| \\
&\le \big\|DB\big(r, J_n Y^{s,y}(r)\big) - DB\big(r, Y^{s,y}(r)\big)\big\|\big\|J_n DY^{s,y}(r)g\big\| \\
&\quad + \big\|DB\big(r, Y^{s,y}(r)\big)\big[J_n DY^{s,y}(r)g - DY^{s,y}(r)g\big]\big\| \\
&\le K_{DB}\big\|J_n Y^{s,y}(r) - Y^{s,y}(r)\big\|\big\|DY^{s,y}(r)g\big\| \\
&\quad + \big\|DB\big(r, Y^{s,y}(r)\big)\big[J_n DY^{s,y}(r)g - DY^{s,y}(r)g\big]\big\|,
\end{aligned}
$$

where the last line goes to zero thanks to Assumption 5.2 again, and therefore A_{122} goes to zero by the dominated convergence theorem. From (42) and this last

argument it follows that for any fixed $\varepsilon > 0$

$$
(43) \quad
\begin{aligned}
&\|DY_n(T)g - DY(T)g\| \\
&\qquad \leq C \int_s^T \|DB_n\| \|DY_n^{s,y}(r)g - DY^{s,y}(r)g\| \, dr + \varepsilon
\end{aligned}
$$

for n large enough. Since $\|DB_n\|$ is bounded uniformly in n and in r we can use Gronwall's lemma to prove that $\|DY_n^{s,y}(T)g - DY^{s,y}(T)g\| \to 0$, and since $\|D\Phi_n\|$ are uniformly bounded as well we can conclude that also $A_1 \to 0$ as $n \to \infty$. Putting together all the pieces, we just examined we obtain the desired convergence of $\langle Du_n, B_n \rangle$ to $\langle Du, B \rangle$ thanks to the dominated convergence theorem (in the variable ω).

Step 4. All the procedures used in the previous steps apply again to treat the convergence of the term

$$\langle Du_n(s, y), Ay \rangle,$$

no further passages are needed; therefore, we omit the computations and go on to the term involving the second derivatives.

Step 5. We will study only the convergence of

$$D^2 u_n(s, y)(e_1, e_1)$$

since the σ_j's are constants and the passage from one to d dimensions is trivial. We will drop the subscript 1 in the computations to simplify notation. We can proceed as follows (suppressing again s, y, ω and T):

$$
\begin{aligned}
&\left| D^2 u_n(s, y)(e, e) - D^2 u(s, y)(e, e) \right| \\
&\qquad \leq \mathbb{E} \left| D^2 \Phi_n(Y_n)(DY_n e, DY_n e) - D^2 \Phi(Y)(DY e, DY e) \right| \\
&\qquad\quad + \mathbb{E} \left| \langle D\Phi_n(Y_n), D^2 Y_n(e, e) \rangle - \langle D\Phi(Y), D^2 Y(e, e) \rangle \right| \\
&\qquad = \mathbb{E}|C| + \mathbb{E}|D|.
\end{aligned}
$$

The kind of computations needed are similar to those for the terms involving the first derivative. We first write C (for ω fixed) as

$$
\begin{aligned}
C &= \left[D^2 \Phi_n(Y_n)(DY_n e, DY_n e) - D^2 \Phi_n(Y_n)(DY e, DY e) \right] \\
&\quad + \left[D^2 \Phi_n(Y_n)(DY e, DY e) - D^2 \Phi(Y)(DY e, DY e) \right] \\
&= C_1 + C_2.
\end{aligned}
$$

For C_1, just write

$$
\begin{aligned}
|C_1| &\leq \big| D^2 \Phi_n(Y_n)(DY_n e - DY e, DY_n e - DY e) \\
&\qquad + D^2 \Phi_n(Y_n)(DY e, DY_n e - DY e) + D^2 \Phi_n(Y_n)(DY_n e - DY e, DY e) \big| \\
&\leq \|D^2 \Phi_n(Y_n)\| \left[\|DY_n e - DY e\|^2 + 2\|DY e\| \|DY_n e - DY e\| \right]
\end{aligned}
$$

and the last line goes to zero by the same reasoning as in A_1 and the boundedness of $\|D^2\Phi_n(Y_n)\|$ (uniformly in n).

Write C_2 as

$$
\begin{aligned}
C_2 &= D^2\Phi_n(Y_n)(DYe, DYe) - D^2\Phi(Y)(DYe, DYe) \\
&= D^2\Phi(J_nY_n)(J_nDYe, J_nDYe) - D^2\Phi(Y)(DYe, DYe) \\
&= \left[D^2\Phi(J_nY_n)(J_nDYe, J_nDYe) - D^2\Phi(J_nY)(J_nDYe, J_nDYe) \right] \\
&\quad + \left[D^2\Phi(J_nY)(J_nDYe, J_nDYe) - D^2\Phi(Y)(DYe, DYe) \right] \\
&= C_{21} + C_{22}.
\end{aligned}
$$

Now

$$
C_{21} = \left[D^2\Phi(J_nY_n) - D^2\Phi(J_nY) \right](J_nDYe, J_nDYe)
$$

hence

(44)
$$
\begin{aligned}
\|C_{21}\| &\leq \|J_nDYe\|^2 \|D^2\Phi\|_\alpha \|J_nY_n - J_nY\| \\
&\leq C_J^2 \|DYe\| \|D^2\Phi\|_\alpha C_J \|Y_n - Y\|
\end{aligned}
$$

[here $\|D^2\Phi\|_\alpha$ is the α-Hölder norm of $D^2\Phi$ as a map from $\mathcal{D}$ into the set of bilinear forms $L(\mathcal{D}, \mathcal{D}; \mathcal{D})$] which converges to zero thanks to the first step of the proof. For C_{22}, we can write

$$
\begin{aligned}
C_{22} &= \left[D^2\Phi(J_nY) - D^2\Phi(Y) \right](J_nDYe, J_nDYe) \\
&\quad + D^2\Phi(Y)(J_nDYe, J_nDYe) - D^2\Phi(Y)(DYe, DYe) \\
&= \left[D^2\Phi(J_nY) - D^2\Phi(Y) \right](J_nDYe, J_nDYe) \\
&\quad + D^2\Phi(Y)(J_nDYe, J_nDYe - DYe) \\
&\quad + D^2\Phi(Y)(J_nDYe - DYe, DYe) \\
&= \left[D^2\Phi(J_nY) - D^2\Phi(Y) \right](J_nDYe, J_nDYe) \\
&\quad + D^2\Phi(Y)(J_nDYe - DYe, J_nDYe - DYe) \\
&\quad + D^2\Phi(Y)(DYe, J_nDYe - DYe) \\
&\quad + D^2\Phi(Y)(J_nDYe - DYe, J_nDYe - DYe).
\end{aligned}
$$

Last three terms go to zero by Assumption 5.2, while the first one is bounded in norm by

$$
C_J \|D^2\Phi\|_\alpha \|J_nY - Y\|^\alpha \|DYe\|^2
$$

which goes to zero since $\|J_nY - Y\| \to 0$.

We now go on with D:

$$\mathrm{D} = \langle D\Phi_n(Y_n), D^2 Y_n(e,e) - D^2 Y(e,e)\rangle + \langle D\Phi_n(Y_n) - D\Phi(Y), D^2 Y(e,e)\rangle$$
$$= \mathrm{D}_1 + \mathrm{D}_2$$

and D_2 is easy to handle since

$$|\mathrm{D}_2| \le |\langle D\Phi_n(Y_n) - D\Phi_n(Y), D^2 Y(e,e)\rangle| + |\langle D\Phi_n(Y) - D\Phi(Y), D^2 Y(e,e)\rangle|,$$

where the first term is bounded by

$$\|D^2\Phi_n\|\|Y_n - Y\|\|D^2 Y(e,e)\|$$

and, therefore, goes to zero as for A_1, and the second goes to zero since $D^2 Y(e,e)$ is in $\widehat{\mathcal{C}}$ and $D\Phi_n(y)$ converge to $D\Phi(y)$ for any y as functionals on $\widehat{\mathcal{C}}$. Let us now rewrite the right-hand term in the bracket defining D_1 as

$$D^2 Y_n^{s,y}(T)(e,e) - D^2 Y^{s,y}(T)(e,e)$$

(45)
$$= \int_s^T e^{(T-r)A}\big[D^2 B_n\big(r, Y_n^{s,y}(r)\big)\big(DY_n^{s,y}(r)e, DY_n^{s,y}(r)e\big)$$
$$- D^2 B\big(r, Y^{s,y}(r)\big)\big(DY^{s,y}(r)e, DY^{s,y}(r)e\big)\big]\,dr$$
$$+ \int_s^T e^{(T-r)A}\big[DB_n\big(r, Y_n^{s,y}(r)\big)D^2 Y_n^{s,y}(r)(e,e)$$
$$- DB\big(r, Y^{s,y}(r)\big)D^2 Y^{s,y}(r)(e,e)\big]\,dr$$
$$= \mathrm{D}_{11} + \mathrm{D}_{12}.$$

Proceeding in a way similar to before we write the integrand in D_{11} as a sum (suppressing also the variable r)

$$\mathrm{D}_{11} = \big[D^2 B_n(Y_n)(DY_n e, DY_n e) - D^2 B_n(Y_n)(DY e, DY e)\big]$$
$$+ \big[D^2 B_n(Y_n) - D^2 B(Y)\big](DY e, DY e)$$
$$= \mathrm{D}_{111} + \mathrm{D}_{112}$$

and notice that

$$\|\mathrm{D}_{111}\| = \big\| D^2 B_n(Y_n)(DY_n e - DY e, DY_n e - DY e)$$
$$+ D^2 B_n(Y_n)(DY_n e - DY e, DY e) + D^2 B_n(Y_n)(DY e, DY_n e - DY e)\big\|$$
$$\le \|D^2 B_n(Y_n)\|\big[\|DY_n e - DY e\|^2 + 2\|DY e\|\|DY_n e - DY e\|\big]$$

which can be treated as in A_1 since the norms $\|D^2 B_n(r, Y_n^{s,y}(r))\|$ are bounded uniformly in n and r. D_{112} can be treated as we did for C_2, obtaining

$$\mathrm{D}_{112} = \big[D^2 B(J_n Y_n)(J_n DY e, J_n DY e) - D^2 B(J_n Y)(J_n DY e, J_n DY e)\big]$$
$$+ \big[D^2 B(J_n Y)(J_n DY e, J_n DY e) - D^2 B(Y)(DY e, DY e)\big]$$
$$= \mathrm{D}_{1121} + \mathrm{D}_{1122};$$

an estimate analogous to (44) shows how to control the term D_{1121}, while

$$
\begin{aligned}
D_{1122} = &[D^2B(J_nY) - D^2B(Y)](J_nDYe, J_nDYe) \\
&+ D^2B(Y)(J_nDYe - DYe, J_nDYe - DYe) \\
&+ D^2B(Y)(DYe, J_nDYe - DYe) \\
&+ D^2B(Y)(J_nDYe - DYe, J_nDYe - DYe)
\end{aligned}
$$

and these last quantities are shown to go to zero pointwise in r thanks to Assumption 5.2 and to the α-Hölderianity of D^2B_n in the same way as for C_{22}. By dominated convergence, D_{11} is thus shown to converge to 0. To finish studying D_1 (hence D), we need to rewrite the integrand in D_{12} as

$$
\begin{aligned}
&DB_n(Y_n)D^2Y_n(e,e) - DB(Y)D^2Y(e,e) \\
&= DB_n(Y_n)[D^2Y_n - D^2Y](e,e) + [DB_n(Y_n) - DB_n(Y)]D^2Y(e,e) \\
&\quad + [DB_n(Y) - DB(Y)]D^2Y(e,e) \\
&= DB_n(Y_n)[D^2Y_n - D^2Y](e,e) + [DB_n(Y_n) - DB_n(Y)]D^2Y(e,e) \\
&\quad + DB(J_nY)[J_nD^2Y(e,e) - D^2Y(e,e)] \\
&\quad + [DB(J_nY) - DB(Y)]D^2Y(e,e).
\end{aligned}
$$

The second term in this last sum is bounded in norm by

$$
\|D^2B_n(r,\cdot)\| \|Y_n - Y\| \|D^2Y(e,e)\|
$$

which goes to zero since $Y_n \to Y$ and $\|DB_n\|$ are uniformly bounded (as already noticed before); the norm of the third term goes to zero because it is bounded by

$$
\|DB(J_nY)\| \|J_nD^2Y(e,e) - D^2Y(e,e)\|;
$$

the norm of last term goes to zero as well by the Lipschitz property of DB. Taking into account all these observations and the fact that D_{11} has already been shown to converge to zero, we can use Gronwall's lemma in (45) to obtain that

$$
D^2Y_n^{s,y}(T)(e,e) - D^2Y^{s,y}(T)(e,e) \to 0.
$$

This together with the uniform boundedness of $D\Phi_n(Y_n)$ finally yields the convergence to zero of D, hence that of the second-order term.

At last, an application of the dominated convergence theorem with respect to the variable s in all integral terms appearing in the Kolmogorov equation completes the proof. $\square$

REMARK 5.5. Since u is given as an integral of functions which are bounded in the variable t, it is a Lipschitz function, hence differentiable almost everywhere

thanks to a classic result by Rademacher. Therefore, a posteriori it satisfies the differential form of Kolmogorov equation

$$(46) \quad \begin{cases} \dfrac{\partial u}{\partial t}(t, y) + \langle Du(t, y), Ay + B(t, y) \rangle + \dfrac{1}{2} \sum_{j=1}^{d} \sigma_j^2 D^2 u(t, y)(e_j, e_j) = 0, \\ u(T, \cdot) = \Phi \end{cases}$$

for almost every $t \in [0, T]$.

6. Examples. We give here some examples, recalling also those mentioned at the beginning of the paper, to which the theory exposed so far can be applied. In particular, we first consider integral functions and explain why they cannot be treated in the standard Hilbert space setting for our purposes, and then show that the technical Assumption 5.2, which can seem very restrictive when considered in its abstract form, is indeed satisfied by all the usual examples.

6.1. *Examples for the $\mathcal{L}^p$ theory.*

EXAMPLE 6.1 (A negative example). First, we show that, as said before, even the simplest path-dependent functions one can think of, namely integral functional, do not have enough smoothness when considered in the standard $\mathcal{L}^2$ setting.

In dimension $d = 1$, consider the integral functional

$$b_t(\gamma_t) = \int_0^t g(\gamma(s)) \, ds,$$

where $g : \mathbb{R} \to \mathbb{R}$ is a C_b^3 function. Its infinite-dimensional lifting is given by

$$B\left(t, \begin{pmatrix} x \\ \varphi \end{pmatrix}\right) = \begin{pmatrix} \hat{b}\left(t, \begin{pmatrix} x \\ \varphi \end{pmatrix}\right) \\ 0 \end{pmatrix},$$

where

$$\hat{b}\left(t, \begin{pmatrix} x \\ \varphi \end{pmatrix}\right) = \int_0^t g(\varphi(s - t)) \, ds.$$

The second Gâteaux derivative of B with respect to $y = \begin{pmatrix} x \\ \varphi \end{pmatrix}$ is simply

$$D_G^2 B(t, y)\left(\begin{pmatrix} x_1 \\ \psi \end{pmatrix}, \begin{pmatrix} x_2 \\ \chi \end{pmatrix}\right) = \begin{pmatrix} \int_0^t g''(\varphi(s - t)) \psi(s - t) \chi(s - t) \, ds \\ 0 \end{pmatrix}.$$

Given $\begin{pmatrix} x_1 \\ \psi \end{pmatrix}$, $\begin{pmatrix} x_2 \\ \chi \end{pmatrix}$, it is easy to check, by Lebesgue theorem, that this Gâteaux derivative is continuous in y in the $\mathcal{L}^2$ topology; with some additional effort it can be also shown that it is uniformly continuous, in $y \in \mathcal{L}^2$. Presumably, thanks to this result on B, with due effort it can be shown that uniform continuity of Gâteaux derivatives holds true also for the solution Y of the SDE and then for $u(t, y)$. However,

with only such knowledge about the space regularity of u, we do not know how to prove that u satisfies the Kolmogorov equation (we do not know how to control the remainders in Taylor developments). Coherently, with the present literature on the subject, we are able to complete the proof that $u(t, y)$ fulfills the Kolmogorov equation only when the second-order Fréchet differential is uniformly continuous [not only the Gâteaux derivative for given $\binom{x_1}{\psi}$, $\binom{x_2}{\chi}$]. This is false for B as above: integral functionals are not even twice differentiable in Fréchet sense in general. In order for $D_G^2 B(t, y)$ to be the second-order Fréchet differential of B we would need that

$$\lim_{\|w\|_{\mathcal{L}^2} \to 0} \frac{1}{\|w\|_{\mathcal{L}^2}} \left\| DB(t, y+w)z - DB(t, y)z - D_G^2 B(t, y)(z, w) \right\|_{\mathcal{L}^2} = 0$$

uniformly in $z \in \mathcal{L}^2$, that is, for $y = \binom{z}{\varphi}$, $z = \binom{x_1}{\psi}$, $w = \binom{x_2}{\chi}$,

$$\lim_{\|\chi\|_{L^2} \to 0} \frac{1}{\|\chi\|_{L^2}} \left| \int_0^t \left[g'(\varphi(s-t) + \chi(s-t)) - g'(\varphi(s-t)) \right] \psi(s-t) \, ds \right.$$
$$\left. - \int_0^t g''(\varphi(s-t)) \psi(s-t) \chi(s-t) \, ds \right| = 0$$

uniformly in $\psi \in L^2$. Suppose that g'' is not constant, take as φ any continuous function and choose $\psi(s) = s^{-1/3}$ and $\chi_n(s) = s^{-1/3} \mathbb{1}_{[-1/n,0)}(s)$. Then $\chi_n \to 0$ in L^2 as $n \to \infty$ and

$$\lim_{n \to \infty} \frac{1}{\|\chi_n\|_{L^2}} \left| \int_0^t \left[g'(\varphi(s-t) + \chi_n(s-t)) - g'(\varphi(s-t)) \right] \psi(s-t) \, ds \right.$$
$$\left. - \int_0^t g''(\varphi(s-t)) \psi(s-t) \chi_n(s-t) \, ds \right|$$
$$= \lim_{n \to \infty} \frac{1}{\|\chi_n\|_{L^2}} \left| \int_0^t \left[g''(\varphi(s-t)) \chi_n(s-t) \psi(s-t) \right.\right.$$
$$\left.\left. + \frac{1}{2} g'''(\bar{x}) \chi_n(s-t)^2 \psi(s-t) \right] ds \right.$$
$$\left. - \int_0^t g''(\varphi(s-t)) \psi(s-t) \chi_n(s-t) \, ds \right|,$$

where $\bar{x}$ is some point in $\mathbb{R}$. Since g''' is bounded, we have to compute

$$\lim_{n \to \infty} \frac{1}{\|\chi_n\|_{L^2}} \int_0^t |\chi_n(s-t)|^2 |\psi(s-t)| \, ds,$$

but with our choice of χ_n and ψ the functions $|\chi_n|^2 |\psi|$ are not integrable for any n. Therefore, $D_G^2 B(t, y)$ cannot be the differential of second order of B in Fréchet sense.

EXAMPLE 6.2. On the other hand, the infinite-dimensional lifting of integral functionals of the form

$$b_t(\gamma_t) = \int_0^t g(\gamma(t), \gamma(s)) \, ds$$

with g of class $C_b^{2,\alpha}(\mathbb{R}^d \times \mathbb{R}^d; \mathbb{R})$ satisfy the assumptions of Theorem 4.1 for $p = 2 + \alpha$; in particular they are twice Fréchet differentiable with α-Hölder continuous (hence uniformly continuous) second Fréchet differential in $\mathcal{L}^p$ for $p = 2 + \alpha$. Indeed, for $y = \binom{x}{\varphi}$,

$$B(t, y) = \begin{pmatrix} \int_0^t g(x, \varphi(s - t)) \, ds \\ 0 \end{pmatrix},$$

$$D^2 B(t, y)\left(\binom{x_1}{\psi}, \binom{x_2}{\chi}\right) = \binom{a}{0},$$

where (denoting by ∂_1 and ∂_2 the partial derivatives of g in its two arguments)

$$a = \int_0^t \partial_1^2 g(x, \varphi(s - t)) \, ds + \int_0^t \partial_2^2 g(\varphi(s - t)) \psi(s - t) \chi(s - t) \, ds$$

$$+ \int_0^t \partial_1 \partial_2 g(x, \varphi(s - t))(\psi(s - t) + \chi(s - t)) \, ds.$$

For $z = \binom{x_1}{\varphi_1}$, we have to estimate $\|D^2 B(t, y) - D^2 B(t, z)\|_{L(\mathcal{L}^p, \mathcal{L}^p; \mathcal{L}^p)}$ and the most difficult term is

$$\left| \int_0^t \left(\partial_2^2 g(\varphi(s - t)) - \partial_2^2 g(\varphi_1(s - t)) \right) \psi(s - t) \chi(s - t) \, ds \right|$$

$$\leq \|\partial_2^2 g\|_\alpha \int_0^t |\varphi(s - t) - \varphi_1(s - t)|^\alpha |\psi(s - t)| |\chi(s - t)| \, ds$$

$$\leq \|\partial_2^2 g\|_\alpha \||\varphi - \varphi_1|^\alpha\|_{L^{p/\alpha}} \|\psi\|_{L^p} \|\chi\|_{L^p}$$

$$= \|\partial_2^2 g\|_\alpha \|\varphi - \varphi_1\|_{L^p}^\alpha \|\psi\|_{L^p} \|\chi\|_{L^p}$$

which implies

$$\sup_{\substack{\chi, \psi \in \mathcal{L}^p \\ \|\chi\|_{L^p}, \|\psi\|_{L^p} \leq 1}} \left| \int_0^t \left(\partial_2^2 g(\varphi(s - t)) - \partial_2^2 g(\varphi_1(s - t)) \right) \psi(s - t) \chi(s - t) \, ds \right|$$

$$\leq \|\partial_2^2 g\|_{C^\alpha} \|\varphi - \varphi_1\|_{L^p}^\alpha.$$

Since g and its derivatives are bounded, Assumption 2.1 is easily seen to be satisfied.

This argument can be easily extended to include dependence on t and s in g, as in example (i) in the Introduction.

6.2. *Examples for the theory in $\mathcal{D}$.*

EXAMPLE 6.3. We show now that the lifting of the function introduced in Section 1.2, example (ii) satisfies the assumptions of Theorem 5.4. For simplicity, we evaluate any càdlàg curve γ only in two fixed points t_1 and t_2, $0 \le t_1 \le t_2 < T$, that is, we set

$$b_t(\gamma_t) = h_1\big(\gamma(t_1)\big)\mathbb{1}_{[t_1,t_2)}(t) + h_2\big(\gamma(t_1), \gamma(t_2)\big)\mathbb{1}_{[t_2,T]}(t),$$

where $h_1 : \mathbb{R}^d \to \mathbb{R}^d$ and $h_2 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ are in $C_b^{2,\alpha}$ on their respective domains.

Given an element $\binom{x}{\varphi} \in \mathcal{D}$, we will write $\varphi(0)$ for x to avoid the burdensome notation $\varphi(s)\mathbb{1}_{[-T,0)}(s) + x\mathbb{1}_{\{0\}}(s)$ in the following computations, and we will write $\mathbb{1}_{[a,0]}$ for $\binom{1}{\mathbb{1}_{[a,0)}}$ accordingly.

We first check that Assumption 5.2 is satisfied. Here, $\hat{b}$ is given by

$$\hat{b}_t(t, x, \varphi) = h_1\big(\varphi(t_1 - t)\big)\mathbb{1}_{[t_1,t_2)}(t) + h_2\big(\varphi(t_1 - t), \varphi(t_2 - t)\big)\mathbb{1}_{[t_2,T]}(t).$$

Therefore, the Fréchet differential of B with respect to its second argument $\binom{x}{\varphi}$ is given by

$$DB\left(t, \binom{x}{\varphi}\right)\binom{x_1}{\psi} = \begin{pmatrix} D\hat{b}\left(t, \binom{x}{\varphi}\right)\binom{x_1}{\psi} \\ 0 \end{pmatrix},$$

where

$$\begin{aligned}
D\hat{b}&\left(t, \binom{x}{\varphi}\right)\binom{x_1}{\psi} \\
&= Dh_1\big(\varphi(t_1 - t)\big)\psi(t_1 - t)\mathbb{1}_{[t_1,t_2)}(t) \\
&\quad + Dh_2\big(\varphi(t_1 - t), \varphi(t_2 - t)\big)\big(\psi(t_1 - t), \psi(t_2 - t)\big)\mathbb{1}_{[t_2,T]}(t)
\end{aligned}$$

and Dh_j denotes the Jacobian matrix of h_j.

For any fixed $a \in [-T, 0]$ (recall the convention we adopted in Definition 5.1), the first component of $DB(t, \binom{x}{\varphi})J_n\binom{1}{\mathbb{1}_{[a,0)}}$ is given by

$$\begin{aligned}
\big[Dh_1&\big(\varphi(t_1 - t)\big) \cdot J_n\mathbb{1}_{[a,0]}(t_1 - t)\big]\mathbb{1}_{[t_1,t_2)}(t) \\
&+ \big[Dh_2\big(\varphi(t_1 - t), \varphi(t_2 - t)\big) \cdot \big(J_n\mathbb{1}_{[a,0]}(t_1 - t), J_n\mathbb{1}_{[a,0]}(t_2 - t)\big)\big]\mathbb{1}_{[t_2,T]}(t)
\end{aligned}$$

while the second is 0. Therefore,

$$DB\left(t, \binom{x}{\varphi}\right)\left(J_n\binom{1}{\mathbb{1}_{[a,0)}} - \binom{1}{\mathbb{1}_{[a,0)}}\right) \longrightarrow 0$$

if and only if

$$J_n\mathbb{1}_{[a,0]}(t_j - t) \to \mathbb{1}_{[a,0]}(t_j - t),$$

$j = 1, 2$. Fix $j = 1$ (the situation being analogous with $j = 2$); if $t = t_1$, it is straightforward to verify the assumption, therefore, suppose $t \neq t_1$. Then, using the sequence J_n given by (34), if $t_1 > 0$ we have

$$
J_n \mathbb{1}_{[a,0]}(t_1 - t) = \int_{-T}^0 \rho_n\big(\tau_{1/n}(t_1 - t) - y\big)\mathbb{1}_{[a,0]}(y)\,\mathrm{d}y
$$

(47)

$$
= \int_a^0 \rho_n(t_1 - t - y)\,\mathrm{d}y
$$

for n big enough. Now if $t_1 - t < a$ then choosing n large enough we have that $(t_1 - t) + \mathrm{supp}(\rho_n) \cap [a, 0] = \varnothing$, hence the function in (47) equals to 0 definitively as n tends to infinity. Conversely, if $t_1 - t > a$ for n large enough we have that $(t_1 - t) + \mathrm{supp}(\rho_n) \cap [a, 0] = (t_1 - t) + \mathrm{supp}(\rho_n)$ and the function in (47) equals 1 definitively. If $t_1 = 0$, the same procedure applies when $t \neq T$ or $a > -T$, while when $t = T$ and $a = -T$ by the definition of $\tau_{1/n}$ it follows that

$$
J_n \mathbb{1}_{[a,0]}(-T) = \int_{-T}^0 \rho_n\big(\tau_{1/n}(-T) - y\big)\mathbb{1}_{[-T,0)}(y)\,\mathrm{d}y
$$

$$
= \int_{-T}^0 \rho_n\Big(-T + \frac{1}{n} - y\Big)\,\mathrm{d}y = 1.
$$

Therefore, for any $t \in [0, T]$, for any $a \neq t_1 - t$ we have that $J_n \mathbb{1}_{[a,0]}(t_1 - t) = \mathbb{1}_{[a,0]}(t_1 - t)$ definitively as n tends to ∞, as required. It is easy to see that if $a = t_1 - t$ then $J_n \mathbb{1}_{[a,0]}(t_1 - t) \to \frac{1}{2}$.

The second Fréchet differential is given by

$$
D^2 B\left(t, \begin{pmatrix} x \\ \varphi \end{pmatrix}\right)\left(\begin{pmatrix} x_1 \\ \psi \end{pmatrix}, \begin{pmatrix} x_2 \\ \chi \end{pmatrix}\right) = \begin{pmatrix} D^2 \hat{b}\left(t, \begin{pmatrix} x \\ \varphi \end{pmatrix}\right)\left(\begin{pmatrix} x_1 \\ \psi \end{pmatrix}, \begin{pmatrix} x_2 \\ \chi \end{pmatrix}\right) \\ 0 \end{pmatrix},
$$

where

$$
D^2 \hat{b}\left(t, \begin{pmatrix} x \\ \varphi \end{pmatrix}\right)\left(\begin{pmatrix} x_1 \\ \psi \end{pmatrix}, \begin{pmatrix} x_2 \\ \chi \end{pmatrix}\right)
$$

$$
= D^2 h_1\big(\varphi(t_1 - t)\big)\big(\psi(t_1 - t), \chi(t_1 - t)\big)\mathbb{1}_{[t_1, t_2)}(t)
$$

$$
+ D^2 h_2\big(\varphi(t_1 - t), \varphi(t_2 - t)\big)
$$

$$
\times \big((\psi(t_1 - t), \psi(t_2 - t)), (\chi(t_1 - t), \chi(t_2 - t))\big)\mathbb{1}_{[t_2, T]}(t)
$$

and $D^2 h_j$ denotes the Hessian tensor of h_j; it can be easily seen that this differential satisfies the requirements of Assumption 5.2 reasoning as above.

It is also immediate to check that since h_1 and h_2 are in $C_b^{2,\alpha}$ Assumption 2.1 is satisfied by this example.

EXAMPLE 6.4. We can use evaluation at fixed times also in the terminal condition for the path-dependent Kolmogorov equation (4) (see also Section 7): given a smooth function $q : \mathbb{R}^{(n+1)d} \to \mathbb{R}$, bounded with bounded derivatives, consider

$$f(\gamma_T) = q(\gamma(t_0), \gamma(t_1), \ldots, \gamma(t_n), \gamma(T)).$$

Its infinite-dimensional lifting is then given by

$$\Phi\begin{pmatrix} x \\ \varphi \end{pmatrix} = \begin{pmatrix} \hat{f}\begin{pmatrix} x \\ \varphi \end{pmatrix} \\ 0 \end{pmatrix},$$

where

$$\hat{f}\left(\begin{pmatrix} x \\ \varphi \end{pmatrix}\right) = q(\varphi(t_0 - T), \varphi(t_1 - T), \ldots, \varphi(t_n - T), x).$$

From Example 6.3, it is immediate to see that such a Φ satisfies Assumption 5.2 and, therefore, it can be chosen as terminal condition in Theorem 5.4.

EXAMPLE 6.5. From Examples 6.3 and 6.4, it follows also that Theorem 5.4 can be applied when the drift or the terminal condition in the Kolmogorov equation (or both) are delayed functions of the form

$$b_t(\gamma_t) = g(\gamma(t), \gamma(t - \delta))\mathbb{1}_{[\delta, T]}(t), \qquad f(\gamma_T) = q(\gamma(T), \gamma(T - \delta))$$

for g and q sufficiently regular and with values in $\mathbb{R}^d$ and $\mathbb{R}$, respectively, and $0 < \delta < T$, since in this case we have that

$$B\left(t, \begin{pmatrix} x \\ \varphi \end{pmatrix}\right) = \begin{pmatrix} g(x, \varphi(-\delta)) \\ 0 \end{pmatrix}\mathbb{1}_{[\delta, T]}(t) \qquad \forall t \in [0, T]$$

and

$$\Phi\begin{pmatrix} x \\ \varphi \end{pmatrix} = \begin{pmatrix} q(x, \varphi(-\delta)) \\ 0 \end{pmatrix}.$$

REMARK 6.6. The theory exposed here cannot be applied to example (iv) in Section 1.2, that is the functional

$$b_t(\gamma_t) = \sup_{s \in [0,t]} \gamma(s)$$

since the supremum is not Fréchet differentiable as a function of the path.

7. Comparison with path-dependent calculus. We conclude this work establishing some connections between our results and objects and those defined by Dupire and successively developed by Cont and Fournié. We recall here the definitions of the pathwise derivatives given in Cont and Fournié (2013). For a function

$v = \{v_t\}_t$, $v_t \colon D([0, T]; \mathbb{R}^d) \to \mathbb{R}^d$, the ith *vertical derivative* at γ_t ($i = 1, \dots, d$) is defined as

$$(48) \qquad \mathscr{D}_i v_t(\gamma_t) = \lim_{h \to 0} \frac{v_t(\gamma_t^{h e_i}) - v_t(\gamma_t)}{h},$$

where $\gamma_t^{h e_i}(s) = \gamma_t(s) + h e_i \mathbb{1}_{\{t\}}(s)$; we denote the *vertical gradient* at γ_t by

$$\mathscr{D} v_t(\gamma_t) = \big(\mathscr{D}_1 v_t(\gamma_t), \dots, \mathscr{D}_d v_t(\gamma_t)\big);$$

higher order vertical derivatives are defined in a straightforward way. The *horizontal derivative* at γ_t is defined as

$$(49) \qquad \mathscr{D}_t v(\gamma_t) = \lim_{h \to 0^+} \frac{v_{t+h}(\gamma_{t,h}) - v_t(\gamma_t)}{h},$$

where $\gamma_{t,h}(s) = \gamma_t(s) \mathbb{1}_{[0,t]}(s) + \gamma_t(t) \mathbb{1}_{(t,t+h]}(s) \in D([0, t + h]; \mathbb{R}^d)$. The connection between a functional b of paths and the operator B was essentially a matter of definition, as carried out in (6)–(11). To establish some relations between Fréchet differentials of B and horizontal and vertical derivatives of b is much less obvious; some results are given by the following theorem.

THEOREM 7.1. *Suppose $u \colon [0, T] \times \mathcal{D} \to \mathbb{R}$ is given and define, for each $t \in [0, T]$, $v_t \colon D([0, t]; \mathbb{R}^d) \to \mathbb{R}$ as $v_t(\gamma) := u(t, \gamma(t), L_t \gamma)$, in the same way as in (11). Then the vertical derivatives of v_t coincide with the partial derivatives of u with respect to the second variable (i.e., the present state), that is,*

$$(50) \qquad \mathscr{D}_i v_t(\gamma) = \frac{\partial}{\partial x} u(t, x, L_t \gamma), \qquad i = 1, \dots, d.$$

The same result holds true also if u is given from v as in (10). Furthermore let $\gamma_t \in C_b^1([0, t]; \mathbb{R}^d)$ and let again u be given and define v as above. Then

$$\mathscr{D}_t v(\gamma_t) = \frac{\partial u}{\partial t}\big(t, \gamma(t), L_t \gamma_t\big) + \langle Du\big(t, \gamma(t), L_t \gamma_t\big), (L_t \gamma_t)'_+ \rangle,$$

where $\langle \cdot, \cdot \rangle$ is the duality between D and D', Du is the Fréchet derivative of u with respect to φ and the lower script $+$ denotes right derivative.

PROOF. Both claims in the theorem are proved through explicit calculations starting from the definition of derivatives. From the definition of vertical derivative, one gets

$$\mathscr{D}_i v_t(\gamma) = \lim_{h \to 0} \frac{1}{h}\big[v_t(\gamma^h) - v_t(\gamma)\big]$$

$$= \lim_{h \to 0} \frac{1}{h}\big[u\big(t, \gamma^h(t), L_t \gamma^h\big) - u\big(t, \gamma(t), L_t \gamma\big)\big]$$

$$= \lim_{h\to 0} \frac{1}{h} [u(t, \gamma(t) + h, L_t \gamma^h) - u(t, \gamma(t), L_t \gamma)]$$

$$= \frac{\partial}{\partial x_i} u(t, x, L_t \gamma).$$

This proves the first part of the theorem.

For the second part, suppose first that there is no explicit dependence on t in u. Then

$$\mathscr{D}_t v(\gamma_t) = \lim_{h\to 0} \frac{1}{h} [u(\gamma_{t,h}(t), L_{t+h}\gamma_{t,h}) - u(t, \gamma_t(t), L_t \gamma_t)]$$

$$= \lim_{h\to 0} \frac{1}{h} [u(\gamma_t(t), L_{t+h}\gamma_{t,h}) - u(t, \gamma_t(t), L_t \gamma_t)]$$

$$= \lim_{h\to 0} \frac{1}{h} \left[u\left(\gamma_t(t), \begin{cases} \gamma_{t,h}(t+s), & [-t-h,0), \\ \gamma_{t,h}(0), & [-T,-t-h) \end{cases}\right) \right.$$
$$\left. - u\left(\gamma_t(t), \begin{cases} \gamma_t(t+s), & [-t,0), \\ \gamma_t(0), & [-T,-t) \end{cases}\right) \right]$$

$$= \lim_{h\to 0} \frac{1}{h} \left[u\left(\gamma_t(t), \begin{cases} \gamma_t(t), & [-h,0), \\ \gamma_t(t+s+h), & [-t,-h), \\ \gamma_t(t+s+h), & [-t-h,-t), \\ \gamma_t(0), & [-T,-t-h) \end{cases}\right) \right.$$
$$\left. - u\left(\gamma_t(t), \begin{cases} \gamma_t(t+s), & [-h,0), \\ \gamma_t(t+s), & [-t,-h), \\ \gamma_t(0), & [-t-h,-t), \\ \gamma_t(0), & [-T,-t-h) \end{cases}\right) \right].$$

Last line can be written as

(51) $$\lim_{h\to 0} \frac{1}{h} [u(\gamma_t(t), L_t\gamma_t + N_{t,h}\gamma_t) - u(\gamma_t(t), L_t\gamma_t)],$$

where

(52) $$N_{t,h}\gamma_t(s) = \begin{cases} 0, & [-T,-t-h), \\ \gamma_t(t+h+s) - \gamma_t(0), & [-t-h,-t), \\ \gamma_t(t+h+s) - \gamma_t(t+s), & [-t,-h), \\ \gamma_t(t) - \gamma(t+s), & [-h,0). \end{cases}$$

$N_{t,h}\gamma_t$ is a continuous function that goes to 0 as $h \to 0$; moreover, recalling that in the definition of horizontal derivative h is greater than zero, we see that:

(i) for $s \in [-T, -t)$ $\exists \bar{h}$ s.t. $s < -t - \bar{h}$, hence $N_{t,h}\gamma(s) = 0 \ \forall h < \bar{h}$ and

$$\lim_{h\to 0^+} \frac{1}{h} N_{t,h}\gamma(s) = 0 = (L_t\gamma)'(s);$$

(ii) for $s = -t$, since $N_{t,h} \gamma(-t) = \gamma(h) - \gamma(0)$ we have

$$\lim_{h \to 0^+} \frac{1}{h} N_{t,h} \gamma_t(-t) = \left(\frac{d^+}{ds} L_t \gamma_t \right)(-t) = (L_t \gamma_t)'_+(-t) = \gamma'_+(0);$$

(iii) for $s \in (-t, 0)$ $\exists \bar{h}$ s.t. $s < -\bar{h} < 0$, hence

$$\lim_{h \to 0^+} \frac{1}{h} N_{t,h} \gamma_t(s) = \lim_{h \to 0^+} \frac{1}{h} [\gamma_t(t + s + h) - \gamma_t(t + s)]$$

$$= \gamma'_+(t + s) = \gamma'(t + s) = (L_t \gamma_t)'(s).$$

Therefore,

(53)
$$\frac{1}{h} N_{t,h} \gamma_t(s) \xrightarrow{h \to 0^+} (L_t \gamma_t)'_+(s)$$

and, since $\gamma \in C_b^1$,

$$(L_t \gamma_t)'_+(s) = (L_t \gamma_t)'(s) \qquad \forall s \neq -t.$$

Again since $\gamma_t \in C^1$ with bounded derivative, $\frac{1}{h} N_{t,h} \gamma_t$ converges to $(L_t \gamma_t)'_+$ also uniformly. Keeping into account (51) and the definition of Fréchet derivative, one gets

$$\mathscr{D}_t v(\gamma_t) = \lim_{h \to 0} \frac{1}{h} [u(\gamma_t(t), L_t \gamma_t + N_{t,h} \gamma_t) - u(\gamma_t(t), L_t \gamma_t)]$$

$$= \lim_{h \to 0} \frac{1}{h} [\langle Du(\gamma_t(t), L_t \gamma_t), N_{t,h} \gamma_t \rangle + \xi(h)],$$

where ξ is infinitesimal with respect to $\|N_{t,h} \gamma_t\|$ as $h \to 0$,

$$= \lim_{h \to 0} \frac{1}{h} \langle Du(\gamma_t(t), L_t \gamma_t), N_{t,h} \gamma_t \rangle + \lim_{h \to 0} \frac{\|N_{t,h} \gamma_t\|}{h} \frac{\xi(h)}{\|N_{t,h} \gamma_t\|}$$

$$= \langle Du(\gamma_t(t), L_t \gamma_t), (L_t \gamma_t)'_+ \rangle$$

by the dominated convergence theorem.

If now u depends explicitly on t just write

$$\frac{1}{h} [v_{t+h}(\gamma_{t,h}) - v_t(\gamma)] = \frac{1}{h} [u(t + h, \gamma(t), L_{t+h} \gamma_{t,h}) - u(t, \gamma(t), L_t \gamma)]$$

$$= \frac{1}{h} [u(t + h, \gamma(t), L_{t+h} \gamma_{t,h}) - u(t, \gamma(t), L_{t+h} \gamma_{t,h})]$$

$$+ \frac{1}{h} [u(t, \gamma(t), L_{t+h} \gamma_{t,h}) - u(t, \gamma(t), L_t \gamma)];$$

the first term in the last line converges to the time derivative of u while the second can be treated exactly as above. $\square$

Thanks to this result we can reinterpret equation (46), which is the differential form of the infinite-dimensional Kolmogorov equation (26), in terms of the horizontal and vertical derivatives introduced in the previous section.

Consider the Kolmogorov equation with horizontal and vertical derivatives, namely

$$(54) \qquad \begin{cases} \mathscr{D}_t v(\gamma_t) + b_t(\gamma_t) \cdot \mathscr{D} v_t(\gamma_t) + \dfrac{1}{2} \sum_{j=1}^{d} \sigma_j^2 \mathscr{D}_j^2 v_t(\gamma_t) = 0, \\[2mm] v_T(\gamma_T) = f(\gamma_T). \end{cases}$$

THEOREM 7.2. *Let X^{γ_t} be the solution to equation*

$$(5) \qquad \begin{cases} \mathrm{d}X(t) = b_t(X_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W(t), & \text{for } t \in [t_0, T], \\ X_{t_0} = \gamma_{t_0}. \end{cases}$$

Associate to b_t and f the operators B and Φ as in (16); *if such B and Φ satisfy the assumptions of Theorem* 5.4 *then, for almost every t, the function*

$$(55) \qquad v_t(\gamma_t) = \mathbb{E}\big[f\big(X^{\gamma_t}(T) \big) \big]$$

is a solution of the path-dependent Kolmogorov equation (54) *for all $\gamma \in C_b^1([0,T]; \mathbb{R}^d)$ such that $\gamma'(0) = 0$.*

PROOF. Lift equation (5) to the infinite-dimensional SDE (14′) defining the operators A, B and Σ as in the previous sections; associate then to this last equation the PDE (26) with terminal condition given by

$$\Phi\left(\begin{pmatrix} x \\ \varphi \end{pmatrix} \right) = f\left(\widetilde{M} \begin{pmatrix} x \\ \varphi \end{pmatrix} \right).$$

Fix t: with our choice of γ the element $y = (\gamma(t), L_t \gamma_t)$ is in $\mathrm{Dom}(A_{\frown})$, therefore, if B and Φ satisfy Assumptions 2.1 and 5.2, Theorem 5.4 guarantees that $u(s, y) = \mathbb{E}[\Phi(Y^{s,y}(T))]$ is a solution to the Kolmogorov equation. Notice that solving this equation for $s \geq t$ involves only a piece (possibly all) of the path γ_t, so that our "artificial" lengthening by means of L_t is used only for defining all objects in the right way but does not come into the solution of the equation. Of course, in principle one can solve the infinite-dimensional PDE for any $s \in [0, T]$, anyway we are interested in solving it at time t: indeed if we now define v through u by means of (11) we have that

$$\begin{aligned} v_t(\gamma_t) &= u\big(t, \gamma(t), L_t \gamma_t\big) \\ &= \mathbb{E}\big[f\big(\widetilde{M}(Y^{t,y}(T)) \big) \big] \\ &= \mathbb{E}\big[f\big(X^{\gamma_t}(T) \big) \big]. \end{aligned}$$

Recalling Remark 5.5 and noticing that $(L_t \gamma_t)'_+ = A(L_t \gamma_t)$ thanks to the assumption that $\gamma'(0) = 0$, we can apply for almost every t Theorem 7.1 obtaining that equations (46) and (54) coincide. $\square$

REMARK 7.3. If in the above proof one can show that the function u which solves (26) is in fact differentiable with respect to t for every $t \in [0, T]$, then Theorem 7.2 holds everywhere, that is, the function ν defined by (55) solves equation (54) for every $t \in [0, T]$.

REMARK 7.4. The restriction $\gamma'(0) = 0$ is only technical and is likely avoidable with some effort. We intend to address this matter in the future to obtain full generality in our result.

APPENDIX: PROOF OF THEOREM 2.3

Thanks to Theorem 2.2 we can work path by path. Therefore, we consider ω fixed throughout the proof.

We start from a simple estimate; for $y, k \in E$ we have

$$\|Y^{t_0,y+k}(t) - Y^{t_0,y}(t)\|_E$$

$$= \left\| e^{(t-t_0)A}k + \int_{t_0}^t e^{(t-s)A}\big[B\big(s, Y^{t_0,y+k}(s)\big) - B\big(s, Y^{t_0,y}(s)\big)\big]\,\mathrm{d}s \right\|_E$$

$$\leq C\|k\|_E + C\|DB\|_\infty \int_{t_0}^t \|Y^{t_0,y+k}(s) - Y^{t_0,y}(s)\|_E\,\mathrm{d}s$$

hence, by Gronwall's lemma,

$$\text{(A1)} \qquad \sup_t \|Y^{t_0,y+k}(t) - Y^{t_0,y}(t)\|_E \leq \widetilde{C}_Y\|k\|_E.$$

First derivative. We introduce the following equation for the unknown $\xi^{t_0,y}(t)$ taking values in the space of linear bounded operators $L(E, E)$

$$\xi^{t_0,y}(t) = e^{(t-t_0)A} + \int_{t_0}^t e^{(t-s)A}DB\big(s, Y^{t_0,y}(s)\big)\xi^{t_0,y}(s)\,\mathrm{d}s.$$

Existence and uniqueness of a solution in $L^\infty(0, T; L(E, E))$ follow again easily from the contraction mapping principle, since

$$\left\| \int_{t_0}^t e^{(t-s)A}DB\big(s, Y^{t_0,y}(s)\big)\big[\xi_1(s) - \xi_2(s)\big]\,\mathrm{d}s \right\|_{L(E,E)}$$

$$\leq C\|DB\|_\infty \int_{t_0}^t \|\xi_1(s) - \xi_2(s)\|_{L(E,E)}\,\mathrm{d}s.$$

Moreover, by Gronwall's lemma, $\|\xi^{t_0,y}(t)\|_{L(E,E)} \leq C_\xi$ uniformly in t. Now for $k \in E$ we compute

$$r^{t_0,y,k}(t) := Y^{t_0,y+k}(t) - Y^{t_0,y}(t) - \xi^{t_0,y}(t)k$$

$$= \int_{t_0}^t e^{(t-s)A}\big[B\big(s, Y^{t_0,y+k}(s)\big) - B\big(s, Y^{t_0,y}(s)\big)\big]\,\mathrm{d}s$$

$$- \int_{t_0}^t e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\xi^{t_0,y}(s)k \, \mathrm{d}s$$

$$= \int_{t_0}^t e^{(t-s)A}\bigg[\int_0^1 DB\big(s, \alpha Y^{t_0,y+k}(s)$$

$$+ (1-\alpha)Y^{t_0,y}(s)\big)\big(Y^{t_0,y+k}(s) - Y^{t_0,y}(s)\big)\, \mathrm{d}\alpha$$

$$- DB\big(s, Y^{t_0,y}(s)\big)\xi^{t_0,y}(s)k\bigg]\, \mathrm{d}s$$

$$= \int_{t_0}^t e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)r^{t_0,y,k}(s) \, \mathrm{d}s$$

$$+ \int_{t_0}^t e^{(t-s)A}\bigg[\int_0^1 DB\big(s, \alpha Y^{t_0,y+k}(s) + (1-\alpha)Y^{t_0,y}(s)\big)\, \mathrm{d}\alpha$$

$$- DB\big(s, Y^{t_0,y}(s)\big)\bigg]\big(Y^{t_0,y+k}(s) - Y^{t_0,y}(s)\big)\, \mathrm{d}s.$$

Recalling (A1), we get

$$\big\|r^{t_0,y,k}(t)\big\|_E \le C\|DB\|_\infty \int_{t_0}^t \big\|r^{t_0,y,k}(s)\big\|_E \, \mathrm{d}s$$

$$+ C \cdot \widetilde{C}_Y \|k\|_E \int_{t_0}^t \bigg\|\int_0^1 DB\big(s, \alpha Y^{t_0,y+k}(s) + (1-\alpha)Y^{t_0,y}(s)\big)\, \mathrm{d}\alpha$$

$$- DB\big(s, Y^{t_0,y}(s)\big)\bigg\|_{L(E,E)} \, \mathrm{d}s$$

$$\le C\|DB\|_\infty \int_{t_0}^t \big\|r^{t_0,y,k}(s)\big\|_E \, \mathrm{d}s$$

$$+ C \cdot \widetilde{C}_Y \|k\|_E \|D^2 B\|_\infty \int_{t_0}^t \int_0^1 \alpha\big\|Y^{t_0,y+k}(s) + Y^{t_0,y}(s)\big\|_E \, \mathrm{d}\alpha \, \mathrm{d}s$$

$$\le C\|DB\|_\infty \int_{t_0}^t \big\|r^{t_0,y,k}(s)\big\|_E \, \mathrm{d}s + C \cdot \widetilde{C}_Y (T - t_0)\big\|D^2 B\big\|_\infty \|k\|_E^2$$

which yields, by Gronwall's lemma,

$$\big\|r^{t_0,y,k}(t)\big\|_E \le \widetilde{C}\|k\|^2{}_E.$$

Therefore,

$$\xi^{t_0,y}(t)k = DY^{t_0,y}(t)k \qquad \forall k \in E.$$

We proceed with an estimate about the continuity of $\xi^{t_0,y}(t)$ with respect to the initial condition y. For $h, k \in E$

$$\big\|\xi^{t_0,y+k}(t)h - \xi^{t_0,y}(t)h\big\|_E$$

$$= \bigg\|\int_{t_0}^t e^{(t-s)A}\big[DB\big(s, Y^{t_0,y+k}(s)\big)\xi^{t_0,y+k}(s)h - DB\big(s, Y^{t_0,y}(s)\big)\xi^{t_0,y}h\big]\, \mathrm{d}s\bigg\|_E$$

$$\leq \left\| \int_{t_0}^t e^{(t-s)A} \big[DB\big(s, Y^{t_0, y+k}(s)\big)\xi^{t_0, y+k}(s)h \right.$$

$$\left. - DB\big(s, Y^{t_0, y+k}(s)\big)\xi^{t_0, y}(s)h \big] \mathrm{d}s \right\|_E$$

$$+ \left\| \int_{t_0}^t e^{(t-s)A} \big[DB\big(s, Y^{t_0, y+k}(s)\big)\xi^{t_0, y}(s)h \right.$$

$$\left. - DB\big(s, Y^{t_0, y}(s)\big)\xi^{t_0, y}(s)h \big] \mathrm{d}s \right\|_E$$

$$\leq C\|DB\|_\infty \int_{t_0}^t \big\| \xi^{t_0, y+k}(s)h - \xi^{t_0, y}(s)h \big\|_E \, \mathrm{d}s$$

$$+ C \int_{t_0}^t \big\| DB\big(s, Y^{t_0, y+k}(s)\big) - DB\big(s, Y^{t_0, y}(s)\big) \big\|_{L(E,E)} \big\| \xi^{t_0, y}(s)h \big\|_E \, \mathrm{d}s$$

$$\leq C\|DB\|_\infty \int_{t_0}^t \big\| \xi^{t_0, y+k}(s)h - \xi^{t_0, y}(s)h \big\|_E \, \mathrm{d}s$$

$$+ C \cdot C_\xi \|h\|_E \|D^2 B\|_\infty \int_{t_0}^t \big\| Y^{t_0, y+k}(s) - Y^{t_0, y}(s) \big\|_E \, \mathrm{d}s$$

$$\leq C\|DB\|_\infty \int_{t_0}^t \big\| \xi^{t_0, y+k}(s)h - \xi^{t_0, y}(s)h \big\|_E \, \mathrm{d}s$$

$$+ C \cdot C_\xi \|D^2 B\|_\infty \widetilde{C}_Y (t - t_0)\|h\|_E \|k\|_E.$$

Again by Gronwall's lemma, we get

$$\text{(A3)} \qquad\qquad \big\| \xi^{t_0, y+k}(t)h - \xi^{t_0, y}(t)h \big\|_E \leq \widetilde{C}_\xi \|h\|_E \|k\|_E.$$

Therefore, $\xi^{t_0, y}(t)$ is uniformly continuous in y uniformly in t.

Second derivative. Let us consider the operator $\mathcal{U}$ defined on the space $C([t_0, T]; L(E, E; E))$ through the equation

$$\mathcal{U}(Y)(t)(h, k) = \int_{t_0}^t e^{(t-s)A} D^2 B\big(s, Y^{t_0, y}(s)\big)\big(\xi^{t_0, y}(s)h, \xi^{t_0, y}(s)k\big) \, \mathrm{d}s$$

$$+ \int_{t_0}^t e^{(t-s)A} DB\big(s, Y^{t_0, y}(s)\big)Y(s)(h, k) \, \mathrm{d}s$$

for $h, k \in E$; it is immediate to check that $\mathcal{U}(Y)$ belongs to $C([t_0, T]; L(E, E; E))$.
    Since

$$\sup_{t, h, k} \big\| \mathcal{U}(Y_1)(t)(h, k) - \mathcal{U}(Y_2)(t)(h, k) \big\|_E$$

$$\leq C\|DB\|_\infty T \sup_{t, h, k} \big\| Y_1(t)(h, k) - Y_2(t)(h, k) \big\|_E$$

there exists a unique fixed point for $\mathcal{U}$, which will be denoted by $\eta^{t_0, y}(t)(h, k)$; furthermore simple calculations yield that $\|\eta^{t_0, y}(t)\|_{L(E, E; E)} \leq C_\eta$ uniformly in t.

We now compute:

$$\tilde{r}^{t_0,y,h,k}(t) := \xi^{t_0,y+k}(t)h - \xi^{t_0,y}(t)h - \eta^{t_0,y}(t)(h,k)$$

$$= \int_{t_0}^t e^{(t-s)A} DB\big(s, Y^{t_0,y+k}(s)\big)\xi^{t_0,y+k}(s)h \, ds$$

$$- \int_{t_0}^t e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\xi^{t_0,y}(s)h \, ds$$

$$- \int_{t_0}^t e^{(t-s)A} D^2 B\big(s, Y^{t_0,y}(s)\big)\big(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k\big) \, ds$$

$$- \int_{t_0}^t e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\eta^{t_0,y}(s)(h,k) \, ds$$

$$= \int_{t_0}^t e^{(t-s)A} DB\big(s, Y^{t_0,y+k}(s)\big)\xi^{t_0,y+k}(s)h \, ds$$

$$- \int_{t_0}^t e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\xi^{t_0,y+k}(s)h \, ds$$

$$+ \int_{t_0}^t e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\xi^{t_0,y+k}(s)h \, ds$$

$$- \int_{t_0}^t e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\xi^{t_0,y}(s)h \, ds$$

$$- \int_{t_0}^t e^{(t-s)A} D^2 B\big(s, Y^{t_0,y}(s)\big)\big(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k\big) \, ds$$

$$- \int_{t_0}^t e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\eta^{t_0,y}(s)(h,k) \, ds$$

$$= \int_{t_0}^t e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\tilde{r}^{t_0,y,h,k}(s) \, ds$$

$$+ \int_{t_0}^t e^{(t-s)A} \big[DB\big(s, Y^{t_0,y+k}(s)\big) - DB\big(s, Y^{t_0,y}(s)\big)\big]\xi^{t_0,y+k}(s)h \, ds$$

$$- \int_{t_0}^t e^{(t-s)A} D^2 B\big(s, Y^{t_0,y}(s)\big)\big(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k\big) \, ds$$

$$= \int_{t_0}^t e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\tilde{r}^{t_0,y,h,k}(s) \, ds$$

$$+ \int_{t_0}^t e^{(t-s)A}\bigg[\int_0^1 D^2 B\big(s, \alpha Y^{t_0,y+k}(s) + (1-\alpha)Y^{t_0,y}(s)\big) \, d\alpha$$

$$\times \big(\xi^{t_0,y+k}(s)h, Y^{t_0,y+k}(s) - Y^{t_0,y}(s)\big)$$

$$- D^2 B\big(s, Y^{t_0,y}(s)\big)\big(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k\big)\Big] ds$$

$$= \int_{t_0}^t e^{(t-s)A} D B\big(s, Y^{t_0,y}(s)\big)\tilde{r}^{t_0,y,h,k}(s) \, ds$$

$$+ \int_{t_0}^t e^{(t-s)A}\bigg[\int_0^1 D^2 B\big(s, \alpha Y^{t_0,y+k}(s) + (1-\alpha)Y^{t_0,y}(s)\big) \, d\alpha$$

$$- D^2 B\big(s, Y^{t_0,y}(s)\big)\bigg]\big(\xi^{t_0,y+k}(s)h, Y^{t_0,y+k}(s) - Y^{t_0,y}(s)\big) \, ds$$

$$+ \int_{t_0}^t e^{(t-s)A}\big[D^2 B\big(s, Y^{t_0,s}(s)\big)\big(\xi^{t_0,y+k}(s)h, Y^{t_0,y+k}(s) - Y^{t_0,y}(s)\big)$$

$$- D^2 B\big(s, Y^{t_0,y}(s)\big)\big(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k\big)\big] \, ds$$

$$= \int_{t_0}^t e^{(t-s)A} D B\big(s, Y^{t_0,y}(s)\big)\tilde{r}^{t_0,y,h,k}(s) \, ds$$

$$+ \int_{t_0}^t e^{(t-s)A}\bigg[\int_0^1 D^2 B\big(s, \alpha Y^{t_0,y+k}(s) + (1-\alpha)Y^{t_0,y}(s)\big) \, d\alpha$$

$$- D^2 B\big(s, Y^{t_0,y}(s)\big)\bigg]\big(\xi^{t_0,y+k}(s)h, Y^{t_0,y+k}(s) - Y^{t_0,y}(s)\big) \, ds$$

$$+ \int_{t_0}^t e^{(t-s)A} D^2 B\big(s, Y^{t_0,y}(s)\big)$$

$$\times \big[\big(\xi^{t_0,y+k}(s)h, Y^{t_0,y+k}(s) - Y^{t_0,y}(s)\big) - \big(\xi^{t_0,y+k}(s)h, \xi^{t_0,y}(s)k\big)$$

$$+ \big(\xi^{t_0,y+k}(s)h, \xi^{t_0,y}(s)k\big) - \big(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k\big)\big] \, ds$$

$$= \int_{t_0}^t e^{(t-s)A} D B\big(s, Y^{t_0,y}(s)\big)\tilde{r}^{t_0,y,h,k}(s) \, ds$$

$$+ \int_{t_0}^t e^{(t-s)A}\bigg[\int_0^1 D^2 B\big(s, \alpha Y^{t_0,y+k}(s) + (1-\alpha)Y^{t_0,y}(s)\big) \, d\alpha$$

$$- D^2 B\big(s, Y^{t_0,y}(s)\big)\bigg]\big(\xi^{t_0,y+k}(s)h, Y^{t_0,y+k}(s) - Y^{t_0,y}(s)\big) \, ds$$

$$+ \int_{t_0}^t e^{(t-s)A} D^2 B\big(s, Y^{t_0,y}(s)\big)$$

$$\times \big(\xi^{t_0,y+k}(s)h, Y^{t_0,y+k}(s) - Y^{t_0,y}(s) - \xi^{t_0,y}(s)k\big) \, ds$$

$$+ \int_{t_0}^t e^{(t-s)A} D^2 B\big(s, Y^{t_0,y}(s)\big)$$

$$\times \big(\xi^{t_0,y+k}(s)h - \xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k\big) \, ds.$$

These calculations together with (A1) and (A2) imply that

$$
\left\| \tilde{r}^{t_0, y, h, k}(t) \right\|_E
$$

$$
\leq C \|DB\|_\infty \int_{t_0}^t \left\| \tilde{r}^{t_0, y, h, k}(s) \right\|_E \, \mathrm{d}s
$$

$$
+ C \int_{t_0}^t \left\| \int_0^1 D^2 B\big(s, \alpha Y^{t_0, y+k}(s) + (1-\alpha) Y^{t_0, y}\big) \, \mathrm{d}\alpha \right.
$$

$$
\left. - D^2 B\big(s, Y^{t_0, y}(s)\big) \right\|_{L(E, E; E)} \left\| \xi^{t_0, y+k}(s) h \right\|_E
$$

$$
\times \left\| Y^{t_0, y+k}(s) - Y^{t_0, y}(s) \right\|_E \, \mathrm{d}s
$$

$$
+ C \|D^2 B\|_\infty \int_{t_0}^t \left\| \xi^{t_0, y+k}(s) h \right\|_E
$$

$$
\times \left\| Y^{t_0, y+k}(s) - Y^{t_0, y}(s) - \xi^{t_0, y}(s) k \right\|_E \, \mathrm{d}s
$$

$$
+ C \|D^2 B\|_\infty \int_{t_0}^t \left\| \xi^{t_0, y+k}(s) h - \xi^{t_0, y}(s) h \right\|_E \cdot \left\| \xi^{t_0, y}(s) k \right\|_E \, \mathrm{d}s
$$

$$
\leq C \|DB\|_\infty \int_{t_0}^t \left\| \tilde{r}^{t_0, y, h, k}(s) \right\|_E \, \mathrm{d}s + C \cdot C_\xi \widetilde{C}_Y \|h\|_E \|k\|_E
$$

$$
\times \int_{t_0}^t \left\| \int_0^1 D^2 B\big(s, \alpha Y^{t_0, y+k}(s) + (1-\alpha) Y^{t_0, y}\big) \, \mathrm{d}\alpha \right.
$$

$$
\left. - D^2 B\big(s, Y^{t_0, y}(s)\big) \right\|_{L(E, E; E)} \, \mathrm{d}s
$$

$$
+ C \cdot C_\xi \|D^2 B\|_\infty \|h\|_E
$$

$$
\times \int_{t_0}^t \left\| \int_0^1 \xi^{t_0, \alpha(y+k)+(1-\alpha)y}(s) k \, \mathrm{d}\alpha - \xi^{t_0, y}(s) k \right\|_E \, \mathrm{d}s
$$

$$
+ C \cdot C_\xi \widetilde{C}_\xi T \|D^2 B\|_\infty \|h\|_E \|k\|_E^2
$$

$$
\leq C \|DB\|_\infty \int_{t_0}^t \left\| \tilde{r}^{t_0, y, h, k}(s) \right\|_E \, \mathrm{d}s + C_1 \|h\|_E \|k\|_E
$$

$$
\times \int_{t_0}^t \left\| \int_0^1 D^2 B\big(s, \alpha Y^{t_0, y+k}(s) + (1-\alpha) Y^{t_0, y}\big) \, \mathrm{d}\alpha \right.
$$

$$
\left. - D^2 B\big(s, Y^{t_0, y}(s)\big) \right\|_{L(E, E; E)} \, \mathrm{d}s
$$

$$
+ C_2 \|h\|_E \int_{t_0}^t \left\| \int_0^1 \xi^{t_0, y+\alpha k}(s) \, \mathrm{d}\alpha - \xi^{t_0, y}(s) \right\|_{L(E, E)} \, \mathrm{d}s \, \|k\|_E
$$

$$
+ C_3 \|h\|_E \|k\|_E^2.
$$

Finally, by an application of Gronwall's lemma

$$\frac{\|\tilde{r}^{t_0,y,h,k}(t)\|_E}{\|k\|_E} \le C_4 \|h\|_E$$

$$\times \left[\int_{t_0}^t \left\| \int_0^1 D^2 B(s, \alpha Y^{t_0,y+k}(s) + (1-\alpha)Y^{t_0,y}) \, d\alpha \right. \right.$$

$$\left. - D^2 B(s, Y^{t_0,y}(s)) \right\|_{L(E,E;E)} ds$$

$$\left. + \int_{t_0}^t \left\| \int_0^1 \xi^{t_0,y+\alpha k}(s) \, d\alpha - \xi^{t_0,y}(s) \right\|_{L(E,E)} ds + \|k\|_E \right]$$

and such quantity goes to 0 uniformly in $\|h\|_E \le N \; \forall N > 0$ when $\|k\|_E$ goes to 0 by Lebesgue's dominated convergence theorem.

Our next step is to study the continuity of the second derivative computed above. We have

$$\eta^{t_0,y}(t)(h,k) - \eta^{t_0,w}(t)(h,k)$$

$$= \int_{t_0}^t e^{(t-s)A} \big[D^2 B(s, Y^{t_0,y}(s))(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k)$$

$$- D^2 B(s, Y^{t_0,w}(s))(\xi^{t_0,w}(s)h, \xi^{t_0,w}(s)k) \big] ds$$

$$+ \int_{t_0}^t e^{(t-s)A} \big[DB(s, Y^{t_0,y}(s)) \eta^{t_0,y}(s)(h,k)$$

$$- DB(s, Y^{t_0,w}(s)) \eta^{t_0,w}(s)(h,k) \big] ds$$

$$= I_1 + I_2;$$

then

$$I_1 = \int_{t_0}^t e^{(t-s)A} \big[D^2 B(s, Y^{t_0,y}(s))(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k)$$

$$- D^2 B(s, Y^{t_0,w}(s))(\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k)$$

$$+ D^2 B(s, Y^{t_0,w}(s))(\xi^{t_0,y}(s)h, \xi^{t_0,y}k)$$

$$- D^2 B(s, Y^{t_0,w}(s))(\xi^{t_0,w}(s)h, \xi^{t_0,w}(s)k) \big] ds$$

$$= \int_{t_0}^t e^{(t-s)A} \big[D^2 B(s, Y^{t_0,y}(s)) - D^2 B(s, Y^{t_0,w}(s)) \big] (\xi^{t_0,y}(s)h, \xi^{t_0,y}(s)k) \, ds$$

$$+ \int_{t_0}^t e^{(t-s)A} D^2 B(s, Y^{t_0,w}(s))([\xi^{t_0,y}(s) - \xi^{t_0,w}]h, \xi^{t_0,y}k) \, ds$$

$$+ \int_{t_0}^t e^{(t-s)A} D^2 B(s, Y^{t_0,w}(s))(\xi^{t_0,w}(s)h, [\xi^{t_0,y}(s) - \xi^{t_0,w}(s)]k) \, ds$$

and

$$I_2 = \int_{t_0}^{t} e^{(t-s)A} DB\big(s, Y^{t_0,y}(s)\big)\big[\eta^{t_0,y}(s)(h,k) - \eta^{t_0,w}(s)(h,k)\big] \, \mathrm{d}s$$

$$+ \int_{t_0}^{t} e^{(t-s)A}\big[DB\big(s, Y^{t_0,y}(s)\big) - DB\big(s, Y^{t_0,w}(s)\big)\big]\eta^{t_0,w}(s)(h,k) \, \mathrm{d}s.$$

Recalling all the estimates previously obtained and the fact that both $\|Y^{t_0,y}(t)\|_E$ and $\|\xi^{t_0,y}(t)\|_{L(E,E)}$ are bounded uniformly in t, denoting with C_H the Hölder constant of $D^2 B$, we get

$$\big\|\eta^{t_0,y}(t)(h,k) - \eta^{t_0,w}(t)(h,k)\big\|_E$$

$$\leq C \cdot C_H \int_{t_0}^{t} \big\|Y^{t_0,y}(s) - Y^{t_0,w}(s)\big\|_E^{\alpha} \big\|\xi^{t_0,y}(s)h\big\|_E \big\|\xi^{t_0,y}(s)k\big\|_E \, \mathrm{d}s$$

$$+ C\|D^2 B\|_{\infty} \int_{t_0}^{t} \big\|\xi^{t_0,y}(s) - \xi^{t_0,w}(s)\big\|_{L(E,E)}^{\alpha} \big\|\xi^{t_0,y}(s) - \xi^{t_0,w}(s)\big\|_{L(E,E)}^{1-\alpha}$$

$$\times \big[\|h\|_E\big\|\xi^{t_0,y}(s)k\big\|_E + \big\|\xi^{t_0,w}(s)h\big\|_E\|k\|_E\big] \, \mathrm{d}s$$

$$+ C\|DB\|_{\infty} \int_{t_0}^{t} \big\|\eta^{t_0,y}(s)(h,k) - \eta^{t_0,w}(s)(h,k)\big\|_E \, \mathrm{d}s$$

$$+ C\|DB\|_{\infty} \int_{t_0}^{t} \big\|Y^{t_0,y}(s) - Y^{t_0,w}(s)\big\|_E^{\alpha} \big\|Y^{t_0,y}(s) - Y^{t_0,w}(s)\big\|_E^{1-\alpha}$$

$$\times \big\|\eta^{t_0,w}(s)(h,k)\big\|_E \, \mathrm{d}s$$

$$\leq C_5\|h\|_E\|k\|_E\|y - w\|_E^{\alpha} + C_6 \int_{t_0}^{t} \big\|\eta^{t_0,y}(s)(h,k) - \eta^{t_0,w}(s)(h,k)\big\|_E \, \mathrm{d}s$$

hence

$$\big\|\eta^{t_0,y}(t)(h,k) - \eta^{t_0,w}(t)(h,k)\big\|_E \leq C_7\|h\|_E\|k\|_E\|y - w\|_E^{\alpha}$$

which shows that the second Fréchet derivative of the map $y \mapsto Y^{t_0,y}(t)$ is α-Hölder continuous.

Continuity with respect to the initial time. Fix $t \in [0, T]$, $\omega \in \Omega_0$ (that we do not write, as before) and $\varepsilon > 0$ and consider two initial times s_1 and s_2, with $s_1 < s_2$ for simplicity. Since we assume that $y \in \mathcal{L}^p$ or $y \in \widehat{\mathcal{C}}$, we can find δ such that

$$\big\|Y^{s_2,y}(t) - Y^{s_1,y}(t)\big\|_E$$

$$\leq \Big\|e^{(t-s_2)A}\big(1 - e^{(s_2-s_1)A}\big)y$$

$$+ \int_{s_2}^{t} e^{(t-r)A}\big[B\big(r, Y^{s_2,y}(r)\big) - B\big(r, Y^{s_1,y}(r)\big)\big] \, \mathrm{d}r$$

$$- \int_{s_1}^{s_2} e^{(t-r)A} B\big(r, Y^{s_1,y}(r)\big)\,\mathrm{d}r - \int_{s_1}^{s_2} e^{(t-r)A} \Sigma\,\mathrm{d}W(r) \bigg\|_E$$

$$\leq C \big\|(1 - e^{(s_2-s_1)A})y\big\|_E + C\|DB\|_\infty \int_{s_2}^{t} \big\|Y^{s_2,y}(r) - Y^{s_1,y}(r)\big\|_E\,\mathrm{d}r$$

$$+ C\|B\|_\infty |s_2 - s_1| + C\|\Sigma\|_\infty \big|W(s_2) - W(s_1)\big|$$

$$\leq C\|DB\|_\infty \int_{s_2}^{t} \big\|Y^{s_2,y}(r) - Y^{s_1,y}(r)\big\|_E\,\mathrm{d}r + C\varepsilon$$

for $|s_2 - s_1| < \delta$, because e^{sA} is strongly continuous and $W(\cdot, \omega)$ is continuous.

The conclusion follows using Gronwall's lemma, ε being arbitrary.

Acknowledgments. We are grateful to the referees for their careful comments and stimulating questions that lead us to improve the exposition and revise several details for the final version of our work.

REFERENCES

BENSOUSSAN, A., DA PRATO, G., DELFOUR, M. C. and MITTER, S. K. (1992). *Representation and Control of Infinite Dimensional Systems*. Birkhäuser, Boston.

CONT, R. and FOURNIE, D. (2010a). A functional extension of the Ito formula. *C. R. Math. Acad. Sci. Paris* **348** 57–61. MR2586744

CONT, R. and FOURNIÉ, D.-A. (2010b). Change of variable formulas for non-anticipative functionals on path space. *J. Funct. Anal.* **259** 1043–1072. MR2652181

CONT, R. and FOURNIÉ, D.-A. (2013). Functional Itô calculus and stochastic integral representation of martingales. *Ann. Probab.* **41** 109–133. MR3059194

COSSO, A. (2012). Viscosity solutions of path-dependent PDEs and non-Markovian forward-backward stochastic equations. Preprint. Available at arXiv:1202.2502.

DA PRATO, G. and ZABCZYK, J. (1992). *Stochastic Equations in Infinite Dimensions*. Cambridge Univ. Press, Cambridge. MR1207136

DI GIROLAMI, C., FABBRI, G. and RUSSO, F. (2014). The covariation for Banach space valued processes and applications. *Metrika* **77** 51–104. MR3152020

DI GIROLAMI, C. and RUSSO, F. (2010). Infinite dimensional stochastic calculus via regularization. HAL INRIA 00473947.

DI GIROLAMI, C. and RUSSO, F. (2012). Generalized covariation and extended Fukushima decomposition for Banach space-valued processes. Applications to windows of Dirichlet processes. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **15** 1250007, 50. MR2957134

DI GIROLAMI, C. and RUSSO, F. (2014). Generalized covariation for Banach space valued processes, Itô formula and applications. *Osaka J. Math.* **51** 729–783. MR3272616

DUPIRE, B. (2009). Functional Itô calculus. Portfolio Research Paper 2009-04, Bloomberg.

EKREN, I., TOUZI, N. and ZHANG, J. (2016a). Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part I. *Ann. Probab.* **44** 1212–1253.

EKREN, I., TOUZI, N. and ZHANG, J. (2016b). Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part II. *Ann. Probab.* **44** 2507–2553.

EKREN, I., KELLER, C., TOUZI, N. and ZHANG, J. (2014). On viscosity solutions of path dependent PDEs. *Ann. Probab.* **42** 204–236. MR3161485

FEDERICO, S., GOLDYS, B. and GOZZI, F. (2010). HJB equations for the optimal control of differential equations with delays and state constraints, I: Regularity of viscosity solutions. *SIAM J. Control Optim.* **48** 4910–4937. MR2735510

FUHRMAN, M., MASIERO, F. and TESSITORE, G. (2010). Stochastic equations with delay: Optimal control via BSDEs and regular solutions of Hamilton–Jacobi–Bellman equations. *SIAM J. Control Optim.* **48** 4624–4651. MR2683901

GOZZI, F. and MARINELLI, C. (2006). Stochastic optimal control of delay equations arising in advertising models. In *Stochastic Partial Differential Equations and Applications—VII. Lect. Notes Pure Appl. Math.* **245** 133–148. Chapman & Hall/CRC, Boca Raton, FL. MR2227226

PENG, S. and WANG, F. (2011). BSDE, path-dependent PDE and nonlinear Feynman–Kac formula. Preprint. Available at arXiv:1108.4317.

TANG, S. and ZHANG, F. (2013). Path-dependent optimal stochastic control and viscosity solution of associated Bellman equations. *Discrete Contin. Dyn. Syst.* **35** 5521–5553.

DIPARTIMENTO DI MATEMATICA
UNIVERSITÀ DI PISA
LARGO BRUNO PONTECORVO 5
56127 PISA IT
ITALY
E-MAIL: flandoli@dma.unipi.it
        zanco@mail.dm.unipi.it