An infinite-dimensional approach to path-dependent Kolmogorov’s equations
Franco Flandoli∗ Giovanni Zanco†
Dipartimento di Matematica, Università di Pisa
December 24, 2013

Abstract
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’s 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’s 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
\[ \begin{cases} \frac{dX(t)}{dt} = b_t(X_t) dt + \sigma dW(t) & \text{for } t \in [t_0, T], \\ X_{t_0} = \gamma_{t_0} \end{cases} \tag{1} \]
where \( \{W(t)\}_{t \geq 0} \) is a Brownian motion in \( \mathbb{R}^d \), \( \sigma \) 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 \( \mathbb{R}^d \), \( \gamma_{t_0} \) is a given path on \( [0, t_0] \).

After the insightful ideas proposed by Dupire (\[Dup09\]) and Cont and Fournié (\[CF13\], \[CF10a\], \[CF10b\]), 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 generalising some classical concept to this setting, like forward-backward systems and viscosity solutions (see \[PW11\], \[ITZ13\], \[EKTZar\], \[ETZ13a\], \[ETZ13b\]). Also, depending on the approach, there are some similarities with investigations about delay equations, see for instance \[FGG10\], \[GM06\], \[FMT10\]. Some of these works formulate a path-dependent Kolmogorov equation associated to the path-dependent SDE \( (1) \).

Several issues about such Kolmogorov’s 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.

∗flandoli@dma.unipi.it
†zanco@mail.dm.unipi.it
1.1 Notation

We will use the following notations 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 $r$, $r < t$. Stochastic processes will be denoted with upper-case letters, while greek lower-case letters will be used for deterministic processes, 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 $\gamma$ as defined on the whole interval $[0, T]$ and read $\gamma_r$ 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 non-anticipative 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 $\sigma$-field on $D([0, t]; \mathbb{R}^d)$. Some possible examples of path-dependent functions are the following:

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

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

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

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

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

(iii) 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 a 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

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

(2)

(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 $\beta$ is a finite-dimensional Brownian motion (section 2.1).

We associate to it the backward Kolmogorov equation in integral form with terminal condition $\Phi$

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

(3)

and the related concept of solution (section 3).

Our main result, under suitable regularity assumptions on $B$ and $\Phi$, as explained in section 5, is the following:

Theorem. The function

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

where $Y^{s,y}(t)$ solves equation (2), is of class $C^{2,\alpha}$ and solves the backward Kolmogorov equation in the space of continuous functions.
Since under our assumptions all the integrands appearing in (3) are in $L^\infty$, 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’s backward equation in its differential form
\[
\frac{\partial u}{\partial t}(t, y) + (Du(t, y), Ay + B(t, y)) + \frac{1}{2} \sum_{j=1}^{d} \sigma_j^2 D^2 u(t, y)(e_j, e_j), \quad u(T, \cdot) = \Phi.
\]
We moreover exhibit some links between our results and the path-dependent calculus developed by Cont and Fournié (section 6). In particular, thanks to the theorem stated above, we can prove the following result (again under some regularity assumptions compatible with those of the previous theorem):

**Theorem.** The function $\nu_s(\gamma_s) = \mathbb{E}[f(X_{\gamma_s}(T))]$, where $X_{\gamma_s}(t)$ is the solution to equation (1), solves the path-dependent backward Kolmogorov equation
\[
\begin{align*}
D_t \nu_t(\gamma_t) + b_t(\gamma_t) \cdot D \nu_t(\gamma_t) + \frac{1}{2} \sum_{j=1}^{d} \sigma_j^2 D^2 \nu_t(\gamma_t) &= 0, \\
\nu_T(\gamma_T) &= f(\gamma_T).
\end{align*}
\]
in which the derivatives are understood as horizontal and vertical derivatives as defined in [CF13].

### 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 $L^2$. 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 $\mathcal{C}$ defined as:

\[
\mathcal{C} := \left\{ y = (\varphi) \in \mathbb{R}^d \times C([-T, 0); \mathbb{R}^d) \text{ s.t. } x = \lim_{s \to 0} \varphi(s) \right\}.
\]

This leads to two issues: first, the operator $B$ (our abstract realization of the functional $b$) takes values in a space larger than $\mathcal{C}$, 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

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

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 for example [DGRarb], [DGRara], [DGFRar], [RDG10]), 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 the integral functionals) and we develop rigorously the relation between the SDE and the PDE in this framework (section 4). We then introduce a keen 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 onwards fix a time horizon $0 < T < \infty$ and a filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\right)$.

We introduce the following spaces:

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

\[ \widehat{C}: = \left\{ y = (x, \varphi) \in C \text{ s.t. } x = \lim_{s\uparrow 0} \varphi(s) \right\}, \]

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

\[ D_1: = \left\{ y = (x, \varphi) \in D \text{ s.t. } \varphi \text{ is discontinuous at most in the only point } t \right\}, \]

\[ L^p: = \mathbb{R}^d \times L^p([-T,0); \mathbb{R}^d), \quad p \geq 2. \]

All of them apart from $L^p$ are Banach spaces with respect to the norm $\| (x, \varphi) \|^2 = |x|^2 + \| \varphi \|^2_{C_b}$, while $L^p$ is a Banach space with respect to the norm $\| (x, \varphi) \|^2 = |x|^2 + \| \varphi \|^2_{L^p}$; the space $D$ 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

\[ \widehat{C} \subset C \subset D \subset L^p \]

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

Notice that the space $\widehat{C}$ 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

\[
\begin{align*}
\text{d}X(t) &= b_t(X_t) \, \text{d}t + \sigma \, \text{d}W(t) \quad \text{for } t \in [t_0,T], \\
X_{t_0} &= \gamma_{t_0},
\end{align*}
\]

(5)

where $\sigma$ 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 $D = \cup_t D([0,t]; \mathbb{R}^d)$.

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

\[ M_t : D([-T,0); \mathbb{R}^d) \to D([0,t); \mathbb{R}^d) \]

\[ M_t(\varphi)(s) = \varphi(s - t), \quad s \in [0,t) \]  

(6)

and the backward extension operator

\[ L_t : D([0,t); \mathbb{R}^d) \to D([-T,0); \mathbb{R}^d) \]

\[ L_t(\gamma)(s) = \gamma(0)1_{[-T,-t)}(s) + \gamma(t + s)1_{[-t,0)}(s), \quad s \in [-T,0). \]  

(7)

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

\[ M_t L_t \gamma = \gamma \]  

(8)

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
\begin{equation}
\tilde{M}_t(x, \varphi)(s) = \begin{cases} M_t \varphi(s) & s \in [0, t) \\ x & s = t. \end{cases}
\end{equation}

Now given a functional $b$ on $D$ as in (5) one can define a function $\hat{b}$ on $[0, T] \times D$ setting
\begin{equation}
\hat{b}(t, (x, \varphi)) = \tilde{b}(t, x, \varphi) := b_t \left( \tilde{M}_t(x, \varphi) \right);
\end{equation}
conversely if $\hat{b}$ is given one can obtain a functional $b$ on $D$ setting
\begin{equation}
b_t(\gamma) := \hat{b}(t, \gamma(t), L_t \gamma).
\end{equation}
The idea is simply to translate 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 $[0, T] \times D$ given by
\begin{equation}
\hat{b}(t, (x, \varphi)) = \int_0^t g(\varphi(s - t)) \, ds.
\end{equation}

We consider again the path-dependent SDE (5) with the initial condition given now by a path $\varphi$ on $[\psi(t_0), t_0]$ and its terminal value $x = \psi(t_0)$,
\begin{equation}
\begin{aligned}
\frac{dX(s)}{ds} &= 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_0 - T, t_0);
\end{aligned}
\end{equation}
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 (12) (in some space), for $t \in [t_0, T]$ we set
\begin{equation}
Y(t) = \left( X(t) \right)_{\{X(t+s)\}_{s \in [-T,0]}}
\end{equation}
and then differentiate with respect to $t$ formally obtaining
\begin{equation}
\frac{dY(t)}{dt} = \left( \frac{X(t)}{X(t+s)} \right)_{s \in [-T,0]} = \left( \begin{array}{c}
0 \\
\left( \dot{X}(t+s) \right)_{s \in [-T,0]}
\end{array} \right) + \left( \begin{array}{c}
b_t(X_t) \\
\sigma \dot{W}(t)
\end{array} \right) = \dot{Y}(t).
\end{equation}

It is therefore natural to define the operators
\begin{equation}
A(x, \varphi) := \left( \begin{array}{c}
0 \\
\varphi
\end{array} \right),
\end{equation}
\begin{equation}
B(t, (x, \varphi)) := \left( \hat{b}(t, (x, \varphi)) \right)
\end{equation}
and
\begin{equation}
\sqrt{\Sigma}(x, \varphi) := \left( \begin{array}{c}
\sigma x \\
0
\end{array} \right)
\end{equation}
and to formulate the infinite dimensional SDE
\begin{equation}
\frac{dY(t)}{dt} = AY(t) \, dt + B(t, Y(t)) \, dt + \sqrt{\Sigma} \, d\beta(t), \quad t \in [t_0, T],
\end{equation}
where $\beta$ is given by
\begin{equation}
\beta(t) = \left( \begin{array}{c}
W(t) \\
0
\end{array} \right).
\end{equation}
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
\[
Y(t) = e^{(t-t_0)A}y + \int_{t_0}^t e^{(t-s)A}B(s, Y(s)) \, ds + \int_{t_0}^t e^{(t-s)A}\sqrt{\Sigma} \, d\beta(s). \tag{13'}
\]
It is not difficult to show that if \( Y \) solves \((13')\) then its first coordinate \( X(t) \) solves the original SDE \((12)\).

### 2.2 Some properties of the convolution integrals

The operator \( A \) has different domains depending on the space we choose:
\[
\text{Dom}(A) = \left\{ (x, \varphi) \in \mathcal{L}^p : \varphi \in W^{1,p}(-T, 0; \mathbb{R}^d), \varphi(0) = x \right\},
\]
\[
\text{Dom}(A_{\subset}) = \left\{ (x, \varphi) \in \subset \mathcal{C} : \varphi \in C^1([-T, 0); \mathbb{R}^d) \right\};
\]
one can think to define \( A \) on \( \mathcal{L}^p \) and then consider its restriction to \( D \) or to \( \subset \mathcal{C} \), as the notation above emphasizes.
It is well known (see theorem 4.4.2 in [BDPD92]) that \( A \) is the infinitesimal generator of a strongly continuous semigroup both in \( \mathcal{L}^p \) and in \( \subset \mathcal{C} \); it’s easy to check that it still generates a semigroup in \( \mathcal{D} \) which is not uniformly continuous. Indeed we have that
\[
e^{tA}(x, \varphi) = \left( \begin{array}{c} x \\ \varphi \\
\end{array} \right) = \left( \begin{array}{c} x \\ \varphi(\xi + t) 1_{[-T,-t]}(\xi) + x 1_{[-t,0]}(\xi) \end{array} \right)_{\xi \in [-T,0]} . \tag{18}
\]
This formula comes from the trivial delay equation
\[
\begin{align*}
\frac{dx(t)}{dt} &= 0, \quad t \geq 0 \\
x(0) &= x, \quad x(\xi) = \varphi(\xi) \text{ for } \xi \in [-T,0];
\end{align*}
\]
its solution, for \( t \geq 0 \), is simply \( x(t) = x \). If we introduce the pair
\[
y(t) := \left( \begin{array}{c} x(t) \\ x|_{[t-T,t]} \end{array} \right)
\]
then
\[
y(t) = e^{tA}(x, \varphi).
\]
However it still holds that
\[
\|e^{tA}\|_{\mathcal{L}(\mathcal{D}, \mathcal{D})} \leq C \text{ for } t \in [0, T] \tag{19}
\]
with \( C \) not depending on \( t \). Moreover it is evident from [18] that \( e^{tA} \) maps \( \mathcal{L}^p \) into \( \mathcal{L}^p \), \( \mathcal{D} \) into \( \mathcal{D} \) and \( \subset \mathcal{C} \) into \( \subset \mathcal{C} \), but it maps \( \mathcal{C} \) into \( \mathcal{D}_{\subset} \), 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
\[
Z^\sigma(t) := \int_{t_0}^t e^{(t-s)A}\sqrt{\Sigma} \, d\beta(s) = \int_{t_0}^t e^{(t-s)A} \left( \sigma \, dW(s) \right), \quad t \geq t_0.
\]
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 formulae. From now on we work in a set $\Omega_0 \subseteq \Omega$ of full probability on which $W$ has continuous trajectories. For any $\omega \in \Omega_0$ and $x \in \mathbb{R}^d$ we have

$$e^{(t-s)A} \sqrt{\Sigma} \left( \begin{array}{c} x \\ 0 \end{array} \right) = \left( \begin{array}{c} \sigma x \\ \sigma \xi \end{array} \right)_{\xi \in [-T,0]}$$

hence

$$Z^{t_0}(t) = \left( \begin{array}{c} \int_{t_0}^t \sigma W(s) \\ \int_{t_0}^t 1_{[-t_0,0]}(\xi) \sigma dW(s) \end{array} \right) = \left( \begin{array}{c} \sigma(W(t) - W(t_0)) \\ \sigma((t+\cdot) \wedge t_0) - W(t_0) \end{array} \right)$$

because

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

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

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

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

$$E\left( \|Z^{t_0}(t)\|_{\mathcal{C}}^2 \right) \leq 2^4 \mathbb{E} \left( \sup_{s \in [0,t-t_0]} |\sigma W(s)|^4 \right) \leq C' \mathbb{E} \left( |W(t-t_0)|^4 \right) \leq C'' (t-t_0)^2$$

where $C'$ and $C''$ are suitable constants. Consequently the same property holds in $\mathcal{C}$ (possibly with a different constant) by continuity of the embedding $\mathcal{C} \subset \mathcal{L}^p$. Moreover from (20) we obtain that for $\omega$ fixed

$$\|Z^{t_0}(t) - Z^{t_0}(s)\|_{\mathcal{C}} = C \left( |W(t) - W(s)| + \sup_{\xi \in [-T,0]} |W((t+\xi) \wedge t_0) - W((s+\xi) \wedge t_0)| \right).$$

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

$$W((t+\xi) \wedge t_0) - W((s+\xi) \wedge t_0) = \left\{ \begin{array}{ll} 0 & \text{for } \xi \in [-T,0] \\ W(t+\xi) - W(t_0) & \text{for } \xi \in [t_0-t, t_0-s] \\ W(t+\xi) - W(s+\xi) & \text{for } \xi \in [t_0-s, 0] \end{array} \right.$$ 

and

$$\sup_{\xi \in [t_0-t, t_0-s]} |W(t+\xi) - W(t_0)| = \sup_{\eta \in [0,t-s]} |W(\eta)|,$$

therefore $Z^{t_0}$ is a continuous process in $\mathcal{C}$, since any fixed trajectory of $W$ is uniformly continuous. The same property holds then in $\mathcal{L}^p$ again by continuity of the embedding $\mathcal{C} \subset \mathcal{L}^p$. We can argue in a similar way for $F^{t_0} : [t_0,T] \times \mathcal{L}^\infty([t_0,T]; \mathcal{D}) \to \mathcal{D}$,

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

From (15) using (18) one deduces that

$$e^{(t-s)A} B(s,\theta(s)) = \left( \begin{array}{c} b_s(\theta(s)) \\ b_s(\theta(s)) 1_{[-t,s]}(\xi) \end{array} \right)$$

and therefore

$$\int_{t_0}^t e^{(t-s)A} B(s,\theta(s)) \, ds = \left( \begin{array}{c} \int_{t_0}^t b_s(\theta(s)) \, ds \\ \int_{t_0}^t b_s(\theta(s)) \, ds \end{array} \right).$$
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

\[ dY(t) = AY(t) \, dt + B(t, Y(t)) \, dt + \sqrt{\Sigma} \, d\beta(t), \quad Y(t_0) = y, \]

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 \(L^p\) and satisfies (13) in \(D\). Although all these theorems are analogous to well known results for stochastic equations in Hilbert spaces (see for example [DPZ92]), 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 \(L^p\) and in \(D\); 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 \(E\).

We will make the following assumption:

**Assumption 2.1.**

\[ B \in L^\infty\left(0, T; C_b^{2,\alpha}(E, E)\right) \]

for some \(\alpha \in (0, 1)\), where we have denoted by \(C_b^{2,\alpha}(E, E)\) the space of twice Fréchet differentiable functions \(\varphi\) from \(E\) to \(E\), bounded with their first and second differentials, such that \(x \mapsto D^2\varphi(x)\) is \(\alpha\)-Hölder continuous from \(E\) to \(L(E, E; E)\) (the space of bilinear forms on \(E\)). The \(L^\infty\) property in time means that the differentials are measurable in \((t, x)\) and both the function, the two differentials and the Hölder parameters 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 continuity modulus.

**Theorem 2.2.** Equation (13) 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 \([t_0, T] \ni t \to Y_{t_0-y}(\omega, t) \in E\) which satisfies identity (13’)

\[ Y_{t_0-y}(\omega, t) = e^{(t-t_0)A}y + \int_{t_0}^t e^{(t-s)A}B(s, Y_{t_0-y}(\omega, s)) \, ds + \int_{t_0}^t e^{(t-s)A}\sqrt{\Sigma} \, d\beta(\omega, s). \] (13’)

Such a function is continuous if \(E = L^p\); it is only in \(L^\infty\) if \(E = D\).

**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 \(D\) is due to the fact that the semigroup \(e^{tA}\) is not strongly continuous in \(D\).

**Theorem 2.3.** For almost every \(\omega \in \Omega\), 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 \(y \mapsto D^2Y_{t_0-y}(t, \omega)\) is \(\alpha\)-Hölder continuous from \(E\) to \(L(E, E; E)\).

**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 [DPZ92]. 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. □

In section 4 we will need the notion of continuity modulus 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 $R : E → L (E, E; ℝ)$, we define its continuity modulus

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

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

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

for every $x, x_0 ∈ E$. Indeed,

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

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

$$|r_v(x, x_0)| = \left| (D^2v(ξ_{v, x, x_0}) - D^2v(x_0))(x - x_0, x - x_0) \right|$$

$$\leq \|D^2v(ξ_{v, x, x_0}) - D^2v(x_0)\|_{L(E^2, ℝ)} \|x - x_0\|_E^2$$

$$\leq \omega (D^2v, \|x - x_0\|_E) \|x - x_0\|_E^2.$$ 

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

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

then

$$\omega (D^2v, \|x - x_0\|_E) \leq M \|x - x_0\|_E^2$$

and thus

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

### 3 The Kolmogorov equation

In this and the following section 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 6.

Suppose for a moment we are working in a standard Hilbert-space setting, that is, in a space $H = ℝ × H$ where $H$ is a Hilbert space. Then (see again [DPZ92]) the backward Kolmogorov equation, for the unknown $u : [0, T] × H → ℝ$, is

$$\frac{∂u}{∂t}(t, y) + \frac{1}{2} \text{Tr} (ΣD^2u(t, y)) + \langle Du(t, y), Ay + B(t, y) \rangle = 0, \quad u(T, ·) = Φ,$$ 

(22)

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

$$u(t, y) = E \left[ Φ \left( Y^t,w(T) \right) \right]$$ 

(23)
where \( Y^{t,y}(t) \) solves the associated SDE
\[
    dY(s) = [AY(s) + B(s, Y(s))] \, ds + \sqrt{\sum} \, d\beta(s), \quad s \in [t, T], \quad Y(t) = y
\]
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 \( \widehat{\mathcal{C}} \), but since \( B(t, y) \) belongs to \( \mathbb{R}^d \times \{0\} \not\subset \widehat{\mathcal{C}} \), in order to give meaning to the term \( (Du(t, y), B(t, y)) \) 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 \( \Phi \) to be defined at least on \( \cup_{s \in [t, T]} \mathcal{D}_{-t+s} \) in order for a function of the form \( (23) \) to be well defined. However the space \( \mathcal{D}_s \) 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 \( \mathcal{D}^\prime \) and \( \mathcal{D} \).

For the trace term, if we denote by \( e_1, \ldots, e_d \) an orthonormal basis of \( \mathbb{R}^d \) where \( \sigma \) diagonalizes, i.e. \( \sigma e_j = \sigma_j e_j \) for some real \( \sigma_j \) (in any of the spaces considered up to now), we could complete it to an orthonormal system \( \{e_n\} \) in \( \mathcal{H} \) obtaining that
\[
    \text{Tr} \left( \sum D^2 u(t, y) \right) = \sum_j \sigma_j^2 (D^2 u(t, y)e_j, e_j);
\]
hence, by analogy, also when working in \( \mathcal{D} \) we interpret the trace term as
\[
    \text{Tr} \left( \sum D^2 u(t, y) \right) = \sum_{j=1}^d \sigma_j^2 D^2 u(t, y)(e_j, e_j). \tag{24}
\]
Moreover we consider Kolmogorov equation in its integrated form with respect to time, that is, given a (sufficiently regular; see below) real function \( \Phi \) on \( \mathcal{D} \) we seek for a solution of the PDE
\[
    u(t, y) - \Phi(y) = \int_t^T (Du(s, y), Ay + B(s, y)) \, ds + \frac{1}{2} \int_t^T \sum_{j=1}^d \sigma_j^2 D^2 u(s, y)(e_j, e_j) \, ds. \tag{25}
\]

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, some results in this direction are appearing in the literature recently).

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

All these observations lead to our definition of solution to \( (25) \); first we say that a functional \( u \) on \([0, T] \times \mathcal{D} \) belongs to
\[
    L^\infty \left( 0, T; C_b^{2,\alpha} (\mathcal{D}, \mathbb{R}) \right)
\]
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 \( \alpha \)-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 parameters 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 final condition \( \Phi \) if
\[
    u \in L^\infty \left( 0, T; C_b^{2,\alpha'} (\mathcal{D}, \mathbb{R}) \right) \cap C \left( [0, T] \times \mathcal{D}, \mathbb{R} \right)
\]
for some \( \alpha' \in (0, 1) \), and satisfies identity \( (25) \) for every \( t \in [0, T] \) and \( y \in \text{Dom}(A_{-c}) \), with the duality terms understood with respect to the topology of \( \mathcal{D} \).
It will be clear in section 5 that the restriction \( y \in \text{Dom} \left( A^<_C \right) \) 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

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

solves equation (25).

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 \( L^p \), then we will show that if the problem is formulated in \( D \) it is possible to approximate it with a sequence of \( L^p \) problems; the solutions to such approximating problems will be finally shown to converge to a function that solves Kolmogorov’s backward PDE in the sense of definition 3.1.

All the above discussion about the meaning of Kolmogorov’s equation applies verbatim to the space \( L^p \). A solution in \( L^p \) is defined in a straightforward way as follows:

**Definition 3.2.** Given \( \Phi \in C^{2,0}_{b} (L^p, \mathbb{R}) \), we say that \( u : [0, T] \times L^p \rightarrow \mathbb{R} \) is a solution of the Kolmogorov equation in \( L^p \) with final condition \( \Phi \) if

\[
  u \in L^{\infty} \left( 0, T; C^{2,\alpha'}_{b} (L^p, \mathbb{R}) \right) \cap C \left( [0, T] \times L^p, \mathbb{R} \right)
\]

for some \( \alpha' \in (0, 1) \), and satisfies identity (25) for every \( t \in [0, T] \) and \( y \in \text{Dom} (A) \), with the duality terms understood with respect to the topology of \( L^p \).

4 **Solution in \( L^p \)**

The choice to work in a general \( L^p \) space instead of working with the Hilbert space \( L^2 \) could seem unjustified at first sight. As long as solving Kolmogorov’s equation in \( L^p \) is only a step towards solving it in \( D \) through approximations it would be enough to develop the theory in \( L^2 \), where the results needed are well known. Nevertheless we give and prove here this more general statement for \( L^p \) 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 \( L^2 \): recall for instance example (i) with its “delay”-version

\[
  B \left( t, \left( \frac{x}{\varphi} \right) \right) = \left( \tilde{b}(t, (\frac{x}{\varphi})), 0 \right)
\]

where \( \tilde{b} \) is given in (i)’. The second Fréchet derivative of \( B \) with respect to \( y = (\frac{x}{\varphi}) \) is simply

\[
  D^2 B \left( t, y \right) \left( \left( \frac{x_1}{\psi_{1}}, \frac{x_2}{\psi_{2}} \right) \right) = \left( \int_0^t g''(\varphi(s-t))\psi(s-t)\chi(s-t) \, ds \right)
\]

which is not uniformly continuous in \( (\frac{x}{\varphi}) \) in \( L^2 \), since for \( z = (\frac{x}{\varphi}) \)

\[
  \| D^2 B(t, y) - D^2 B(t, z) \|_{L(L^2, L^2; L^2)} =
  \sup_{\| \chi \|, \| \psi \| \leq 1} \int_0^t \left| g''(\varphi(s-t) - \varphi_1(s-t)) \right| \cdot \| \psi(s-t) \| \cdot \| \chi(s-t) \| \, ds \leq \| g'' \| \sup_{\| \chi \|, \| \psi \| \leq 1} \int_0^t \varphi(s-t) - \varphi_1(s-t) \cdot \| \chi(s-t) \| \cdot \| \psi(s-t) \| \, ds
\]
can not be bounded in terms of \(\|\varphi - \varphi_1\|_{L^2}\) if \(\varphi, \varphi_1, \chi\) and \(\psi\) are only in \(L^2\), and uniform continuity is essential in our proof, as it is in all classical cases. \(D^2B\) is however uniformly continuous in \(L^4\). This shows that proving the result in \(L^p\) 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 \(E = L^p\), theorems 2.2, 2.3 and 2.4 yield that the SDE

\[
dY(s) = \left[AY(s) + B(s, Y(s))\right]ds + \sqrt{\Sigma}d\beta(s), \quad s \in [t, T], \quad Y(t) = y
\]

admits a unique mild solution \(Y^{t_0:y}(t)\) in \(L^p\) which is continuous in time, \(C^{2,\alpha}_0\) with respect to \(y\) and has the Markov property.

**Theorem 4.1.** Let \(\Phi : L^p \rightarrow \mathbb{R}\) be in \(C^{2,\alpha}_0\) and let assumption 2.1 hold in \(L^p\). Then the function

\[
u(t, y) := E\left[\Phi \left( Y^{t_0:y}(T) \right) \right], \quad (t, y) \in [0, T) \times L^p,
\]

is a solution of the Kolmogorov equation in \(L^p\) with final condition \(\Phi\).

**Proof.** Throughout this proof \(\|\cdot\|\) will denote the norm in \(L^p\) and \(\langle \cdot, \cdot \rangle\) will denote duality between \(L^p\) and \(L^q\), where \(\frac{1}{p} + \frac{1}{q} = 1\). Since \(\Phi \in C^{2,\alpha}_0(L^p, \mathbb{R})\), theorem 2.3 assures that the function \(\nu\) has the regularity properties required by the definition of solution. We have thus to show that it satisfies equation 2.1. 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

\[
u(t_0, y) = E\left[\nu(t_1, Y^{t_0:y}(t_1))\right]
\]

because

\[
E\left[\Phi \left( Y^{t_0:y}(T) \right) \right] = E\left[ E\left[ \Phi \left( Y^{t_0:y}(T) \right) \mid Y^{t_0:y}(t_1) \right] \right] = E\left[ E\left[ \Phi \left( \nu(t_1, T) \right) \right] \right] = E\left[ \nu(t_1, Y^{t_0:y}(t_1)) \right].
\]

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

\[
u(t_1, Y^{t_0:y}(t_1)) - \nu(t_1, e^{(t_1-t_0)A} y) = \left\langle D\nu \left( t_1, e^{(t_1-t_0)A} y \right), Y^{t_0:y}(t_1) - e^{(t_1-t_0)A} y \right\rangle
\]

\[
+ \frac{1}{2} D^2\nu \left( t_1, e^{(t_1-t_0)A} y \right) \left(Y^{t_0:y}(t_1) - e^{(t_1-t_0)A} y, Y^{t_0:y}(t_1) - e^{(t_1-t_0)A} y \right)
\]

\[
+ \frac{1}{2} \tau_u(t_1, \cdot) \left( Y^{t_0:y}(t_1), e^{(t_1-t_0)A} y \right)
\]

where

\[
\left| \tau_u(t_1, \cdot) \left( Y^{t_0:y}(t_1), e^{(t_1-t_0)A} y \right) \right| \leq \omega \left( D^2\nu \left( t_1, \cdot \right), \left\| Y^{t_0:y}(t_1) - e^{(t_1-t_0)A} y \right\| \right) \left\| Y^{t_0:y}(t_1) - e^{(t_1-t_0)A} y \right\|^2
\]

(for the definitions of \(\tau\) and \(\omega\) see Remark 2.5 in the appendix). Due to the \(C^{2,\alpha}_0(L^p, \mathbb{R})\)-property, uniform in time, we have

\[
\left| \tau_u(t_1, \cdot) \left( Y^{t_0:y}(t_1), e^{(t_1-t_0)A} y \right) \right| \leq M \left\| Y^{t_0:y}(t_1) - e^{(t_1-t_0)A} y \right\|^{2+\alpha}.
\]

Recall that

\[
Y^{t_0:y}(t_1) - e^{(t_1-t_0)A} y = F^{t_0} \left( t_1, Y^{t_0:y} \right) + Z^{t_0} \left( t_1 \right)
\]

\[
F^{t_0} \left( t_1, Y^{t_0:y} \right) = \int_{t_0}^{t_1} e^{(t_1-s)A} B \left( S, Y^{t_0:y} \left( s \right) \right) ds
\]
and

\[ E[Z^{t_0}(t_1)] = 0 \]
\[ E[\|Z^{t_0}(t_1)\|^4] \leq C_2^4 (t_1 - t_0)^2 \]
\[ \|F^{t_0}(t_1, Y^{t_0,y})\| \leq C_A \|B\|_{\infty, L^p} (t_1 - t_0) \]

Hence, recalling \( u(0,0) = E[u(t_1, Y^{t_0,y}(t_1))] \),

\[ u(t_0,y) - u(t_1, e^{(t_1-t_0)A} y) = \]
\[ = \left\langle D u \left( t_1, e^{(t_1-t_0)A} y \right), E \left[ F^{t_0}(t_1, Y^{t_0,y}) \right] \right\rangle \]
\[ + \frac{1}{2} E \left[ D^2 u \left( t_1, e^{(t_1-t_0)A} y \right) \right. \]
\[ \left. \left( F^{t_0}(t_1, Y^{t_0,y}) + Z^{t_0}(t_1) \right) \right) \]
\[ + \frac{1}{2} E[r_{u(t_1,\cdot)}(Y^{t_0,y}(t_1), e^{(t_1-t_0)A} y)] \].

**Step 2.** Now let us explain the strategy. Given \( t \in [0,T] \), taken a sequence of partitions \( \pi_n \) of \([t,T]\) of the form \( t = t^n_1 \leq \ldots \leq t^n_{k_n+1} = T \) of \([t,T]\) with \( |\pi_n| \to 0 \), we take \( t_0 = t^n_0 \) and \( t_1 = t^n_{1+1} \) in the previous identity and sum over the partition \( \pi_n \) to get

\[ u(t,y) = \Phi(y) + I^{(1)}_n + I^{(2)}_n + I^{(3)}_n + I^{(4)}_n \]

where

\[ I^{(1)}_n := \sum_{i=1}^{k_n} \left( u(t^{n}_{i+1}, y) - u(t^{n}_{i+1}, e^{(t^{n}-t^{n}_i)A} y) \right) \]
\[ I^{(2)}_n := \sum_{i=1}^{k_n} \left\langle D u \left( t^{n}_{i+1}, e^{(t^{n}-t^{n}_i)A} y \right), E \left[ F^{t_1}(t^{n}_{i+1}, Y^{t^{n}_i,y}) \right] \right\rangle \]
\[ I^{(3)}_n := \frac{1}{2} \sum_{i=1}^{k_n} E \left[ D^2 u \left( t^{n}_{i+1}, e^{(t^{n}-t^{n}_i)A} y \right) \right. \]
\[ \left. \left( F^{t_1}(t^{n}_{i+1}, Y^{t^{n}_i,y}) + Z^{t_1}(t^{n}_{i+1}) \right) \right) \]
\[ I^{(4)}_n := \frac{1}{2} \sum_{i=1}^{k_n} E \left[ r_{u(t^{n}_{i+1}, \cdot)} \left( Y^{t^{n}_i,y}(t^{n}_{i+1}), e^{(t^{n}-t^{n}_i)A} y \right) \right] \].

We want to show that

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

2. \( \lim_{n \to \infty} I^{(2)}_n = \int_t^T \langle Du(s,y), B(s,y) \rangle \, ds \),

3. \( \lim_{n \to \infty} I^{(3)}_n = \frac{1}{2} \int_t^T \sum_{j=1}^d \sigma^2_j D^2 u(s,y) (e_j, e_j) \, ds \),

4. \( \lim_{n \to \infty} I^{(4)}_n = 0 \).
Step 3. We have, for \(y \in \text{Dom} (A)\) (in this case \(\frac{d}{dt} e^{tA} y = Ae^{tA} y\))

\[
\sum_{i}^{k_n} u \left( t^{n}_{i+1}, y \right) - u \left( t^{n}_{i+1}, e^{(s_i - t^n_i)A} y \right) = \sum_{i}^{k_n} \int_{t^{n}_{i+1}}^{t^{n}_{i+1} + t^n_i} \langle Du \left( t^{n}_{i+1}, e^{(s-t^n_i)A} y \right), Ae^{sA} y \rangle \, ds
\]

\[
= - \sum_{i}^{k_n} \int_{t^{n}_{i+1}}^{t^{n}_{i+1} + t^n_i} \langle Du \left( t^{n}_{i+1}, e^{(s-t^n_i)A} y \right), Ae^{(s-t^n_i)A} y \rangle \, ds
\]

\[
= - \int_{t}^{T} \sum_{i}^{k_n} \langle Du \left( t^{n}_{i+1}, e^{(s-t^n_i)A} y \right), Ae^{(s-t^n_i)A} y \rangle \mathbb{1}_{[t^n_i, t^n_{i+1}]}(s) \, ds
\]

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

Step 4. The function \(\mathcal{B}\) By standard properties of the Bochner integral we have

\[
\sum_{i}^{k_n} \int_{t^{n}_{i+1}}^{t^{n}_{i+1} + t^n_i} \langle Du \left( t^{n}_{i+1}, e^{(s-t^n_i)A} y \right), e^{(t^n_i-s)A} \mathcal{B} \left( s, Y^{t^n_i}(s) \right) \rangle \, ds
\]

\[
= \sum_{i}^{k_n} \int_{t^{n}_{i+1}}^{t^{n}_{i+1} + t^n_i} \langle Du \left( t^{n}_{i+1}, e^{(s-t^n_i)A} y \right), e^{(t^n_i-s)A} \mathcal{B} \left( s, Y^{t^n_i}(s, \omega) \right) \rangle \, ds
\]

\[
= \mathbb{E} \int_{t}^{T} \sum_{i}^{k_n} \langle Du \left( t^{n}_{i+1}, e^{(s-t^n_i)A} y \right), e^{(t^n_i-s)A} \mathcal{B} \left( s, Y^{t^n_i}(s, \omega) \right) \rangle \mathbb{1}_{[t^n_i, t^n_{i+1}]}(s) \, ds
\]

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:

\[
D^2 u \left( t^{n}_{i+1}, e^{(t^n_i-t^n_{i-1})A} y \right) \left( F^{t^n_i} (t^{n}_{i+1}, Y^{t^n_i}(s)) + Z^{t^n_i} (t^{n}_{i+1}), F^{t^n_i} (t^{n}_{i+1}, Y^{t^n_i}(s)) + Z^{t^n_i} (t^{n}_{i+1}) \right)
\]

\[
= D^2 u \left( t^{n}_{i+1}, e^{(t^n_i-t^n_{i-1})A} y \right) \left( F^{t^n_i} (t^{n}_{i+1}, Y^{t^n_i}(s)) + F^{t^n_i} (t^{n}_{i+1}, Y^{t^n_{i-1}}(s)) \right) +
\]

\[
D^2 u \left( t^{n}_{i+1}, e^{(t^n_i-t^n_{i-1})A} y \right) \left( Z^{t^n_i} (t^{n}_{i+1}), F^{t^n_i} (t^{n}_{i+1}, Y^{t^n_{i-1}}(s)) \right) +
\]

\[
D^2 u \left( t^{n}_{i+1}, e^{(t^n_i-t^n_{i-1})A} y \right) \left( Z^{t^n_i} (t^{n}_{i+1}), Z^{t^n_i} (t^{n}_{i+1}) \right)
\]

Let us give the main estimates. We have

\[
\mathbb{E} \left[ D^2 u \left( t, e^{(t-t_0)A} y \right) \left( F^{t_0} (t, Y^{t_0}(s)) + Z^{t_0} (t) \right) \right] \leq \left\| D^2 u \right\|_{\infty, L^p} \mathbb{E} \left[ \left\| F^{t_0} (t, Y^{t_0}(s)) \right\|^2 \right]
\]

\[
\leq \left\| D^2 u \right\|_{\infty, L^p} C_A^2 \mathbb{E} \left[ B \right]^2 \mathbb{E} \left[ \left\| Z^{t_0} (t) \right\|^2 \right]^{1/2}
\]

and

\[
\mathbb{E} \left[ D^2 u \left( t, e^{(t-t_0)A} y \right) \left( F^{t_0} (t, Y^{t_0}(s)) + Z^{t_0} (t) \right) \right] \leq \left\| D^2 u \right\|_{\infty, L^p} \mathbb{E} \left[ \left\| F^{t_0} (t, Y^{t_0}(s)) \right\|^2 \right]^{1/2} \mathbb{E} \left[ \left\| Z^{t_0} (t) \right\|^2 \right]^{1/2}
\]

hence the first three terms give no contribution when summing up over \(i\), because they are estimated by a power of \(t^{n}_{i+1} - t_i\) greater than 1. Therefore it remains to show that the term

\[
\sum_{i}^{k_n} \mathbb{E} \left[ D^2 u \left( t^{n}_{i+1}, e^{(t^{n}_{i+1}-t^n_i)A} y \right) \left( Z^{t^n_i} (t^{n}_{i+1}), Z^{t^n_i} (t^{n}_{i+1}) \right) \right]
\]

(26)
converges to $\int_t^s \sigma^2 D^2 u(s, y)(e, e) \, ds$. To this aim we recall that

$$Z^n_{t_{i+1}}(t^n_{i+1}) = \int_{t^n_i}^{t^n_{i+1}} e^{(t^n_{i+1}-t^n_i)A} \left( \sigma \, dW(r) \right)$$

$$= \left( \sigma \left( W \left( (t^n_{i+1} + t^n_i) \right) - W(t^n_i) \right) \right)$$

$$= \sigma \left( Z^n_0 \right).$$

We split again (26) into

$$\sum_{i=1}^{k_n} E \left[ D^2 u \left( t^n_{i+1}, e^{(t^n_{i+1}-t^n_i)A} y \right) \left( \left( \begin{array}{c} Z^n_0 \\ 0 \end{array} \right), \left( \begin{array}{c} Z^n_i \\ 0 \end{array} \right) \right) \right]$$

$$+ D^2 u \left( t^n_{i+1}, e^{(t^n_{i+1}-t^n_i)A} y \right) \left( \left( \begin{array}{c} Z^n_0 \\ 0 \end{array} \right), \left( \begin{array}{c} Z^n_i \\ 0 \end{array} \right) \right)$$

$$+ D^2 u \left( t^n_{i+1}, e^{(t^n_{i+1}-t^n_i)A} y \right) \left( \left( \begin{array}{c} 0 \\ Z^n_0 \end{array} \right), \left( \begin{array}{c} 0 \\ Z^n_i \end{array} \right) \right)$$

$$+ D^2 u \left( t^n_{i+1}, e^{(t^n_{i+1}-t^n_i)A} y \right) \left( \left( \begin{array}{c} 0 \\ Z^n_0 \end{array} \right), \left( \begin{array}{c} 0 \\ Z^n_i \end{array} \right) \right).$$

For the first term we have, using Itô isometry, that

$$\sum_{i=1}^{k_n} E \left[ D^2 u \left( t^n_{i+1}, e^{(t^n_{i+1}-t^n_i)A} y \right) \left( \left( \begin{array}{c} Z^n_0 \\ 0 \end{array} \right), \left( \begin{array}{c} Z^n_i \\ 0 \end{array} \right) \right) \right]$$

$$= \sum_{j=1}^{d} \sigma_j^2 \sum_{i=1}^{k_n} D^2 u \left( t^n_{i+1}, e^{(t^n_{i+1}-t^n_i)A} y \right) \left( e_j, e_j \right) \left( t^n_{i+1} - t^n_i \right)$$

and the right-hand side in this equation converges to $\sum_{j=1}^{d} \sigma_j^2 \int_t^s D^2 (s, y)(e_j, e_j) \, ds$ thanks to the strong continuity of $e^{tA}$.

For the second term we can write

$$E \left[ D^2 u \left( t^n_{i+1}, e^{(t^n_{i+1}-t^n_i)A} y \right) \left( \left( \begin{array}{c} Z^n_0 \\ 0 \end{array} \right), \left( \begin{array}{c} Z^n_i \\ 0 \end{array} \right) \right) \right] \leq \| D^2 u \|_{\infty, \mathcal{L}^p} E \left[ \left| W \left( t^n_{i+1} \right) - W \left( t^n_i \right) \right| \right]$$

$$\leq \| D^2 u \|_{\infty, \mathcal{L}^p} \left( \int_0^{t^n_{i+1} - t^n_i} |W(r)|^p \, dr \right)^{\frac{1}{p}}$$

$$\leq \| D^2 u \|_{\infty, \mathcal{L}^p} \left( \int_0^{t^n_{i+1} - t^n_i} |W(r)| \, dr \right)^{\frac{1}{p}}$$

$$\leq \| D^2 u \|_{\infty, \mathcal{L}^p} \left( t^n_{i+1} - t^n_i \right)^{\frac{1}{2}} \left( t^n_i \right)^{\frac{1}{2}}$$

$$\leq \| D^2 u \|_{\infty, \mathcal{L}^p} \left( t^n_{i+1} - t^n_i \right)^{\frac{1}{p} + \frac{1}{2}},$$

using Itô isometry and Burkholder-Davis-Gundy inequality, thus it converges to zero when summing over $i$ and letting $n$ go to $\infty$.

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

$$E \left[ D^2 u \left( t^n_{i+1}, e^{(t^n_{i+1}-t^n_i)A} y \right) \left( \left( \begin{array}{c} 0 \\ Z^n_0 \end{array} \right), \left( \begin{array}{c} 0 \\ Z^n_i \end{array} \right) \right) \right] \leq \left( t^n_{i+1} - t^n_i \right)^{1 + \frac{1}{p}}.$$
hence it follows that also this term gives no contribution when passing to the limit.

**Step 6.** Since
\[ \left| r_{u(t, \cdot)} \left( Y^{t_0, y}(t), e^{(t-t_0)A}y \right) \right| \leq M \left\| Y^{t_0, y}(t) - e^{(t-t_0)A}y \right\|^{2+\alpha} \]
we have that
\[ \left| \mathbb{E} \left[ r_{u(t, \cdot)} \left( Y^{t_0, y}(t), e^{(t-t_0)A}y \right) \right] \right| \leq M \mathbb{E} \left[ \left\| Y^{t_0, y}(t) - e^{(t-t_0)A}y \right\|^{2+\alpha} \right] \]
\[ \leq C \left( \mathbb{E} \left[ \left\| Y^{t_0, y} \right\|^{4} \right] + \mathbb{E} \left[ \left\| Z^{t_0} (t) \right\|^{4} \right] \right)^{\frac{2+\alpha}{4}} \]
\[ \leq C (t - t_0)^{1+\frac{\alpha}{2}} \]
and from this one proves that \( \lim_{n \to \infty} I_n^{(4)} = 0 \). \qed

**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], can not be carried out when working with the sup-norm: if we start again from [27] using the norm of \( \mathcal{D} \) we would end up with the estimate
\[ \mathbb{E} \left| D^2 u \left( t_{i+1}^{-}, e^{(\bar{c}_{i+1} - t_{i}^{-})A}y \right) \left( \left( \frac{z_{i}^{\cdot}}{\sigma_{i}^{\cdot}} \right), \left( \frac{0}{z^{\cdot}_{i}} \right) \right) \right| \leq \left\| D^2 u \right\|_{\infty, \mathcal{L}^p} (t_{i+1}^{n} - t_{i}^{n}) \] (28)
which is not enough to obtain the convergence to 0 that we need.

## 5 Solution in \( \hat{\mathcal{C}} \)

We now show how to use \( \mathcal{L}^p \) approximations in order to obtain classical solutions of Kolmogorov’s equations in the sense of definition [3]. As in the previous we will assume that \( B \) satisfied assumption [2.1] for \( E = \mathcal{D} \), that is
\[ B \in \mathcal{L}^{\infty} \left( 0, T; C^{2+\alpha}_{L_b} (\mathcal{D}, \mathcal{D}) \right) \]
for some \( \alpha \in (0, 1) \). Suppose we have a sequence \( \{ J_n \} \) of linear continuous operators from \( \mathcal{L}^p(-T, 0; \mathbb{R}^d) \) into \( \mathcal{C}([-T, 0]; \mathbb{R}^d) \) such that \( J_n \varphi \xrightarrow{n \to \infty} \varphi \) uniformly for any \( \varphi \in \mathcal{C}([-T, 0]; \mathbb{R}^d) \). By Banach-Steinhaus theorem we have that \( \sup \left\| J_n \right\|_{\mathcal{L}(\mathcal{C}, \mathcal{C})} < \infty \); however we need a slightly stronger property, namely that
\[ \left\| J_n f \right\|_{\infty} < C_n \left\| f \right\|_{\infty} \]
for all \( f \) with at most one jump, uniformly in \( n \). Then we can define the sequence of operators
\[ B_n : [0, T] \times \mathcal{L}^p \to \mathbb{R}^d \times \{0\} \]
\[ B_n(t, y) = B_n(t, (\varphi)) = B_n(t, x, \varphi) : = B(t, x, J_n \varphi) \]. (29)
We will often write \( J_n (\varphi) \) for \( (j_{n, \varphi}) \) 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
\[ dY_n(t) = AY_n(t) \, dt + B_n(t, Y_n(t)) \, dt + \sqrt{\sigma} \, d\tilde{\beta}(t), \quad Y_n(0) = y \in \mathcal{L}^p \] (30)
by theorem [2.2] it admits a unique path by path mild solution \( Y_{n, y}^{s, y} \) such that, thanks to theorem [2.3] the map \( t \mapsto Y_{n, y}^{s, y}(t) \) is in \( C^{2+\alpha}_{L_b} \). Suppose also we are given a final condition \( \Phi : \mathcal{D} \to \mathbb{R} \) for the backward Kolmogorov equation [25] associated to the original problem with \( B \); approximations \( \Phi_n \) can be defined in the exact same way. We have then a sequence of approximated backward Kolmogorov’s equations in \( \mathcal{L}^p \), namely
\[ u_n (t, y) - \Phi (y) = \int_t^T (D u_n (s, y), A y + B_n (s, y)) \, ds + \frac{1}{2} \int_t^T \sum_{j=1}^d \sigma_j^2 D^2 u_n (s, y) (e_j, e_j) \, ds \] (31)
with final condition \( u_n(T, \cdot) = \Phi_n \). Theorem 4.1 yields in fact that for each \( n \) the function
\[
u_n(s, y) = \mathbb{E} [\Phi_n (Y^{\alpha,y}_n(T))] \tag{32}
\]
is a solution to equation (31) in \( L^p \). If we choose the initial condition \( y \) in the space \( \hat{C} \), then \( Y^{\alpha,y}_n(t) \in \hat{C} \) as well for every \( n \) and every \( t \in [s, T] \) and therefore we have uniform convergence of \( J_n Y^{\alpha,y}_n \) to \( Y^{\alpha,y} \).

An example of a sequence \( \{ J_n \} \) satisfying the required properties can be constructed as follows: for any \( \varepsilon \in (0, \frac{1}{2}) \) define a function \( \alpha_\varepsilon : [-T, 0] \to [-T, 0] \) as
\[
\alpha_\varepsilon(x) = \begin{cases} 
-T + \varepsilon & \text{if } x \in [-T, -T + \varepsilon] \\
x & \text{if } x \in [-T + \varepsilon, -\varepsilon] \\
-\varepsilon & \text{if } x \in [-\varepsilon, 0].
\end{cases}
\]
Then choose any \( C^\infty(\mathbb{R}; \mathbb{R}) \) function \( \rho \) such that \( \|\rho\|_1 = 1, 0 \leq \rho \leq 1 \) and \( \text{supp}(\rho) \subseteq [-1, 1] \) and define a sequence \( \{\rho_n\} \) of mollifiers by \( \rho_n(x) := n \rho(nx) \). Finally, set, for any \( \varphi \in L^1(-T, 0; \mathbb{R}^d) \)
\[
J_n \varphi(x) := \int_{-T}^{0} \rho_n (\alpha_\varepsilon(x) - y) \varphi(y) \, dy. \tag{33}
\]
We will need one further assumption, together with the required properties for \( J_n \) that we write again for future reference.

**Assumption 5.1.** 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 \( J_n \varphi \xrightarrow{n \to \infty} \varphi \) uniformly for any \( \varphi \in C([-T, 0]; \mathbb{R}^d) \) and \( \sup_n \|J_n \varphi\|_\infty < C_\varphi \|\varphi\| \) for every \( \varphi \) that has at most one jump and is continuous elsewhere.

The drift \( B \) and the final condition \( \Phi \) are such that for any \( s \in [-T, 0] \), any \( r \geq s \), any \( y \in \hat{C} \) and for almost every \( a \in [-T, 0] \) the following hold:
\[
DB(r, y) J_n \left( \frac{1}{1_{[a,0)}} \right) \longrightarrow DB(r, y) \left( \frac{1}{1_{[a,0)}} \right);
\]
\[
\langle D\Phi(y), J_n \left( \frac{1}{1_{[a,0)}} \right) \rangle \longrightarrow \langle D\Phi(y), \left( \frac{1}{1_{[a,0)}} \right) \rangle;
\]
\[
D^2\Phi(y) \left( J_n \left( \frac{1}{1_{[a,0)}} \right) - \left( \frac{1}{1_{[a,0)}} \right) \right) \longrightarrow 0;
\]
\[
D^2\Phi(y) \left( \left( \frac{1}{1_{[a,0)}} \right), J_n \left( \frac{1}{1_{[a,0)}} \right) - \left( \frac{1}{1_{[a,0)}} \right) \right) \longrightarrow 0;
\]
\[
D^2\Phi(y) \left( \left( \frac{1}{1_{[a,0)}} \right), J_n \left( \frac{1}{1_{[a,0)}} \right) \right) \longrightarrow 0.
\]

**Remark 5.2.** Assumption 5.1 implies that the same set of properties holds if we substitute \( \left( \frac{1}{1_{[a,0)}} \right) \) with any element \( q = \left( \psi(0) \right)_{\frac{1}{\psi}} \in D_{-a}, \) that is, it has at most one jump and no other discontinuities; this happens because any such \( \psi \) is the sum of a continuous function and an indicator function, and all the derivatives appearing in the above assumptions are linear.

**Remark 5.3.** The infinite dimensional operators associated to examples (7) and (11) in section 1 through (15) satisfy this assumption if we choose \( J_n \) as in (33). Example (6) does not satisfy assumption 5.1 for every \( a \) but only for \( a \neq t_j, n = 1, \ldots, n \).

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

**Theorem 5.4.** Let \( \Phi \in C^{2,\alpha}(D, \mathbb{R}) \) be given and let assumption 2.1 hold for \( E = D \). Under assumption 5.1 the function \( u : [0, T] \times D \to \mathbb{R} \) given by
\[
u(t, y) = \mathbb{E} \left[ \Phi \left( Y^{\alpha,y}(T) \right) \right], \tag{34}
\]
where $Y^{s,y}$ is the solution to equation \[31\] in $\mathcal{D}$, is a solution of the Kolmogorov equation with final condition $\Phi$, that is, for every $(t, y) \in [0, T] \times \text{Dom} \left( A_{\mathcal{C}} \right)$ it satisfies identity

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

**Proof.** Let $B_n$, $\Phi_n$, $Y_n$ and $u_n$ be as above. The proof will be divided into some steps that will prove the following: for $y \in \text{Dom} \left( A_{\mathcal{C}} \right)$

- $Y_n^{s,y}(t) \to Y^{s,y}(t)$ in $\widehat{\mathcal{C}}$ for every $t$ uniformly in $\omega$;
- $u_n(s, y) \to u(s, y) = E[\Phi(Y^{s,y}(T))]$ for every $s$ pointwise in $y$;
- equation \[31\] converges to equation \[25\] for any $t \in [0, T]$.

**Step 1** We first need to compute

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

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

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

(35)

(36)

For the term (35) recall that

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

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 $B$

$$
B(r, J_n Y^{s,y}(r)) \to B(r, Y^{s,y}(r))
$$

(37)

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

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

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

(38)

for $n$ big enough. Consider now (36):

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

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

$$
\leq C \int_s^t K_B \| Y_n^{s,y}(r) - Y^{s,y}(r) \| \, dr.
$$
because, for any $\psi \in C$, $\|J_n\psi\|_\infty \leq C_J \|\psi\|_\infty$ and therefore $\|J_n y\| \leq C_J \|y\|$. Hence this and (38) yield, by Gronwall’s lemma,

$$\|Y_{n}^{s,y}(t) - Y_{\infty}^{s,y}(t)\|_C \leq e^{TC_K B}$$

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

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

$$\left| u_n(s, y) - u(s, y) \right| \leq E \left| \Phi_n (Y_{n}^{s,y}(T)) - \Phi_n (Y_{\infty}^{s,y}(T)) \right| + E \left| \Phi_n (Y_{n}^{s,y}(T)) - \Phi (Y_{\infty}^{s,y}(T)) \right|$$

and for almost any $\omega$

$$\left| \Phi_n (Y_{n}^{s,y}(T)) - \Phi (Y_{\infty}^{s,y}(Y)) \right| \leq K \left( \|Y_{n}^{s,y}(T) - Y_{\infty}^{s,y}(Y)\| \right)$$

and

$$\left| \Phi_n (Y_{\infty}^{s,y}(T)) - \Phi (Y_{\infty}^{s,y}(T)) \right| \leq K \left( \|J_n Y_{\infty}^{s,y}(T) - Y_{\infty}^{s,y}(T)\| \right),$$

both of which are arbitrarily small for $n$ large enough; now since $B$ is bounded and we assumed that $E \|Z\|^4$ is finite, we can apply again the dominated convergence theorem (integrating in the variable $\omega$) 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 $R$, to which we associate the corresponding sequence $g_n = \left( \frac{\tilde{g}_n - \tilde{g}}{n} \right) \in C \subset D$. We remark here that the duality $D' \langle Du_n(s, y), g_n \rangle_D$ is well defined and equals $D' \langle Du_n(s, y), g_n \rangle_D$; a simple proof of this fact is the following: $u_n$ is Fréchet differentiable both on $D$ and on $L^p$ and its Gâteaux derivatives along the directions in $D$ are of course the same in $D$ and in $L^p$, therefore also the Fréchet derivatives must be equal. Now

$$\left| D' \langle Du_n, g_n \rangle_D - D' \langle Du, g \rangle_D \right| = \left| \langle Du_n, g_n - g \rangle + \langle Du_n - Du, g \rangle \right| \leq \left| \langle Du_n - Du, g \rangle \right| + \left| \langle Du_n, g_n - g \rangle \right| = |A| + |B|.$$

We show that for $s, y$ fixed the set $\{ Du_n(s, y) \}_{n}$ is bounded in $D'$. From the definition of $u$ we have that for $h \in D$

$$\langle Du_n(s, y), h \rangle = \langle D\Phi_n (Y_{n}^{s,y}(T)), DY_{n}^{s,y}(T)h \rangle$$

and it is easily shown that

$$D\Phi_n(g) = D\Phi(J_n g)J_n$$

for any $g \in D$. $D\Phi$ is bounded by assumption, whereas by the required properties of $J_n$

$$\|J_n DY_{n}^{s,y}(T)h\|_D \leq C \|DY_{\infty}^{s,y}(T)h\|_D;$$

since the $\|DY_n\|$ are uniformly bounded by a constant depending only on $e^{tA}$ and on $DB$, we have that the $Du_n$’s are uniformly bounded as well and therefore $B \to 0$ as $g_n \to g$. The term $A$ requires some work: first write (suppressing indexes $s, y$ and $T$)

$$A = \langle D\Phi_n (Y_n), DY_n g \rangle - \langle D\Phi(Y), DY g \rangle$$

$$= \langle D\Phi_n (Y_n), (DY_n - DY)g \rangle + \langle D\Phi_n (Y_n) - D\Phi(Y), DY g \rangle = A_1 + A_2;$$

$$A_2 = \langle D\Phi_n (Y_n) - D\Phi(Y), DY g \rangle + \langle D\Phi_n (Y_n) - D\Phi(Y), DY g \rangle = A_{21} + A_{22}. $$

Since the Lipschitz constants of $D\Phi_n$ are uniformly bounded in $C$ we have that

$$\|A_{21}\| \leq \|D\Phi_n (Y_n) - D\Phi_n (Y)\|_D, \|DY g\|_D$$

$$\leq C \|Y_n - Y\| \|DY g\|$$
and the last line goes to zero as $n$ goes to infinity. For $A_{22}$ write

$$|A_{22}| = |(\langle \Phi(J_n^1) \rangle_{\nu_n}, DYg) - \langle \Phi(Y), DYg \rangle|$$

$$\leq |(\langle \Phi(J_n^1) \rangle_{\nu_n}, DYg) - \langle \Phi(Y)J_n, DYg \rangle|$$

$$|\langle \Phi(Y)J_n, DYg \rangle - \langle \Phi(Y), DYg \rangle|$$

$$\leq K_D \|J_n - Y\| \|DYg\| + |\langle \Phi(Y)J_n, DYg \rangle - \langle \Phi(Y), DYg \rangle|;$$

the first term goes to zero by properties of $J_n$, the second one thanks to assumption 5.1: this is because from the defining equation for $DY$ one easily sees that for any $\left( \frac{\alpha}{\alpha n} \right) \in C$ 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 $A_1$:

$$DY_n^{s,y}(T)g - DY^{s,y}(T)g =$$

$$\int_s^T e^{(T-r)A} \{ DB_n (r, Y_n^{s,y}(r)) - DB (r, Y^{s,y}(r)) \} \|DY^{s,y}(r)\| dr$$

$$\int_s^T e^{(T-r)A} \{ DB_n (r, Y_n^{s,y}(r)) - DB (r, Y^{s,y}(r)) \} \|DY^{s,y}(r)\| dr$$

and $A_{12}$ can be written as

$$A_{12} = \int_s^T e^{(T-r)A} \{ DB_n (r, Y_n^{s,y}(r)) - DB (r, Y^{s,y}(r)) \} \|DY^{s,y}(r)\| dr$$

whence

$$\|A_{12}\| \leq C \int_s^T \|DY^{s,y}(r)g\| \|DB (r, J_nY_n^{s,y}(r)) - DB (r, Y_n^{s,y}(r))\| dr$$

$$\leq C \int_s^T \|DY^{s,y}(r)g\| \|DB (r, \cdot)\| \|J_nY_n^{s,y}(r) - J_nY^{s,y}(r)\|$$

$$\leq C \int_s^T \|DY^{s,y}(r)g\| \|DB (r, \cdot)\| \|Y_n^{s,y}(r) - Y^{s,y}(r)\| dr$$

that goes to zero; for $A_{122}$

$$\|DB_n (r, Y_n^{s,y}(r)) - DB (r, Y^{s,y}(r))\|DYg\|$$

$$\leq \|DB (r, J_nY_n^{s,y}(r)) - DB (r, Y_n^{s,y}(r))\| \|J_nDY^{s,y}(r)g\|$$

$$+ \|DB (r, Y_n^{s,y}(r)) - DB (r, Y^{s,y}(r))\| \|DY^{s,y}(r)g\|$$

$$\leq K_D \|J_n - Y\| \|DY_n^{s,y}(r)g - DY^{s,y}(r)g\|$$

$$\leq K_D \|J_n - Y\| \|DY_n^{s,y}(r)g - DY^{s,y}(r)g\|$$

where the last line goes to zero thanks to assumption 5.1 again, and therefore $A_{122}$ goes to zero by the dominated convergence theorem. From (39) and this last argument it follows that for any fixed $\varepsilon > 0$

$$\|DY_n(T)g - DY(T)g\| \leq C \int_s^T \|DB_n\| \|DY_n^{s,y}(r)g - DY^{s,y}(r)g\| dr + \varepsilon$$

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 $(Du_n, B_n)$ to $(Du, B)$.

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

$$(Du_n(s,y), Ay),$$

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

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

since the $\sigma_j$’s are constants and the passage from one to $d$ dimensions is trivial. We will drop the subscript 1 in the computations to simplify notations. We can proceed as follows (suppressing again $s$, $y$ and $T$):

$$|D^2u_n(s,y)(e, e) - D^2u(s,y)(e, e)| \leq |D^2\Phi_n(Y_n)(DY_n e, DY_n e) - D^2\Phi(Y)(DY e, DY e)| + |\langle D\Phi_n(Y_n), D^2Y_n(e, e) \rangle - \langle D\Phi(Y), D^2Y(e, e) \rangle|$$

$$= |C| + |D|.$$

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

$$C = [D^2\Phi_n(Y_n)(DY_n e, DY_n e) - D^2\Phi_n(Y_n)(DY e, DY e)] + [D^2\Phi_n(Y_n)(DY e, DY e) - D^2\Phi(Y)(DY e, DY e)]$$

$$= C_1 + C_2.$$

For $C_1$ just write

$$|C_1| \leq \|D^2\Phi_n(Y_n)(DY_n e - DY e, DY_n e - DY e) + D^2\Phi_n(Y_n)(DY e, DY_n e - DY e) + D^2\Phi_n(Y_n)(DY_n e - DY_n e, DY e)\|$$

$$\leq \|D^2\Phi_n(Y_n)\| \left[\|DY_n e - DY e\|^2 + 2\|DY e\| \|DY_n e - DY e\|\right]$$

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 $n$). For $C_2$

$$C_2 = [D^2\Phi(J_n Y) - D^2\Phi(Y)] (J_n DY e, J_n DY e) + D^2\Phi(Y)(J_n DY e, J_n DY e) + D^2\Phi(Y)(J_n DY e, J_n DY e - DY e) + D^2\Phi(Y)(J_n DY e - DY e, DY e)$$

$$= [D^2\Phi(J_n Y) - D^2\Phi(Y)] (J_n DY e, J_n DY e) + D^2\Phi(Y)(J_n DY e, J_n DY e - DY e) + D^2\Phi(Y)(J_n DY e - DY e, DY e) + D^2\Phi(Y)(J_n DY e - DY e, J_n DY e - Dy e).$$

Last three terms go to zero by assumption [5.1] while the first one is bounded in norm by

$$C_2 \|D^2\Phi\|_{a_1} \|J_n Y - Y\|_a \|DY e\|^2$$

which goes to zero since $\|J_n Y - Y\| \to 0$. We now go on with $D$:

$$D = \langle D\Phi_n(Y_n), D^2Y_n(e, e) - D^2Y(e, e) \rangle + \langle D\Phi_n(Y_n) - D\Phi(Y), D^2Y(e, e) \rangle = D_1 + D_2$$

and $D_2$ is easy to handle since

$$|D_2| \leq |\langle D\Phi(Y_n) - D\Phi_n(Y), D^2Y(e, e) \rangle| + |\langle D\Phi_n(Y) - D\Phi(Y), D^2Y(e, e) \rangle|$$

$$\leq \|D\Phi_n(Y_n) - D\Phi(Y)\| \|D^2Y(e, e)\|.$$
where the first term is bounded by
\[ \|D\Phi_n\| \|Y_n - Y\| \|D^2Y(e, e)\| \]

and therefore goes to zero as for \(\Lambda_1\), and the second goes to zero since \(D^2Y(e, e)\) is in \(\mathcal{C}\) and \(D\Phi_n(y)\) converge to \(D\Phi(y)\) for any \(y\) as functionals on \(\mathcal{C}\). Let’s now rewrite the right-hand term in the bracket defining \(D_1\) as
\[
D^2Y_n(y)(T)(e, e) - D^2Y^{*, y}(T)(e, e) = \int_s^T \left[ D^2B_n(r, Y_n^{*, y}(r)) (DY_n^{*, y}(r)e, DY_n^{*, y}(r)e) + \right.
\]
\[ - D^2B(r, Y^{*, y}(r)) (DY^{*, y}(r)e, DY^{*, y}(r)e) \right] \, dr + \]
\[ + \int_s^T \left[ DB_n(r, Y_n^{*, y}(r)) D^2Y_n^{*, y}(r)(e, e) + \right.
\]
\[ - DB(r, Y^{*, y}(r)) D^2Y^{*, y}(r)(e, e) \right] \, dr
\]
\[ = D_{11} + D_{12}. \]

Proceeding in a way similar to before we write the integrand in \(D_{11}\) as a sum
\[ D^2B_n(Y_n)(DY\epsilon, DY\epsilon) - D^2B_n(Y_n)(DY\epsilon, DY\epsilon) + [D^2B_n(Y_n) - D^2B(Y)] (DY\epsilon, DY\epsilon) = D_{111} + D_{112} \]

and notice that
\[
\|D_{111}\| = \|D^2B_n(Y_n)(DY\epsilon - DY\epsilon, DY\epsilon - DY\epsilon) + D^2B_n(Y_n)(DY\epsilon - DY\epsilon, DY\epsilon) + \]
\[ + D^2B_n(Y_n)(DY\epsilon, DY\epsilon - DY\epsilon)\]
\[ \leq \|D^2B_n(Y_n)\| \|[DY\epsilon, DY\epsilon]\|^2 + 2\|DY\epsilon\| \|D^2Y\epsilon - DY\epsilon\| \]

which can be treated as in \(\Lambda_1\) since the norms \(\|D^2B_n(Y_n^{*, y}(r))\|\) are bounded uniformly in \(n\) and \(r\), and that \(D_{112}\) can be shown to go to zero pointwise in \(r\) thanks to assumption [5] and to the \(\alpha\)-Hölderianity of \(D^2B_n\) in the same way as for \(C_2\). 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
\[
DB_n(Y_n)D^2Y_n(e, e) - DB(Y)D^2Y(e, e) = \]
\[ = DB_n(Y_n) [D^2Y_n - D^2Y] (e, e) + \]
\[ + [DB_n(Y_n) - DB_n(Y)] D^2Y(e, e) + [DB_n(Y) - DB(Y)] D^2Y(e, e) \]
\[ = DB_n(Y_n) [D^2Y_n - D^2Y] (e, e) + [DB_n(Y_n) - DB_n(Y)] D^2Y(e, e) + \]
\[ + DB(J_nY) [J_nD^2Y(e, e) - D^2Y(e, e)] + [DB(J_nY) - DB(Y)] D^2Y(e, e). \]

The second term in last sum is bounded in norm by
\[ \|DB_n(r, \cdot)\| \|Y_n - Y\| \|D^2Y(e, e)\| \]

which goes to zero since \(Y_n \rightarrow Y\) 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 [11] to obtain that
\[ D^2Y_n^{*, y}(T)(e, e) - D^2Y^{*, y}(T)(e, e) \rightarrow 0. \]
This together with the uniform boundedness of $D\Phi_{\nu}(Y_t)$ finally yields the convergence to zero of $D$.

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

\[ \partial_t u(t,y) + (Du(t,y), Ay + B(t,y)) + \frac{1}{2}\sum_{j=1}^{d} \sigma_j^2 D^2 u(t,y)(e_j, e_j) \cdot u(T,\cdot) = \Phi. \]

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

\section{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 \cite{CF13}. For a function $\nu = \{\nu_t\}_{t \in [0,T]}$, $\nu_t : D([0,T]; \mathbb{R}^d) \rightarrow \mathbb{R}^d$ the $i$-th \textit{vertical derivative} at $\gamma_t$ ($i = 1, \ldots, d$) is defined as

\[ \mathcal{D}_i \nu_t(\gamma_t) = \lim_{h \rightarrow 0} \frac{\nu_t(\gamma_t + h e_i) - \nu_t(\gamma_t)}{h} \quad (42) \]

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

\[ \mathcal{D}_\gamma \nu_t(\gamma_t) = (\mathcal{D}_1 \nu_t(\gamma_t), \ldots, \mathcal{D}_d \nu_t(\gamma_t)); \]

higher order vertical derivatives are defined in a straightforward way. The \textit{horizontal derivative} at $\gamma_t$ is defined as

\[ \mathcal{D}_\gamma \nu(\gamma_t) = \lim_{h \rightarrow 0^+} \frac{\nu_{t+h}(\gamma_{t+h}) - \nu_t(\gamma_t)}{h} \quad (43) \]

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 \cite{2} - \cite{11}. To establish some relations between Fréchet derivatives of $B$ and horizontal and vertical derivatives of $b$ is much less obvious; some results are given by the following theorem.

\textbf{Theorem 6.1.} Suppose $u : [0, T] \times \mathcal{D} \rightarrow \mathbb{R}$ is given and define, for each $t \in [0, T]$, $u_t : D([0, t]; \mathbb{R}^d) \rightarrow \mathbb{R}$ as $u_t(\gamma) := u(t, \gamma(t), L_t \gamma)$, in the same way as in \cite{11}. Then the vertical derivatives of $u_t$ coincide with the partial derivatives of $u$ with respect to the second variable (i.e. the present state), that is,

\[ \mathcal{D}_i u(t, x, L_t \gamma) = \frac{\partial}{\partial x_i} u(t, x, L_t \gamma), \quad i = 1, \ldots, d \quad (44) \]

The same result holds true also if $u$ is given from $\nu$ as in \cite{10}. Furthermore let $\gamma_t \in C^1_b([0, t])$ and let again $u$ be given and define $\nu$ as above. Then

\[ \mathcal{D}_\nu \nu(\gamma_t) = \frac{\partial u}{\partial t}(t, \gamma(t), L_t \gamma_t) + (Du(t, \gamma(t), L_t \gamma_t), (L_t \gamma_t)_+'); \]

where $\langle \cdot, \cdot \rangle$ is the duality between $D$ and $D'$, $Du$ is the Fréchet derivative of $u$ with respect to $\varphi$ 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

\[
\mathcal{D}_t \nu_t(\gamma) = \lim_{h \to 0} \frac{1}{h} \left[ \nu_t(\gamma^h) - \nu_t(\gamma) \right]
\]

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

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

\[
= \frac{\partial}{\partial x_i} u(t, x, L_t \gamma)
\]

This proves the first part of the theorem.

For the second part suppose first that there is no explicite dependence on \( t \) in \( u \). Then

\[
\mathcal{D}_t b(\gamma_t) = \lim_{h \to 0} \frac{1}{h} \left[ u(\gamma_{t,h}(t), L_t \gamma_{t,h}) - u(\gamma_t(t), L_t \gamma) \right]
\]

\[
= \lim_{h \to 0} \frac{1}{h} \left[ u(\gamma_t(t), L_t \gamma_{t,h}) - u(\gamma_t(t), L_t \gamma) \right]
\]

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

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

Last line can be written as

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

where

\[ N_{t,h} \gamma_t(s) = \begin{cases} 0 & \begin{cases} -T, -t - h \end{cases} \\ \gamma_t(t + h + s) - \gamma_t(0) & \begin{cases} -T, -t - h \end{cases} \\ \gamma_t(t + h + s) - \gamma_t(t + s) & \begin{cases} -T, -t - h \end{cases} \\ \gamma_t(t) - \gamma_t(t + s) & \begin{cases} -h, 0 \end{cases} \end{cases} \] (46)

\[ N_{t,h} \gamma_t \] 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 \( s \in [-T, -t) \) \( \exists \tilde{h} \) s.t. \( s < -t - \tilde{h} \), hence \( N_{t,h} \gamma(s) = 0 \forall h < \tilde{h} \) 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 \( \frac{1}{h} N_{t,h} \gamma(-t) \to \left( \frac{d}{dx} L_t \gamma \right)(-t) = (L_t \gamma)'(-t) = \gamma'_+(0) \);

(iii) for \( s \in (-t, 0) \) \( \exists \tilde{h} \) s.t. \( s < -\tilde{h} < 0 \), hence \( \frac{1}{h} N_{t,h} \gamma(s) = \frac{1}{h} [\gamma_t(t + s + h) - \gamma_t(t + s)] \to \gamma'_+(t + s) = \gamma'(t + s) = (L_t \gamma_t)'(s) \);

Therefore

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

and, since \( \gamma \in \mathcal{C}^1_b \),

\[ (L_t \gamma_t)'_+(s) = (L_t \gamma_t)'(s) \quad \forall s \neq -t. \]
Again since $\gamma_t \in \mathcal{C}^1$ with bounded derivative $\frac{1}{h} N_{t,h,\gamma_t}$ converges to $(L_t \gamma_t)_+'$ also uniformly. Keeping into account (45) and the definition of Fréchet derivative, one gets

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

$$= \lim_{h \to 0} \frac{1}{h} [D_u u(\gamma_t(t), L_t \gamma_t) \cdot N_{t,h,\gamma_t} + \xi(h)]$$

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

$$= \lim_{h \to 0} \frac{1}{h} \int_0^t \nabla_u u(\gamma_t(t), L_t \gamma_t(s)) (N_{t,h,\gamma_t}(s)) \, ds + \lim_{h \to 0} \|N_{t,h,\gamma_t}\| \frac{\xi(h)}{\|N_{t,h,\gamma_t}\|}$$

$$= \langle \nabla_u u(\gamma_t(t), L_t \gamma_t), (L_t \gamma_t)' \rangle$$

by the dominated convergence theorem. If now $u$ depends explicitly on $t$ just write

$$\frac{1}{h} [u_{t+h}(\gamma_t, h) - u_t(\gamma)] = \frac{1}{h} [u(t, t+h, \gamma(t), L_t \gamma) - u(t, \gamma(t), L_t \gamma)]$$

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

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

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

Thanks to this result we can reinterpret the infinite dimensional Kolmogorov equation (25) in terms of the horizontal and vertical derivatives introduced in the previous section. Consider the Kolmogorov equation with horizontal and vertical derivatives, namely

$$\begin{cases}
\partial_t \nu_t(\gamma_t) + b_t(\gamma_t) \cdot \nabla \nu_t(\gamma_t) + \frac{1}{2} \sum_{j=1}^d \sigma_j^2 \partial_j^2 \nu_t(\gamma_t) = 0 , \\
\nu_t(\gamma_T) = f(\gamma_T).
\end{cases}$$

(48)

**Theorem 6.2.** Let $X^{\gamma_t}$ be the solution to equation

$$\begin{cases}
\frac{dX(t)}{X_{t_0}} = b_t(X_t) \, dt + \sigma \, dW(t) \\
X_{t_0} = \gamma_{t_0}
\end{cases}$$

(5)

Associate to $b_t$ and $f$ the operators $B$ and $\Phi$ as in (15); if such $B$ and $\Phi$ satisfy the assumptions of theorem 5.4 then, for almost every $t$, the function

$$\nu_t(\gamma_t) = E[f( X^{\gamma_t}(T)]$$

is a solution of the path dependent Kolmogorov equation (48) for all $\gamma \in C^1_b$ such that $\gamma(0) = 0$.

**Proof.** Lift equation (5) to the infinite dimensional SDE (13) defining the operators $A, B$ and $\Sigma$ as in the previous sections; associate then to this last equation the PDE (25) with final condition given by

$$\Phi ((\frac{\sigma}{\sigma})) = f \left( \tilde{M} \left( \frac{\sigma}{\sigma} \right) \right) \ .$$

Fix $t$: with our choice of $\gamma$ the element $y = (\gamma(t), L_t \gamma_t)$ is in $\tilde{C}$ therefore, if $B$ and $\Phi$ satisfy assumptions 2.1 and 5.1 theorem 5.4 guarantees that $u(s, y) = E[\Phi(Y_x(y)(T))]$ 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 $\gamma_t$, 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 $\nu$ through $u$ by means of (11) we have that

$$
\nu_t(\gamma_t) = u(t, \gamma(t), L_t \gamma_t)
= E\left[f\left(\hat{M}(Y^t,y(T))\right)\right]
= E\left[f\left(X_{\gamma_t}(T)\right)\right].
$$

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 6.1 obtaining that equations (25) and (48) coincide.

**Remark 6.3.** If in the above proof one can show that the function $u$ which solves (25) is in fact differentiable with respect to $t$ for every $t \in [0, T]$, then theorem 6.2 hold everywhere, i.e. the function $\nu$ defined by (49) solves equation (48) for every $t \in [0, T]$.

**Appendix: proof of theorem 2.3**

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

$$
\|Y^{t_0,y+k}(t) - Y^{t_0,y}\|_E =
= \left\|e^{(t-t_0)A}(k) + \int_{t_0}^t e^{(t-s)A}\left[B(s, Y^{t_0,y+k}(s)) - B(s, Y^{t_0,y}(s))\right]ds\right\|_E
\leq C\|k\|_E + C\|DB\|_\infty \int_{t_0}^t \|Y^{t_0,y+k}(s) - Y^{t_0,y}(s)\|_E ds.
$$

ehence, by Gronwall’s lemma,

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

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

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

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

$$
\left\|\int_{t_0}^t e^{(t-s)A} DB(s, Y^{t_0,y}(s)) [\xi_1(s) - \xi_2(s)] ds\right\|_{L(E, E)} \leq C\|DB\|_\infty \int_{t_0}^t \|\xi_1(s) - \xi_2(s)\|_{L(E, E)} ds;
$$
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
\[
\begin{align*}
\xi^{t_0,y,k}(t) &:= Y^{t_0,y+k}(t) - Y^{t_0,y}(t) - \xi^{t_0,y}(t)k \\
&= \int_t^{t_0} e^{(t-s)A} \left[ B(s, Y^{t_0,y+k}(s)) - B(s, Y^{t_0,y}(s)) \right] ds - \\
&\quad \int_t^{t_0} e^{(t-s)A} DB(s, Y^{t_0,y}(s)) \xi^{t_0,y}(s)k \, ds \\
&= \int_t^{t_0} e^{(t-s)A} \left[ \int_0^1 DB(s, \alpha Y^{t_0,y+k}(s) + (1 - \alpha) Y^{t_0,y}(s)) \right. \\
&\quad \left. \int_0^1 DB(s, \alpha Y^{t_0,y+k}(s) + (1 - \alpha) Y^{t_0,y}(s)) \right] ds \\
&\quad \left. \int_0^1 DB(s, \alpha Y^{t_0,y+k}(s) + (1 - \alpha) Y^{t_0,y}(s)) \right] ds \\
&= \int_t^{t_0} e^{(t-s)A} DB(s, Y^{t_0,y}(s)) \xi^{t_0,y}(s)k \, ds \\
&\quad + \int_t^{t_0} e^{(t-s)A} DB(s, Y^{t_0,y}(s)) \xi^{t_0,y}(s)k \, ds \\
&\quad + \int_t^{t_0} e^{(t-s)A} DB(s, Y^{t_0,y}(s)) \xi^{t_0,y}(s)k \, ds.
\end{align*}
\]

Recalling (4.1), we get
\[
\begin{align*}
\|\xi^{t_0,y,k}(t)\|_E &\leq C\|DB\| \int_t^{t_0} \|r^{t_0,y,k}(s)\|_E ds + \\
&\quad + C\|k\|_E \int_t^{t_0} \left\| \int_0^1 DB(s, \alpha Y^{t_0,y+k}(s) + (1 - \alpha) Y^{t_0,y}(s)) \right. \\
&\quad \left. \int_0^1 DB(s, \alpha Y^{t_0,y+k}(s) + (1 - \alpha) Y^{t_0,y}(s)) \right\|_{L(E,E)} ds \\
\end{align*}
\]
which yields, by Gronwall’s lemma,
\[
\|\xi^{t_0,y,k}(t)\| \leq C\|k\|^2.
\]

Therefore
\[
\xi^{t_0,y}(t)k = DY^{t_0,y}(t)k \quad \forall k \in E.
\]

We proceed with an estimate about the continuity of \( \xi^{t_0,y}(t) \) with respect to the initial condition \( y \). For \( h, k \in E \)
\[
\begin{align*}
\|\xi^{t_0,y+k}(t)h - \xi^{t_0,y}(t)h\|_E &\leq \left\| \int_t^{t_0} e^{(t-s)A} DB(s, Y^{t_0,y+k}(s)) \xi^{t_0,y+k}(s)h - DB(s, Y^{t_0,y}(s)) \xi^{t_0,y}h \right\|_E ds \\
&\leq \left\| \int_t^{t_0} e^{(t-s)A} DB(s, Y^{t_0,y+k}(s)) \xi^{t_0,y+k}(s)h - DB(s, Y^{t_0,y+k}(s)) \xi^{t_0,y}h \right\|_E ds \\
&\quad + \left\| \int_t^{t_0} e^{(t-s)A} DB(s, Y^{t_0,y+k}(s)) \xi^{t_0,y+k}(s)h - DB(s, Y^{t_0,y}(s)) \xi^{t_0,y}h \right\|_E ds \\
&\leq C\|DB\| \int_t^{t_0} \|\xi^{t_0,y+k}(s)h - \xi^{t_0,y}(s)h\|_{L(E,E)} ds + \\
&\quad + C \int_t^{t_0} \|DB(s, Y^{t_0,y+k}(s)) - DB(s, Y^{t_0,y}(s))\|_{L(E,E)} \|\xi^{t_0,y}(s)h\|_E ds
\end{align*}
\]
Again by Gronwall’s lemma we get the equation
\[ h \]
for
\[ \text{Second derivative} \]
Therefore
\[ \| \xi_{t_0}^{t_0} h - \xi_{t_0}^{t_0} h \|_E \leq C \| h \|_E \| k \|_E. \] (A2)

Thereby \( \xi_{t_0}^{t_0} h(t) \) is uniformly continuous in \( y \) uniformly in \( t \).

**Second derivative** Let us consider the operator \( U \) from \( C([t_0, T] ; L(E, E)) \) in itself defined through the equation
\[
U(Y)(t)(h, k) = \int_{t_0}^t e^{(t-s)A} D^2 B(s, Y_{t_0}^{t_0}(s)) (\xi_{t_0}^{t_0}(s) h, \xi_{t_0}^{t_0}(s) k) \, ds + \int_{t_0}^t e^{(t-s)A} D B(s, Y_{t_0}^{t_0}(s)) Y(s)(h, k) \, ds \quad (A3)
\]
for \( h, k \in E \) (we identify \( L(E, L(E, E)) \) with \( L(E, E) \) in the usual way). Since
\[
\sup_{t, h, k} \| U(Y_1)(t)(h, k) - U(Y_2)(t)(h, k) \|_E \leq C \| DB \|_{\infty} T \sup_{t, h, k} \| Y_1(t)(h, k) - Y_2(t)(h, k) \|_E
\]
there exists a unique fixed point for \( U \), which will be denoted by \( \eta_{t_0}^{t_0}(t)(h, k) \); furthermore simple calculations yield that \( \| \eta_{t_0}^{t_0}(t) \|_{L(E, E)} \leq C \| \eta \) uniformly in \( t \). We now compute:
\[
\eta_{t_0}^{t_0}(t)(h, k) = \int_{t_0}^t e^{(t-s)A} D B(s, Y_{t_0}^{t_0}(s)) (\xi_{t_0}^{t_0}(s) h, \xi_{t_0}^{t_0}(s) k) \, ds + \int_{t_0}^t e^{(t-s)A} D B(s, Y_{t_0}^{t_0}(s)) Y(s)(h, k) \, ds
\]
\[
\int_0^t \int_0^t e^{-\alpha s - \beta t} \left( (1 - o) Y_s(s) + (1 - o) Y_t(s) - Y_s(Y_t(s) - Y_s(s)) - \xi(s) \right) d\xi(s) ds
\]
These calculations imply that
\[
\|\xi(t)\|_E \leq C_A \|DB\|_E \int_{t_0}^t \|\tilde{\xi}(s)\|_E \, ds
+ C_A \int_{t_0}^t \left\| D^2 B \left( s, \alpha Y^{t_0:y+k}(s) + (1 - \alpha) Y^{t_0:y} \right) \right\|_{L(E,E)} \, ds
+ C_A \|D^2 B\|_E \int_{t_0}^t \left\| \xi(t_0,y+k)(s) h \right\|_E \cdot \left\| Y^{t_0:y+k}(s) - Y^{t_0:y}(s) \right\|_E \, ds
+ C_A \|D^2 B\|_E \int_{t_0}^t \left\| \xi(t_0,y+k)(s) h - \xi(t_0,y)(s) h \right\|_E \cdot \left\| \xi(t_0,y)(s) k \right\|_E \, ds
\leq C_A \|DB\|_E \int_{t_0}^t \|\tilde{\xi}(s)\|_E \, ds
+ C_A \|DB\|_E \int_{t_0}^t \|\tilde{\xi}(s)\|_E \, ds
+ C_A \|DB\|_E \int_{t_0}^t \|\tilde{\xi}(s)\|_E \, ds
+ C_A \|DB\|_E \int_{t_0}^t \|\tilde{\xi}(s)\|_E \, ds
\]
Finally by a application of Gronwall’s lemma

\[
\frac{\|\xi(t)\|_E}{\|k\|_E} \leq C_A \|h\|_E \cdot \left[ \int_{t_0}^t \left\| D^2 B \left( s, \alpha Y^{t_0:y+k}(s) + (1 - \alpha) Y^{t_0:y} \right) \right\|_{L(E,E)} \, ds \right.
+ \int_{t_0}^t \left\| \xi(t_0,y+k)(s) h - \xi(t_0,y)(s) h \right\|_{L(E,E)} \, ds + \|k\|_E \right]
\]
and such quantity goes to 0 uniformly in \(\|h\|_E \leq M \forall M > 0\) 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
\[
\eta^{t_0,y}(t)(h,k) - \eta^{t_0,y}(t)(h,k) =
\int_{t_0}^t e^{(t-s)A} \left[ D^2 B \left( s, Y^{t_0:y}(s) \right) \xi^{t_0,y}(s) h, \xi^{t_0,y}(s) k \right] \, ds
+ \int_{t_0}^t e^{(t-s)A} \left[ DB \left( s, Y^{t_0:y}(s) \right) \eta^{t_0,y}(s)(h,k) - DB \left( s, Y^{t_0:y}(s) \right) \eta^{t_0,y}(s)(h,k) \right] \, ds
= I_1 + I_2;
\]
(A4)
then
\[ I_1 = \int_t^0 e^{(t-s)A} \left[ D^2 B(s, Y^0, (s)) (\xi^{0,y}(s)h, \xi^{0,y}(s)k) - D^2 B(s, Y^{1,0}(s)) (\xi^{0,y}(s)h, \xi^{0,y}(s)k) \right] ds \]
\[ + D^2 B(s, Y^{1,0}(s)) (\xi^{0,y}(s)h, \xi^{0,y}(s)k) - D^2 B(s, Y^{1,0}(s)) (\xi^{1,0}(s)h, \xi^{1,0}(s)k) \right] ds \]
\[ = \int_t^0 e^{(t-s)A} D^2 B(s, Y^{1,0}(s)) (\xi^{1,0}(s)h, \xi^{1,0}(s)k) ds \]
\[ + \int_t^0 e^{(t-s)A} \left[ DB(s, Y^{1,0}(s)) - DB(s, Y^{1,0}(s)) \right] (\xi^{1,0}(s)h, (\xi^{0,y}(s) - \xi^{1,0}(s)) k) ds \]
and
\[ I_2 = \int_t^0 e^{(t-s)A} DB(s, Y^{1,0}(s)) \left[ \eta^{0,y}(s)(h, k) - \eta^{1,0}(s)(h, k) \right] ds \]
\[ + \int_t^0 e^{(t-s)A} \left[ DB(s, Y^{1,0}(s)) - DB(s, Y^{1,0}(s)) \right] \eta^{1,0}(s)(h, k) ds. \]

Recalling all the previous estimates and the fact that both \( \| Y^{1,0}(t) \|_E \) and \( \| \xi^{1,0}(t) \|_{L(E,E)} \) are bounded uniformly in \( t \), denoting with \( C_H \) the Hölder constant of \( D^2 B \), we get

\[ \| \eta^{0,y}(t)(h, k) - \eta^{1,0}(t)(h, k) \|_E \]
\[ \leq C \cdot C_H \int_t^0 \| Y^{1,0}(s) - Y^{1,0}(s) \|_E^\alpha \| \xi^{0,y}(s)h \|_E \| \xi^{0,y}(s)k \|_E \] \[ + C \| D^2 B \|_\infty \int_t^0 \left( \left\| \xi^{1,0}(s) - \xi^{0,y}(s) \right\|_{L(E,E)}^\alpha \right) \| \xi^{1,0}(s) - \xi^{0,y}(s) \|_{L(E,E)}^{1-\alpha} \| \eta^{0,y}(s)(h, k) \|_E \] \[ + C \| DB \|_\infty \int_t^0 \left( \left\| \eta^{1,0}(s) - \eta^{0,y}(s) \right\|_E \right) \| \eta^{1,0}(s) - \eta^{0,y}(s) \|_{L(E,E)}^{1-\alpha} \| \eta^{0,y}(s)(h, k) \|_E \] \[ + C \| DB \|_\infty \int_t^0 \left( \left\| Y^{1,0}(s) - Y^{1,0}(s) \right\|_E^\alpha \right) \| Y^{1,0}(s) - Y^{1,0}(s) \|_{L(E,E)}^{1-\alpha} \| \eta^{0,y}(s)(h, k) \|_E \] \[ \leq C_1 \| h \|_E \| k \|_E \| y - w \|_E^\alpha + C_2 \int_t^0 \left( \left\| \eta^{1,0}(s)(h, k) - \eta^{0,y}(s)(h, k) \right\|_E \right) \] \[ \leq C \| h \|_E \| k \|_E \| y - w \|_E^2 \]

which shows that the second Fréchet derivative of the map \( y \mapsto Y^{1,0}(t) \) is \( \alpha \)-Hölder.

References

[BDPDM92] Alain Bensoussan, Giuseppe Da Prato, Michel C. Delfour, and Sanjoy K. Mitter. Representation and Control of Infinite Dimensional Systems, volume 1 of Systems & Control: Foundations & Applications. Birkhäuser, Boston, 1992.

[CF10a] Rama Cont and David-Antoine Fournié. A functional extension of the Ito formula. Comptes Rendus Mathématiques de l’Académie des Sciences, 348(1-2):57–61, 2010.
[CF10b] Rama Cont and David-Antoine Fournié. Change of variable formulas for non-anticipative functional on path space. Journal of Functional Analysis, 259:1043–1072, 2010.

[CF13] Rama Cont and David-Antoine Fournié. Functional Ito calculus and stochastic integral representation of martingales. Annals of Probability, 41(1):109–133, 2013.

[DGFRar] Cristina Di Girolami, Giorgio Fabbri, and Francesco Russo. The covariation for Banach space valued processes and applications. Metrika, to appear.

[DGFRara] Cristina Di Girolami and Francesco Russo. Generalized covariation and extended Fukushima decompositions for Banach space valued processes. Application to windows of Dirichlet processes. Infinite Dimensional Analysis, Quantum Probability and Related Topics, to appear.

[DGFRarb] Cristina Di Girolami and Francesco Russo. Generalized covariation for Banach space valued processes and Itô formula. Osaka Journal of Mathematics, to appear.

[DPZ92] Giuseppe Da Prato and Jerzy Zabczyk. Stochastic Equations in Infinite Dimension. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 1992.

[Dup09] Bruno Dupire. Functional Itô calculus. Portfolio Research Paper 2009-04, 2009.

[EKTZar] Ibrahim Ekren, Christian Keller, Nizar Touzi, and Jianfeng Zhang. On Viscosity Solutions of Path Dependent PDEs. Annals of Probability, to appear.

[ETZ13a] Ibrahim Ekrem, Nizar Touzi, and Jianfeng Zhang. Viscosity Solutions of Fully Nonlinear Parabolic Path Dependent PDEs: Part I. 2013, arXiv:1210.0007v2[Math.PR].

[ETZ13b] Ibrahim Ekrem, Nizar Touzi, and Jianfeng Zhang. Viscosity Solutions of Fully Nonlinear Parabolic Path Dependent PDEs: Part II. 2013, arXiv:1210.0006v2[Math.PR].

[FGG10] Salvatore Federico, Ben Goldys, and Fausto Gozzi. HJB Equations for the Optimal Control of Differential Equations with Delays and State Constraints I: Regularity of Viscosity Solutions. SIAM Journal on Control and Optimization, 48:4910–4937, 2010.

[FMT10] Marco Fuhrman, Federica Masiero, and Gianmario Tessitore. Stochastic equations with delay: optimal control via BSDEs and regular solutions of Hamilton-Jacobi-Bellman equations. SIAM Journal of Control and Optimization, 48(7):4624–4651, 2010.

[GM06] Fausto Gozzi and Carlo Marinelli. Stochastic optimal control of delay equations arising in advertising models. Da Prato, Giuseppe (ed.) et al., Stochastic partial differential equations and applications-VII, 245:133–148, 2006.

[PW11] Shige Peng and Falei Wang. BSDE, Path-dependent PDE and Nonlinear Feynman-Kac Formula. 2011, arXiv:1108.4317[Math.PR].

[RDG10] Francesco Russo and Cristina Di Girolami. Infinite dimensional stochastic calculus via regularization. HAL INRIA 00473947, 2010, http://hal.inria.fr/inria-00473947/PDF/InfDimRVApril10.pdf.

[TZ13] Shanjian Tang and Fu Zhang. Path-Dependent Optimal Stochastic Control and Viscosity Solution of Associated Bellman Equations. 2013, arXiv:1210.2078v3[Math.PR].