The existence and uniqueness of the solution for nonlinear Kolmogorov equations

Jianjun Zhou\textsuperscript{a,b}, Bin Liu\textsuperscript{b,}\textsuperscript{*}

\textsuperscript{a} College of Science, Northwest Agriculture and Forestry University, Yangling 712100, Shaanxi, PR China
\textsuperscript{b} School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan 430074, Hubei, PR China

A R T I C L E   I N F O

Article history:
Received 12 April 2010
Revised 19 March 2012
Available online 31 August 2012

MSC:
93E20
60H30
60H15

Keywords:
Kolmogorov equations
Stochastic delay evolution equations
Backward stochastic differential equations
Malliavin calculus
Optimal control

A B S T R A C T

By means of backward stochastic differential equations, the existence and uniqueness of the mild solution are obtained for the nonlinear Kolmogorov equations associated with stochastic delay evolution equations. Applications to optimal control are also given.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

In this paper, we consider stochastic delay evolution equations of the form

\[
\begin{align*}
\frac{dX(s)}{ds} &= AX(s)ds + F(s, X_s)ds + G(s, X_s)dW(s), \quad s \in [t, T], \\
X_t &= x,
\end{align*}
\] (1.1)

Notes:
\textsuperscript{\ddagger} This work was partially supported by NNSF of China (Grant No. 11171122).
\textsuperscript{*} Corresponding author. Fax: +86 27 87543231.
\textsuperscript{E-mail address:} binliu@mail.hust.edu.cn (B. Liu).

0022-0396/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jde.2012.08.026
where

\[ X_s(\theta) = X(s + \theta), \quad \theta \in [-\tau, 0], \text{ and } x \in C([-\tau, 0], H). \]

The Wiener process \( W \) takes values in a Hilbert space \( \mathcal{E} \). \( A \) is the generator of a \( C_0 \) semigroup of bounded linear operator in another Hilbert space \( H \) and the coefficients \( F \) and \( G \) are assumed to satisfy Lipschitz conditions with respect to appropriate norms. Under suitable assumptions, there exists a unique adapted process \( X(s, t, x), s \in [t - \tau, T], \) solution to (1.1).

Our approach to optimal control problems for stochastic delay evolution equations is based on backward stochastic differential equations (BSDEs) which were first introduced by Pardoux and Peng [1]: see [2,3] as general references. In fact, we consider the following forward–backward system

\[
\begin{cases}
    dX(s) = AX(s) \, ds + F(s, X_s) \, ds + G(s, X_s) \, dW(s), & s \in [t, T], \\
    X_t = x, \\
    dY(s) = -\psi(s, X_s, Y(s), Z(s)) \, ds + Z(s) \, dW(s), & s \in [t, T], \\
    Y(T) = \phi(X_T).
\end{cases}
\]

(1.2)

For (1.2), if \( \psi \) and \( \phi \) satisfy suitable conditions, there exists a unique continuous adapted solution, denoted by \( (X(\cdot, t, x), Y(\cdot, t, x), Z(\cdot, t, x)) \) in \( H \times \mathcal{E} \times \mathcal{E} \). We define a deterministic function \( v : [0, T] \times C \rightarrow R \) by \( v(t, x) = Y(t, t, x) \), it turns out that \( v \) is unique mild solution of the generalization nonlinear Kolmogorov equation:

\[
\begin{cases}
    \frac{\partial v(t, x)}{\partial t} + \mathcal{L}_t[v(t, \cdot)](x) = \psi(t, x, v(t, x), \nabla_0 v(t, x) G(t, x)), & t \in [0, T], x \in C, x(0) \in D(A), \\
    v(T, x) = \phi(x),
\end{cases}
\]

(1.3)

where

\[
\mathcal{L}_t[\phi](x) = \mathcal{S}(\phi)(x) + \left\langle Ax(0)_{10} + F(t, x)_{10}, \nabla_x \phi(x) \right\rangle \\
+ \frac{1}{2} \sum_{i=1}^d \nabla^2_x \phi(x) \left( G(t, x)e_i_{10}, G(t, x)e_i_{10} \right).
\]

(1.4)

\( \{e_i\}_{1 \leq i \leq d} \) denotes a basis of \( \mathcal{E} \). Setting \( \psi \equiv 0 \) and \( d = \infty \), if \( v \) is sufficiently regular, we show that \( v \) is a classical solution of the Kolmogorov equation

\[
\begin{cases}
    \frac{\partial v(t, x)}{\partial t} + \mathcal{L}_t[v(t, \cdot)](x) = 0, & t \in [0, T], x \in C, x(0) \in D(A), \\
    v(T, x) = \phi(x).
\end{cases}
\]

(1.5)

Here \( \nabla_x \phi(x) \), \( \nabla^2_x \phi(x) \) denote the extensions of \( \nabla_x \phi(x) \), \( \nabla^2_x \phi(x) \), respectively (see Lemma 2.1 and Lemma 2.2). \( \nabla_0 v(t, x) G(t, x) \) is defined by \( \nabla_0 v(t, x) G(t, x) = \nabla_0 v(t, x)(G(t, x)_{10}) \).

Consider a controlled equation of the form:

\[
\begin{cases}
    dX^u(s) = AX^u(s) \, ds + F(s, X^u_s) \, ds + G(s, X^u_s) R(s, X^u_s, u(s)) \, ds + G(s, X^u_s) \, dW(s), & s \in [t, T], \\
    X^u_t = x.
\end{cases}
\]

(1.6)
The control process $u$ takes values on a measurable space $(U, \mathcal{U})$. The control problem consists of minimizing a cost functional of the form:

$$J(u) = E \int_{t}^{T} q(s, X^u_s, u(s)) \, ds + E\phi(X^u_T),$$

(1.7)

over all admissible controls. Here $q$ and $\phi$ are functions on $[t, T] \times C \times U$ and $C$, respectively. We define the Hamiltonian function relative to the above problem: for all $t \in [0, T]$, $x \in C$, $z \in \mathcal{S}$,

$$\psi(t, x, z) = \inf\{q(t, x, u) + zR(t, x, u) : u \in U\}.$$  

(1.8)

Then, under suitable conditions, we eventually show that $v$ is the value function of the control problem.

Stochastic optimal control problem has been studied by several authors. In the papers [4–8], the authors showed there exists a direct (classical or mild) solution of the corresponding Hamilton–Jacobi–Bellman (HJB) equation, and the optimal feedback law is obtained by the solution. Using regularity properties of the transition semigroup of the associated Ornstein–Uhlenbeck process, F. Gozzi [7, 8] proved there exists unique mild solution of the associated HJB equation, under the diffusion term only satisfies weaker nondegeneracy conditions.

The notion of viscosity solution for HJB equations has been successfully applied to stochastic optimal control problems (see [9–14] and references therein). In [10], P.L. Lions showed that the value function of the stochastic control for Zakai’s equation coincides with the unique viscosity solution of the associated dynamic programming equation. A. Świąȩch [12] proved the existence and uniqueness of the viscosity solution for general unbounded second order partial differential equation. In [15,16], optimal control problems for a special class of nonlinear stochastic differential equations with delay are considered. Mou-Hsiung Chang et al. [17,18] treat an optimal control problem for general stochastic differential equations with a bounded memory and show that the value function is the unique viscosity solution of the HJB equation.

BSDEs are useful tools in the study of control problems, for example, [19–21]. The existence of optimal control for stochastic systems in infinite dimensions has been considered in [22–25]. Using Malliavin calculus and BSDEs, M. Fuhrman and G. Tessitore [22] show that there exists unique mild solution of nonlinear Kolmogorov equations and it coincides with the value function of the control problem. In [24], the existence and uniqueness of the mild solution for the generalizations of the Kolmogorov equations is proved and the existence of optimal control is obtained by the feedback law.

Some authors considered the Kolmogorov equations associated with stochastic evolution equations (see [22] and the following articles) and with stochastic delay differential equations (see [24]). However, as far as we know, there are few authors who concentrated on Eqs. (1.3) and (1.4), for example [26] for $G$ as a constant. In this paper we want to extend the results of [22] and [24] to stochastic delay evolution equations in Hilbert spaces. The main difficulties are that we have to find some extensions $\tilde{f}$ and $\tilde{\beta}$ of the functionals $f : C \to H$ and $\beta : C \times C \to H$ such that, for every $z \in H$ and $t \in (-\tau, 0]$, $\tilde{f}(z(1_{[-\tau,t]}))$ and $\tilde{\beta}(z1_{[-\tau,t]}, z1_{[-\tau,t]})$ are well defined and we should establish a formula for the Malliavin derivative of a $H$-valued functional of a stochastic process with values in $H$. Thanks to Lemma 2.1, Lemma 2.2 and Lemma 2.3, we can consider the optimal control problem of (1.1) and the associated nonlinear Kolmogorov equations (1.3) and (1.4).

The paper is organized as follows. In the next section we introduce the basic notations and prove Lemma 2.3, which is the key of many subsequent result. Section 3 is devoted to proving the regularity in the Malliavin spaces of the solution for stochastic delay evolution equation. The forward–backward system is considered and the formula (4.4) is proved in Section 4. In Section 5, the Kolmogorov equations (1.3) and (1.5) are considered. We showed that (1.5) has a classical solution and $v(t, x)$ is unique mild solution of (1.3). Finally, applications to optimal control are presented in Section 6.
2. Preliminaries

We list some notations that are used in this paper. Let $\mathcal{S}$, $K$ and $H$ denote real separable Hilbert spaces, with scalar products $(\cdot, \cdot)_\mathcal{S}$, $(\cdot, \cdot)_K$ and $(\cdot, \cdot)_H$, respectively. We assume $\mathcal{S}$ is a finite dimensional Hilbert space expected in Section 5.1. We use the symbol $|\cdot|$ to denote the norm in various spaces, with a subscript if necessary. $C = C([-\tau, 0], H)$ denotes the space of continuous functions from $[-\tau, 0]$ to $H$, endowed with the usual norm $|f| = \sup_{\theta \in [-\tau, 0]}|f(\theta)|_H$. $L(\mathcal{S}, H)$ denotes the space of all bounded linear operators from $\mathcal{S}$ into $H$; the subspace of Hilbert–Schmidt operators, with the Hilbert–Schmidt norm, is denoted by $L_2(\mathcal{S}, H)$.

Let $(\Omega, \mathcal{F}, P)$ be a complete space with a filtration $\mathcal{F}_t$ which satisfies the usual condition, i.e., $\mathcal{F}_t$ is a linear isometry into.

We say $f \in C^1(\mathcal{C}; H)$ if $f$ is continuous, Fréchet derivatives and $\forall x : \mathcal{C} \rightarrow L(\mathcal{C}, H)$ is continuous. We say $g \in \mathcal{G}([0, T] \times \mathcal{C}; H)$ if $g$ is continuous, Gâteaux derivatives with respect to $x$ on $[0, T] \times \mathcal{C}$ and $\forall x : \mathcal{C} \rightarrow L(\mathcal{C}, H)$ is strongly continuous.

Let $F_c$ be the vector space of all simple functions of $\mathbb{1}_{[a,c]}$, where $z \in H$, $c \in (a, b]$ and $1_{[a,c]} : [a, b] \rightarrow R$ is the character function of $[a, c]$. It is clearly that $C([a, b], H) \cap F_c = \emptyset$. Form the direct sum $C([a, b], H) \oplus F_c$ and give it the complete norm

$$
\|y + z1_{[a,b]}\| = \sup_{s \in [a,b]} |y(s)| + |z|, \quad y \in C([a, b], H), \quad z \in H.
$$

We say $f : C([a, b], H) \oplus F_c \rightarrow K$ satisfying (V) if $|\alpha|^1_{\alpha \geq 1}$ is a bounded sequence in $C([a, b], H)$ and $y + z1_{[a,c]} \in C([a, b], H) \oplus F_c$ such that $\alpha^0(s) \rightarrow \alpha(s) + z1_{[a,c]}$ as $n \rightarrow \infty$ for all $s \in [a, b]$, and $\sup_{s \in [a,b]} |\alpha^0(s) - \alpha(s), h_i| \lesssim |(z, h_i)|$ for all $i \in N$, where $\{h_i\}_{i \geq 1}$ is a basis of $H$, then $f(\alpha^0) \rightarrow f(y + z1_{[a,c]})$ as $n \rightarrow \infty$.

**Lemma 2.1.** Let $f \in L(C([a, b]; H); K)$. Then, for every $c \in (a, b]$, $f$ has a unique continuous linear extension $\tilde{f} : C([a, b], H) \oplus F_c \rightarrow K$ satisfying (V). Moreover, the extension map $e : L(C([0, T], H), K) \rightarrow L(C([0, T], H) \oplus F_c, K), f \rightarrow \tilde{f}$ is a linear isometry into.

**Proof.** Let $\{h_i\}_{i \geq 1}$ be a basis of $H$. We define the bounded linear functional $f^i$ from $C([a, b], R)$ to $K$ by

$$
f^i(x) = f(h_i, x), \quad x \in C([a, b], R),
$$

and, for every $y \in K$, define $f^y$ from $C([a, b]; H)$ to $R$ by

$$
f^y(x) = f(x)y, \quad x \in C([a, b]; H).
$$

Then, $f^y$ can be extended to $B([a, b], H)$ and get $F^y$, moreover, $\|F^y\| = \|F^y\|$, where $B([a, b], H)$ denotes the space of bounded functions in $[a, b]$ with the norm $\|x\| = \sup_{0 \leq t \leq b} |x(t)|, x \in B([a, b], H)$. Let us define $k^i : \xi \mapsto f^y(h_i1_{[a,\xi]})$ for $\xi \in [a, b]$.

Firstly, for all $z \in H$, $y \in K$ and $\|y\| = \|z\| = 1$, let $\delta_{i, j, y} = \text{sign}((z, h_i)(k^i, j, y) - k^i, j, y(\xi_{j - 1}))$ and $a = \xi_0 < \xi_1 < \cdots < \xi_n = b$, we have that, for every $N \in \mathbb{N}$,

$$
\sum_{j=1}^{n} \sum_{i=N+1}^{\infty} (\xi_i, y)(k^i, j, y) - k^i, j, y(\xi_{j - 1})) = \sum_{j=1}^{n} \sum_{i=N+1}^{\infty} (\delta_{i, j, y}(z, h_i)(k^i, j, y) - k^i, j, y(\xi_{j - 1}))
$$
\[
\sum_{j=1}^{\infty} \sum_{n=N+1}^{\infty} F^j(\epsilon_{i,j,y}(z,h_i)(1_{[a,\xi_j]} - 1_{[a,\xi_{j-1}]})h_i)
\]

\[
= \sum_{j=1}^{\infty} F^j \left( \sum_{n=N+1}^{\infty} \epsilon_{i,j,y}(z,h_i)(1_{[a,\xi_j]} - 1_{[a,\xi_{j-1}]})h_i \right)
\]

\[
= F^j \left( \sum_{j=1}^{n} (1_{[a,\xi_j]} - 1_{[a,\xi_{j-1}]}) \sum_{n=N+1}^{\infty} \epsilon_{i,j,y}(z,h_i)h_i \right)
\]

\[
\leq \| F^j \| \left( \sum_{n=N+1}^{\infty} (z,h_i)^2 \right)^{\frac{1}{2}} \leq \| f \| \left( \sum_{n=N+1}^{\infty} (z,h_i)^2 \right)^{\frac{1}{2}}.
\]

and, for a fixed \( M > 0 \), for any \( \varepsilon > 0 \), there exists an \( n > 0 \) such that

\[
\sum_{i=N+1}^{M} \left| (z,h_i) \right|^{\frac{1}{a}} \left( k^{i,y}(\xi) \right) \leq \varepsilon + \sum_{i=N+1}^{M} \left| (z,h_i) \right| \sum_{j=1}^{n} \left| k^{i,y}(\xi_j) - k^{i,y}(\xi_{j-1}) \right|
\]

\[
\leq \varepsilon + \| f \| \left( \sum_{n=N+1}^{\infty} (z,h_i)^2 \right)^{\frac{1}{2}}.
\]

Therefore, we obtain that

\[
\sum_{i=N+1}^{M} \left| (z,h_i) \right|^{\frac{1}{a}} \left( k^{i,y}(\xi) \right) \leq \| f \| \left( \sum_{n=N+1}^{\infty} (z,h_i)^2 \right)^{\frac{1}{2}}.
\]

Letting \( M \to \infty \) we show that

\[
\left( \sum_{i=N+1}^{\infty} (z,h_i)^2 \right)^{\frac{1}{2}} \leq \| f \| \left( \sum_{n=N+1}^{\infty} (z,h_i)^2 \right)^{\frac{1}{2}}.
\]

By the Riesz representation theorem [27, Theorem VI (7.2)–(7.3), pp. 492–496], for every \( f^i \), there is a unique regular finite measure \( \mu^i : \text{Borel}[a,b] \to K \) such that

\[
f^i(x) = \int_a^b x(s) \, d\mu^i(s) \quad \text{for all } x \in C([a,b], R).
\]

For every \( y \in K \), we have that \( f^i(x)y = \int_a^b x(s) \, d\mu^i(s)y \), therefore, \( |\mu^i(\cdot)y| \leq \| f^i \|^2 \). Define \( \tilde{f} : C([a,b], H) \oplus F_c \to K \) by

\[
\tilde{f}(x + z1_{[a,c]}) = f(x) + \sum_{i=1}^{\infty} (z,h_i)\mu^i([a,c]) \quad x \in C([a,b], H), \quad z \in H, \quad c \in (a,b).
\]

Let \( \{x^n\}_{n \geq 1} \) be a bounded sequence in \( C([a,b], H) \) such that \( x^n(s) \to x(s) + z1_{[a,c]}(s) \) as \( n \to \infty \) for all \( s \in [a,b] \) and \( \sup_{x \in [a,b]} |(x^n(s) - x(s), h_i)| \leq \| (z, h_i) \| \) for all \( i \in N \). Then we have that
\[
\left| f(x^n) - \bar{f}(x + z_{1[a,c]}) \right| \\
\leq \sum_{i=1}^{N} \left| f^i((x^n - x, h_i)) - (z, h_i)\mu^i([a, c]) \right| + \left| f\left( \sum_{i=N+1}^{\infty} (x^n - x, h_i)h_i \right) \right| + \left| \sum_{i=N+1}^{\infty} (z, h_i)\mu^i([a, c]) \right| \\
= \sum_{i=1}^{N} \left| f^i((x^n - x, h_i)) - (z, h_i)\mu^i([a, c]) \right| + \left| f\left( \sum_{i=N+1}^{\infty} (x^n - x, h_i)h_i \right) \right| \\
+ \sup_{y \in K, \|y\|=1} \sum_{i=N+1}^{\infty} (z, h_i)\mu^i([a, c])y \\
\leq \sum_{i=1}^{N} \int_{a}^{b} (x^i(s) - x(s) - z_{1[a,c]}(s), h_i) d\mu^i(s) + \|f\| \left( \sum_{i=N+1}^{\infty} (z, h_i)^2 \right)^{\frac{1}{2}} \\
+ \sup_{y \in K, \|y\|=1} \sum_{i=N+1}^{\infty} \left| (z, h_i) \right| \sqrt{\left( k^j(y) \right)} \\
\leq \sum_{i=1}^{N} \int_{a}^{b} (x^i(s) - x(s) - z_{1[a,c]}(s), h_i) d\mu^i(s) + 2\|f\| \left( \sum_{i=N+1}^{\infty} (z, h_i)^2 \right)^{\frac{1}{2}}.
\]

For any \( \varepsilon > 0 \), firstly, letting \( N \) large enough such that \( 2\|f\|\left( \sum_{i=N+1}^{\infty} (z, h_i)^2 \right)^{\frac{1}{2}} < \frac{\varepsilon}{2} \), then for the fixed \( N \), by the dominated convergence theorem for vector-valued measures [27,28], we have that \( \sum_{i=1}^{N} \|f^i((x^n - x, h_i)) - (z, h_i)\mu^i([a, c])\| \leq \frac{\varepsilon}{2} \) for \( n \) large enough. Therefore, we get that

\[
\lim_{n \to \infty} f(x^n) = \bar{f}(x + z_{1[a,c]}).
\]

To prove uniqueness let \( f' \in L(C([a, b], H) \oplus F_c, K) \) be a continuous linear extension of \( f \) satisfying (V). For any \( z_{1[a,c]} \in F_c \) choose a bounded sequence \( \{x^n\}_{n \geq 0} \) in \( C([a, b], H) \) such that \( x^n(s) \to z_{1[a,c]} \) as \( n \to \infty \) for all \( s \in [a, b] \); e.g. take

\[
x^n(s) = \begin{cases} 
  z, & s \in [a, c - \frac{1}{n}], \\
  nz(c - s), & s \in (c - \frac{1}{n}, c], \\
  0, & s \in (c, b].
\end{cases}
\]

By (V) one has that

\[
f'(x + z_{1[a,c]}) = f'(x) + f'(z_{1[a,c]}) = f(x) + \lim_{n \to \infty} f(x^n) = f(x) + \bar{f}(z_{1[a,c]}) = \bar{f}(x + z_{1[a,c]}).
\]

for all \( x \in C([a, b], H) \). Thus \( f' = \bar{f} \).

The map \( \varepsilon \) from \( L(C([0, T], H), K) \to L(C([0, T], H) \oplus F_c, K) \) to \( f \to \bar{f} \) is clearly linear. Since \( \bar{f} \) is an extension of \( f \), we have that \( \|f\| \leq \|\bar{f}\| \). On the other hand, let \( x = y + z_{1[a,c]} \in C([0, T], H) \oplus F_c \) and construct \( \{x^n\}_{n \geq 1} \) in \( C([a, b], H) \) as above, so we have that

\[
\|\bar{f}(x)\| = \lim_{n \to \infty} \|f(y + x^n)\| \leq \lim_{n \to \infty} \|f\|\|y + x^n\| \\
\leq \lim_{n \to \infty} \|f\|\|y\| + \|x^n\| = \|f\|\|y\| + \|z\| = \|f\|\|x\|
\]

for all \( x \in C([0, T], H) \oplus F_c \). Thus \( \|\bar{f}\| \leq \|f\| \). Therefore, \( \|\bar{f}\| = \|f\| \) and \( \varepsilon \) is an isometry into. \( \square \)
Lemma 2.2. Let \( \beta : C([a, b], H) \times C([a, b], H) \to R \) be a continuous bilinear map. Then, for every \( c \in (a, b] \), \( \beta \) has a unique continuous bilinear extension \( \tilde{\beta} : [C([a, b], H) \oplus F_c] \times [C([a, b], H) \oplus F_c] \to R \) satisfying (W).

**Proof.** We define the bounded bilinear functional \( \beta^{i,j} \) from \( C([a, b], R) \times C([a, b], R) \) to \( R \) by

\[
\beta^{i,j}(x, y) = \beta(xh_i, yh_j), \quad x, y \in C([a, b]; R),
\]

and define \( \beta^{i} \) from \( C([a, b], R) \times C([a, b], H) \) to \( R \) by

\[
\beta^{i}(x, y) = \beta(xh_i, y), \quad x \in C([a, b]; R), \quad y \in C([a, b]; H).
\]

By Theorem VI (7.2)–(7.3) in [28], we have that there is a unique measure \( u^{i,j} : \text{Borel}[a, b] \to C([a, b], R)^{+} \) such that for all \( \xi, y \in C([a, b], R) \)

\[
\beta^{i,j}(\xi, y) = \int_a^b \xi(s) du^{i,j}(s)y,
\]

and by the Riesz representation theorem there is a unique regular finite measure \( \mu^{i,y} : \text{Borel}[a, b] \to R \) such that

\[
\beta^{i}(x, y) = \int_a^b x(s) d\mu^{i,y}(s), \quad x \in C([a, b]; R), \quad y \in C([a, b]; H).
\]

Hence, for all \( x, y \in C([a, b], H), \)

\[
\beta(x, y) = \sum_{i, j=1}^{\infty} \int_a^b (x, h_i)(s) du^{i,j}(s)(y, h_j) = \sum_{i=1}^{\infty} \int_a^b (x(s), h_i) d\mu^{i,y}(s).
\]

If we let \( x, y \) have the form: \( x = (x, h_i)h_i \), \( y = (y, h_j)h_j \), then we get that

\[
\mu^{i,j}(\cdot)(y, h_j) = \mu^{i, (y, h_j)h_j}(\cdot), \quad i, j \in \mathbb{N}.
\]

Define \( \tilde{\beta} : [C([a, b], H) \oplus F_c] \times C([a, b], H) \to R \) by

\[
\tilde{\beta}(x + z1_{[a,c]}), y) = \beta(x, y) + \sum_{i=1}^{\infty} (z, h_i)\mu^{i,y}([a, c]), \quad x, y \in C([a, b], H), \quad z \in H, \quad c \in (a, b].
\]

Firstly, let us show that \( \tilde{\beta}(z1_{[a,c]}, y) \) is bounded linear functional with respect to \( y \) for every \( z \in H \), and \( c \in (a, b] \). Let \( \{x_n\} \) be a sequence satisfying \( \sup_{s \in [a,b]} |(x_n(s), h_i)| \leq |(z, h_i)| \) for all \( i \geq 1 \) such that \( \lim_{n \to \infty} x_n = z1_{[a,c]} \). By Lemma 2.1, we have that
\[
\tilde{\beta}(z_{1[a,c]}, e_1 y_1 + e_2 y_2) = \lim_{n \to \infty} \beta(x_n, e_1 y_1 + e_2 y_2) \\
= \lim_{n \to \infty} e_1 \beta(x_n, y_1) + \lim_{n \to \infty} e_2 \beta(x_n, y_2) \\
= e_1 \tilde{\beta}(z_{1[a,c]}, y_1) + e_2 \tilde{\beta}(z_{1[a,c]}, y_2)
\]

and

\[
\|\tilde{\beta}(z_{1[a,c]}, \cdot)\| = \sup_{y \in C([a,b], H), \|y\|=1} |\tilde{\beta}(z_{1[a,c]}, y)| \leq \sup_{y \in C([a,b], H), \|y\|=1} \|\tilde{\beta}(\cdot, y)\| \|z\| \\
= \sup_{y \in C([a,b], H), \|y\|=1} \sup_{x \in C([a,b], H), \|x\|=1} |\beta(x, y)| \|z\| \\
= \|\beta\| \|z\|.
\]

Let us define \(\beta : C([a, b], H) \oplus F_c \to [C([a, b], H) \oplus F_c]^*\) by

\[
\beta(x + z_{1[a,c]}, y) = \beta(x + z_{1[a,c]}), \quad x \in C([a, b], H), z \in H, c \in (a, b).
\]

Let \(\{x^n\}_{n \geq 1}, \{y^n\}_{n \geq 1}\) be bounded sequences in \(C([a, b], H)\) such that \(x^n(s) \to x(s) + z_{1[a,c]}(s)\) and \(y^n(s) \to y(s) + z_{2[a,c]}(s)\) as \(n \to \infty\) for all \(s \in [a, b]\) and \(\sup_{s \in [a, b]} |(x^n(s) - x(s), h_i)| \leq |(z_1, h_i)|, \sup_{s \in [a, b]} |(y^n(s) - y(s), h_i)| \leq |(z_2, h_i)|\) for all \(i \in N\). Then we have that, for any \(\varepsilon > 0\),

\[
|\beta(x^n, y^n) - \beta(x + z_{1[a,c]}, y + z_{2[a,c]})| \\
\leq |\beta(x^n, y^n) - \tilde{\beta}(x + z_{1[a,c]}, y^n)| + |\tilde{\beta}(x + z_{1[a,c]}, y^n) - \tilde{\beta}(x + z_{1[a,c]}, y + z_{2[a,c]})| \\
\leq \sum_{i=1}^N \sum_{j=1}^M \int_a^b (x^n(s) - x(s) - z_{1[a,c]}(s), h_i) d\mu^{i,j}(y^n, h_j)(s) \\
+ \sum_{i=1}^N \sum_{j=M+1}^\infty (y^n, h_j)(s) \\
+ \sum_{i=1}^N \sum_{j=1}^{M+1} (z_1, h_i)_{1[a,c]} h_i, \sum_{j=M+1}^\infty (y^n, h_j) h_j \\
+ \tilde{\beta}(x + z_{1[a,c]}, y^n) - \tilde{\beta}(x + z_{1[a,c]}, y + z_{2[a,c]}) \\
\leq \sum_{i=1}^N \sum_{j=1}^M \int_a^b (x^n(s) - x(s) - z_{1[a,c]}(s), h_i) d\mu^{i,j}(s) (\|z_2\| + \|y\| c) \\
+ \sum_{i=1}^N \|\beta\| (z_1, h_i) \left[ \left( \sum_{j=M+1}^\infty (z_2, h_j)^2 \right)^{1/2} + \sup_{s \in [a, b]} \left( \sum_{j=M+1}^\infty (y(s), h_j)^2 \right)^{1/2} \right] \\
+ 2\|\beta\| \left( \sum_{i=N+1}^\infty (z_1, h_i)^2 \right)^{1/2} (\|z_2\| + \|y\| c) + |\tilde{\beta}(x + z_{1[a,c]}, y^n) - \tilde{\beta}(x + z_{1[a,c]}, y + z_{2[a,c]})|,
\]
Firstly, letting $N$ be large enough such that $2\|\beta\|\left(\sum_{i=N+1}^{\infty} (z_1, h_i)^2\right)^{\frac{1}{2}} (\|z_2\| + \|y\|_C) < \frac{\varepsilon}{3}$, then for the fixed $N$, letting $M$ be large enough such that $2 \sum_{i=1}^{N} \|\beta\| (z_1, h_i) \left(\sum_{j=M+1}^{\infty} (z_2, h_j)^2\right)^{\frac{1}{2}} + \sup_{s \in [a, b]} (\sum_{j=M+1}^{\infty} (y(s), h_j)^2) < \frac{\varepsilon}{3}$; now for the fixed $M$, $N$, by the dominated convergence theorem and Lemma 2.1 we have that for all $x$, letting $\beta(x) = \sum_{j=M+1}^{\infty} (z_2, h_j)^2 \|z_2\| + \|y\|_C$ for all $x \in \mathbb{R}$, we denote the partial derivatives with respect to the $l$th variable and by $f_1, f_2, \ldots, f_m$ are infinitely differentiable functions $\mathbb{R}^n \rightarrow \mathbb{R}$ bounded together with all their derivatives. The Malliavin derivative $D F$ of $F \in S_K$ is defined as the process $D_s F$, $s \in [0, T]$, where

$$D_s F = \sum_{i=1}^{m} \sum_{l=1}^{n} \partial_l f_i (W(h_1), W(h_2), \ldots, W(h_n)) k_i \otimes h_l(s).$$

By $\partial_l$ we denote the partial derivatives with respect to the $l$th variable and by $k_i \otimes h_l(s)$ denote the operator $u \mapsto k_i(h_l(s), u)_{\mathcal{E}}$. It is well known that the operator $D : S_K \subset L^2(\mathcal{E}; K) \rightarrow L^2(\mathcal{E} \times [0, T]; L_2(\mathcal{E}; K))$ is closable. We denote by $\mathbf{D}^{1,2}(K)$ the domain of its closure, and use the same letter to denote $D$ and its closure:

$$D : \mathbf{D}^{1,2}(K) \subset L^2(\mathcal{E}; K) \rightarrow L^2(\mathcal{E} \times [0, T]; L_2(\mathcal{E}; K)).$$

The adjoint operator of $D$:

$$\delta : \text{dom}(\delta) \subset L^2(\Omega \times [0, T]; L_2(\mathcal{E}; K)) \rightarrow L^2(\Omega, K)$$

is called Skorohod integral.
For \( v \in C([-\tau, 0]; L_2(\mathcal{S}, H)) \) we define \( \nabla_x f(y)v \) by

\[
\nabla_x f(y)v = \sum_{i=1}^{d} (\nabla_f(y)(v, e_i)) \otimes e_i,
\]

and define \( \overline{\nabla_x f(y)}v \) by

\[
\overline{\nabla_x f(y)}v = \sum_{i=1}^{d} (\overline{\nabla_f(y)(v, e_i)}) \otimes e_i.
\]

Here \( \{e_i\}_{1 \leq i \leq d} \) is a basis of \( \mathcal{S} \). Now we prove a lemma which plays a key role in our paper.

**Lemma 2.3.** Let \( f \in C^1([-\tau, T]; K; H) \) and \( \nabla_x f \) be bounded. We assume that \( y \in L_2^2([-\tau, T]; H) \cap L^2([-\tau, T]; D^{1,2}(H)) \) is a continuous process, \( \{D_t y(s), s \in [t, T]\} \) is a continuous process for all \( t \in [0, T] \) satisfying

\[
E \sup_{s \in [-\tau, T]} |y(s)|^2 < \infty; \quad E \int_0^T \sup_{s \in [t, T]} |D_t y(s)|^2 \, dt < \infty.
\]

Then \( f(y) \in D^{1,2}(K) \) and its Malliavin derivative is given by the formula: for a.e. \( t \in [0, T] \) we have that

\[
D_t (f(y)) = \overline{\nabla_x f(y)} D_t (y).
\]

**Proof.** We apply the techniques introduced in Lemma 2.6 of [29]. Firstly, we assume that \( f \) is of the form:

\[
f(y) = g(y(s_1), y(s_2), \ldots, y(s_n)), \tag{2.1}
\]

where \( g \in C(H^n; K), \nabla_x g \) is bounded and \( s_i \in [-\tau, T], i = 1, 2, \ldots, n \), so we have that \( f(y) \in D^{1,2}(K) \) and

\[
D_t f(y) = D_t g(x(s_1), x(s_2), \ldots, x(s_n)) = \sum_{i=1}^{n} \nabla_{x_i} g(x(s_1), x(s_2), \ldots, x(s_n)) D_t (y(s_i)).
\]

On the other hand, by the definition of the Fréchet derivative, for every \( z \in C([-\tau, T]; H) \), we get that

\[
\nabla_x f(y) z = \sum_{i=1}^{n} \nabla_{x_i} g(x(s_1), x(s_2), \ldots, x(s_n)) z(s_i).
\]

From Lemma 2.1 it follows that the above relation is true for all \( z \in C([-\tau, T]; H) \oplus F_t \), where \( t \in (-\tau, T] \) and \( \nabla_x f(y)z \) is replaced by \( \overline{\nabla_x f(y)}z \). Then we get that

\[
D_t f(y) = \overline{\nabla_x f(y)} D_t (y) = \sum_{i=1}^{n} \nabla_{x_i} g(x(s_1), x(s_2), \ldots, x(s_n)) D_t (y(s_i)). \tag{2.2}
\]

In the general case, we define \( P_n : x \in C \rightarrow P_n x \in C \) by
\[ P_n x(s) = x\left(\frac{k(T + \tau)}{n} - \tau\right) + \frac{n}{T + \tau} \left(s - \frac{k(T + \tau)}{n} + \tau\right) \left(x\left(\frac{(k + 1)(T + \tau)}{n} - \tau\right) - x\left(\frac{k(T + \tau)}{n} - \tau\right)\right), \quad s \in \left[\frac{k(T + \tau)}{n} - \tau, \frac{(k + 1)(T + \tau)}{n} - \tau\right) \]

and define \( f^n(x) : x \in \mathcal{C} \to \mathcal{C} \) by

\[ f^n(x) = f(P_n x), \quad x \in \mathcal{C}. \]

It is clearly that \( f^n(x) \) is of the form described by (2.1), so we get that \( f^n \in \mathbf{D}^{1,2}(K) \) and

\[ D_{s} f^n(y) = \nabla_x f^n(y) D_s(y). \tag{2.3} \]

By the definition of Fréchet derivative we can verify that, for \( z \in \mathcal{C} \),

\[ \nabla_x f^n(x) z = \nabla_x f(P_n x) P_n z. \]

We assume that \( z \in C([-\tau, T], H) \oplus F_t \), let us construct \( \{z^m\}_{m \geq 0} \) in \( C([-\tau, T], H) \) as in Lemma 2.1, by the definition of \( P_n \),

\[ \nabla_x f^n(x) z = \lim_{m \to \infty} \nabla_x f^n(x) z^m = \lim_{m \to \infty} \nabla_x f(P_n x) P_n z^m = \nabla_x f(P_n x) P_n z. \tag{2.4} \]

Now we have that, for a suitable constant \( C > 0 \),

\[ \lim_{n \to \infty} E \left\| f^n(y) - f(y) \right\|^2 = \lim_{n \to \infty} E \left\| f(P_n y) - f(y) \right\|^2 \leq \lim_{n \to \infty} CE \sup_{s \in [-\tau, T]} \| P_n y(s) - y(s) \|^2 = 0, \]

and

\[ \lim_{n \to \infty} E \int_{0}^{T} \left| D_{t} f^n(y) - \nabla_x f(y) D_{t}(y) \right|^2_{L_{2}(\mathcal{Z}, H)} dt \]

\[ = \lim_{n \to \infty} E \int_{0}^{T} \left| \nabla_x f(P_n y) P_n D_{t}(y) - \nabla_x f(y) D_{t}(y) \right|^2_{L_{2}(\mathcal{Z}, H)} dt \]

\[ \leq \lim_{n \to \infty} 2E \int_{0}^{T} \left| \left( \nabla_x f(P_n y) - \nabla_x f(y) \right) P_n D_{t}(y) \right|^2_{L_{2}(\mathcal{Z}, H)} dt \]

\[ + \lim_{n \to \infty} 2E \int_{0}^{T} \left| \nabla_x f(y) (P_n D_{t}(y) - D_{t}(y)) \right|^2_{L_{2}(\mathcal{Z}, H)} dt \]

\[ \leq \lim_{n \to \infty} \left[ 2dE \int_{0}^{T} \left| \nabla_x f(P_n y) - \nabla_x f(y) \right|^2 \sup_{s \in [0, T]} \left| D_{t}(y(s)) \right|^2_{L_{2}(\mathcal{Z}, H)} dt + 4dE \int_{0}^{T} \left| \nabla_x f(y) \right|^2 \right. \]

\[ \times \sup_{s \in [0, T]} \left| P_n \left[ D_{t}(y(t)) 1_{(-\tau, t]}(s) + D_{t}(y(s)) \right] - D_{t}(y(t)) 1_{(-\tau, t]}(s) - D_{t}(y(s)) \right|^2_{L_{2}(\mathcal{Z}, H)} dt \]
\[
+ 4E \int_0^T \left| \nabla_x f(y) \left( P_n D_t(y(t)) - D_t(y(t)) 1_{(-\tau,t)}(s) \right)^2_{L^2(\Xi, H)} \right| dt
\]

\[= 0. \]

It follows from the closedness of the operator \( D \) that \( f(y) \in D^{1,2}(K) \) and for \( a.e. \ t \in [0, T] \),

\[D_t(f(y)) = \nabla_x f(y) D_t(y).\]

The proof is finished. \( \square \)

We note that it follows from the proof procedure of Lemma 2.1 that, for \( K = R \),

\[
\left( \sum_{i=1}^{\infty} (z, h_i) k^i(\xi) \right) \leq \sum_{i=1}^{\infty} |(z, h_i)| \left( \sum_{i=1}^{\infty} (k^i(\xi)) \right) \leq \|f\|. \tag{2.5}
\]

Then we have that

\[
f(z \mid_{[a,b]}) = \sum_{i=1}^{\infty} (z, h_i) \int_a^b 1_{[c, b]} d\mu^i(s) = \int_a^b 1_{[c, b]} d\left( \sum_{i=1}^{\infty} (z, h_i) \mu^i(s) \right)
\]

\[= \int_a^b z 1_{[c, b]} d\bar{f}(s), \quad z \in \mathbb{H}, \ c \in [a, b]. \tag{2.6}\]

We also notice that, similarly to Lemma 2.5 in [24], Lemma 2.3 also holds true for \( f \in G^1(\mathbb{C}; H) \).

### 3. The forward equation

In this section we consider the system of stochastic delay evolution equation:

\[
\begin{cases}
  dX(s) = AX(s) ds + F(s, X_s) ds + G(s, X_s) dW(s), & s \in [t, T], \\
  X_t = x \in \mathbb{C}
\end{cases} \tag{3.1}
\]

for \( s \) varying on the time interval \([t, T]\). We make the following assumptions.

**Hypothesis 3.1.** (i) The operator \( A \) is the generator of a strongly continuous semigroup \( \{ e^{tA}, t \geq 0 \} \) of bounded linear operators in the Hilbert space \( H \).

(ii) The mapping \( F: [0, T] \times \mathbb{C} \to H \) is measurable and satisfies, for some constant \( L > 0 \),

\[
|F(t, x)| \leq L(1 + |x|_C),
\]

\[|F(t, x) - F(t, y)| \leq L|x - y|_C, \quad t \in [0, T], \ x, y \in \mathbb{C}. \]

(iii) The mapping \( G: [0, T] \times \mathbb{C} \to L^2(\Xi, H) \) is measurable and satisfies, for every \( t \in [0, T] \) and \( x, y \in \mathbb{C} \),

\[
|G(t, x)|_{L^2(\Xi, H)} \leq L(1 + |x|_C),
\]

\[|G(t, x) - G(t, y)|_{L^2(\Xi, H)} \leq L(|x - y|_C), \tag{3.2}\]

for some constants \( L > 0 \).
We say that \(X\) is a mild solution of Eq. (3.1) if it is a continuous, \(\{\mathcal{F}_t\}_{t \geq 0}\)-predictable process with values in \(H\), and it satisfies: \(P\)-a.s.,

\[
\begin{aligned}
X(s) &= e^{(s-t)A}X(0) + \int_t^s e^{(s-\sigma)A}F(\sigma, X_\sigma) \, d\sigma + \int_t^s e^{(s-\sigma)A}G(\sigma, X_\sigma) \, dW(\sigma), \\
& \quad s \in [t, T], \\
X_t &= x \in \mathcal{C}.
\end{aligned}
\]

For every \(p \in [2, \infty)\), let \(L^p_p(\Omega, \mathcal{C}|[-\tau, T]; H)\) denote the space of predictable process \(X\) with continuous paths in \(H\), such that the norm \(|X|^p = E\sup_{t \in [-\tau, T]}|X(t)|^p\) is finite.

We extend the domain of the solution setting \(X(s, t, x) = x((s - t) \vee (-\tau))\) for \(s \in [-\tau, t)\).

We say \(u \in \overline{L}^{1,p}(H)\) if \(u \in L^p_p(\Omega, \mathcal{C}|[-\tau, T]; H)\) and \(\{D_t u(s), s \in [t, T]\}\) is a continuous process for all \(t \in [0, T]\) and satisfies that

\[
E \int_0^T \sup_{s \in [t, T]} |D_t u(s)|^p \, dt < \infty.
\]

**Lemma 3.3.** If \(X \in \overline{L}^{1,p}(H)\) for \(p > 2\), then the random processes

\[
\begin{aligned}
\int_0^l e^{(l-r)A}F(r, X_r) \, dr, & \quad l \in [0, T], \\
\int_0^l e^{(l-r)A}G(r, X_r) \, dW(r), & \quad l \in [0, T],
\end{aligned}
\]

belong to \(\overline{L}^{1,p}(H)\) and for all \(s\) and \(l\) with \(s < l\), we have that

\[
\begin{aligned}
D_s \int_0^l e^{(l-r)A}F(r, X_r) \, dr &= \int_0^s e^{(l-r)A}\nabla_X F(r, X_r) D_s X_r \, dr, \\
D_s \int_0^l e^{(l-r)A}G(r, X_r) \, dW(r) &= e^{(l-s)A}G(s, X_s) + \int_s^l e^{(l-r)A}\nabla_X G(r, X_r) D_s X_r \, dW(r).
\end{aligned}
\]
Proof. We apply the techniques introduced in Lemma 3.6 of [22]. We only prove (3.5). Let \( u(r) = e^{(l-r)A}G(r, X_r) \) \( (u(r) = 0 \) for \( r > l \) or \( r < 0 \)). We note that

\[
E \int_0^T |u(r)|^P \, dr \leq T E \sup_{r \in [-\tau, T]} |u(r)|^P = T E \sup_{r \in [0, l]} |e^{(l-r)A}G(r, X_r)|^P_{L_2(\mathcal{E}, H)}
\]

\[
\leq T (ML)^P E \sup_{r \in [0, l]} (1 + |X_r|)^P \leq T (ML)^P E \sup_{r \in [-\tau, T]} (1 + |X(r)|)^P < \infty.
\]

From Lemma 2.3 it follows that \( D_s u(r) = e^{(l-r)A}V_x G(r, X_r) D_s X_r \) for all \( s \leq r \), whereas \( D_s u(r) = 0 \) for \( s > r \). By Hypothesis 3.1(iv), we have that

\[
E \int_0^T \int_0^T |D_s u(r)|^P \, dr \, ds = E \int_0^T \int_s^l |e^{(l-r)A}V_x G(r, X_r) D_s X_r|^P_{L_2(\mathcal{E}, L_2(\mathcal{E}, H))} \, dr \, ds
\]

\[
\leq (2LM)^P d^2 E \int_0^T \int_s^l \sup_{\theta \in [-r, 0]} |D_s X_r(\theta)|^P_{L_2(\mathcal{E}, H)} \, dr \, ds
\]

\[
\leq (2LM)^P d^2 T E \int_0^T \sup_{r \in [s, T]} |D_s X(r)|^P_{L_2(\mathcal{E}, H)} \, ds < \infty.
\]

Now from the fact that the Skorohod and the Itô integral coincide for adapted integrands it follows that

\[
\int_0^T E \left| \delta(D_s u) \right|^2 \, ds = \int_0^T E \left| D_s u(r) \, dW(r) \right|^2 \, ds = E \int_0^T \int_0^T |D_s u(r)|^2 \, dr \, ds < \infty.
\]

Since

\[
\delta(u) = \int_0^l e^{(l-r)A}G(r, X_r) \, dW(r), \quad \delta(D_s u) = \int_s^l e^{(l-r)A}V_x G(r, X_r) D_s X_r \, dW(r),
\]

we get that (3.5) holds for all \( l \in [0, T] \) by Proposition 3.4 in [30].
\begin{align*}
\leq c_p T (ML)^p E \sup_{r \in [0,T]} (|1 + |X_r|^p) + c_p (2LM)^p (dT)^2 E \int_0^T \sup_{r \in [s,T]} |D_s X(r)|^p_{L_2(\mathcal{E}, H)} \, ds
< \infty.
\end{align*}

It shows that the process \( \int_0^l e^{(l-r)A} G(r, X_r) \, dW(r) \) belongs to \( \mathcal{L}^{1,p}(H) \). \( \square \)

**Theorem 3.4.** Let Hypothesis 3.1 be satisfied, then \( X(l) \in D^{1,2}(H) \) and

\begin{align*}
D_s X(l) &= e^{(s-l)A} G(s, X_s) + \int_s^l e^{(l-\sigma)A} \nabla_x F(\sigma, X_\sigma) D_\sigma X_\sigma \, d\sigma \\
&\quad + \int_s^l e^{(l-\sigma)A} \nabla_x G(\sigma, X_\sigma) D_\sigma X_\sigma \, dW(\sigma), \quad l \in [s, T].
\end{align*}

(3.6)

**Proof.** Let us consider the Picard approximation of stochastic delay evolution equation (3.3): \( X^0(s) \equiv x(0), \, s \in [t, T] \).

\[ X^{n+1}(s) = e^{(s-t)A} x(0) + \int_t^s e^{(s-\sigma)A} F(\sigma, X^n_\sigma) \, d\sigma + \int_t^s e^{(s-\sigma)A} G(\sigma, X^n_\sigma) \, dW(\sigma), \quad s \in [t, T], \]

and \( X^n(s) = x((s-t) \vee (\tau - t)) \) for \( s \in [-\tau, t) \). By the proof of Theorem 3.2, we get that \( X^n \) converges to the solution of Eq. (3.3) in the space \( L^p_\mathcal{P}(\Omega; C([-\tau, T]; H)) \). From Lemma 3.3 it follows that \( X^n \in \mathcal{L}^{1,p}(H) \) and

\begin{align*}
D_s X^{n+1}(l) &= e^{(l-s)A} G(s, X^n_s) + \int_s^l e^{(l-\sigma)A} \nabla_x F(\sigma, X^n_\sigma) D_\sigma X^n_\sigma \, d\sigma \\
&\quad + \int_s^l \nabla_x \left(e^{(l-\sigma)A} G(\sigma, X^n_\sigma)\right) D_\sigma X^n_\sigma \, dW(\sigma), \quad l \in [s, T].
\end{align*}

Similarly to [22], setting \( I(X^n)_{sl} = e^{(l-s)A} G(s, X^n_s) \) for \( l \geq s \) and \( I(X^n)_{sl} = 0 \) for \( l < s \); \( \Gamma_1(X^n, DX^n)_{sl} = \int_s^l e^{(l-s)A} \nabla_x F(\sigma, X^n_\sigma) D_\sigma X^n_\sigma \, d\sigma \) and \( \Gamma_2(X^n, DX^n)_{sl} = \int_s^l e^{(l-s)A} \nabla_x G(\sigma, X^n_\sigma) D_\sigma X^n_\sigma \, dW(\sigma) \). We define the space \( \mathcal{H}^p \) of processes \( Q_{sl}, 0 \leq s \leq l \leq T \), such that for every \( s \in [t, T] \), \( Q_{sl}, l \in [s, T] \), is a predictable process in \( L_2(\mathcal{E}, H) \) with continuous paths, and such that

\[
\sup_{s \in [0,T]} E \left( \sup_{l \in [s,T]} e^{-\beta p(l-s)} |Q_{sl}|^p_{L_2(\mathcal{E}, H)} \right) < \infty.
\]

We first note that \( I(X^n) \) is a bounded sequence in \( \mathcal{H}^p \). Since

\[
\sup_{s \in [0,T]} E \left( \sup_{l \in [s,T]} e^{-\beta p(l-s)} |e^{(l-s)A} G(s, X^n_s)|^p_{L_2(\mathcal{E}, H)} \right) \leq M^p L^p \sup_{s \in [0,T]} E \left( 1 + |X^n_s|^p \right)
\]

\( X^n \) is a bounded sequence in \( L^p(\Omega, C([-\tau, T]; H)) \). Next we show that
\[
\sup_{s \in [0,T]} \mathbb{E} \left( \sup_{l \in [s,T]} e^{-\beta p(l-s)} \left( \| I_1(X^n, DX^n)_s \|_{L^2(\mathbb{S}, H)}^p + \| I_2(X^n, DX^n)_s \|_{L^2(\mathbb{S}, H)}^p \right) \right) \\
\leq C(\beta) \sup_{s \in [0,T]} \mathbb{E} \left( \sup_{l \in [s,T]} e^{-\beta p(l-s)} \| D_s X^n(l) \|_{L^2(\mathbb{S}, H)}^p \right),
\]

where \( C(\beta) \) depends on \( \beta, p, L, T \) and \( M =: \sup_{l \in [0,T]} |e^{|A|} \), and \( C(\beta) \to 0 \) as \( \beta \to \infty \). For simplicity, we only consider the operator \( I_2 \). Fixed \( s \in [0, T) \) we introduce the space of \( L^2(\mathbb{S}, H) \)-valued continuous adapted processes \( Q_l \), \( l \in [s, T] \) such that the norm \( \| Q \|_p^p = E \sup_{l \in [s,T]} e^{-\beta p(l-s)} \| Q(l) \|_{L^2(\mathbb{S}, H)}^p \) is finite. We will use the so-called factorization method; see [31, Theorem 5.2.5]. Let us take \( p \) and \( \alpha \in (0, 1) \) such that

\[
\frac{1}{p} < \alpha < \frac{1}{2}
\]

and let \( c^{-1}_\alpha = \int_0^s (s-r)^{\alpha-1} (r-\sigma)^{-\alpha} \, dr \), by the stochastic Fubini theorem

\[
I_2(X^n, DX^n)_s = c_\alpha \int_0^r \int_s^l (l-r)^{\alpha-1} (r-\sigma)^{-\alpha} e^{(l-r)A} e^{(r-\sigma)A} \nabla_x G(\sigma, X^n_\sigma) D_s X^n_\sigma \, dr \, dW(\sigma)
\]

\[
= c_\alpha \int_0^r (l-r)^{\alpha-1} e^{(l-r)A} V(r) \, dr,
\]

where

\[
V(r) = \int_s^r (r-\sigma)^{-\alpha} e^{(r-\sigma)A} \nabla_x G(\sigma, X^n_\sigma) D_s X^n_\sigma \, dW(\sigma).
\]

Setting \( q = \frac{p}{(p-1)} \),

\[
\| I_2(X^n, DX^n)_s \| \leq c_\alpha M \int_0^r (l-r)^{\alpha-1} |V(r)| \, dr
\]

\[
\leq c_\alpha M \left( \int_0^r e^{q \beta (r-s)} (l-r)^{q(\alpha-1)} \, dr \right)^{\frac{1}{q}} \left( \int_0^r e^{-p \beta (r-s)} |V(r)|^p \, dr \right)^{\frac{1}{p}},
\]

and we get that

\[
\| I_2(X^n, DX^n)_s \|_p^p \leq c_\alpha^p M^p \int_s^T e^{-q \beta (r-s)} E |V(r)|^p \, dr \sup_{l \in [s,T]} e^{-\beta p(l-s)} \left( \int_s^l e^{q \beta (r-s)} (l-r)^{q(\alpha-1)} \, dr \right)^{\frac{p}{q}}
\]

\[
\leq c_\alpha^p M^p \int_s^T e^{-q \beta (r-s)} E |V(r)|^p \, dr \left( \int_0^T e^{-q \beta r} r^{q(\alpha-1)} \, dr \right)^{\frac{p}{q}}.
\]
Thus

\[
\| \Gamma_2 (X^n, DX^n) \|_s \leq c_\alpha M \left( \int_0^T e^{-p\beta (r-s)} E|V(r)|^p \, dr \right)^{\frac{1}{p}} \left( \int_0^T e^{-q\beta r^q (\alpha - 1)} \, dr \right)^{\frac{1}{q}}.
\]

By the Burkholder–Davis–Gundy inequalities, we have, for some constant \( c_p \) depending only on \( p \),

\[
E|V(r)|^p \leq c_p E \left( \int_0^r (r - \sigma)^{-2\alpha} e^{(r-\sigma)A} \nabla x G(\sigma, X^n_\sigma) D_2 X_\sigma |_{L^2(\mathcal{E})}^2 \right)^{\frac{p}{2}}
\]

\[
\leq (2MLd)^{\frac{1}{2}} c_p \left( \int_0^r (r - \sigma)^{-2\alpha} \sup_{\theta \in [-\tau, 0]} |D_2 X_\sigma (\theta)|^2 \right)^{\frac{p}{2}}
\]

\[
\leq (2MLd)^{\frac{1}{2}} c_p \| DX^n \|_s \left( \int_0^r (r - \sigma)^{-2\alpha} e^{2\beta (\sigma - s)} \, d\sigma \right)^{\frac{p}{2}}.
\]

Then we obtain that

\[
\| \Gamma_2 (X^n, DX^n) \|_s \leq 2M^2 Lc_\alpha d^2 (Tc_p)^{\frac{1}{2}} \left( \int_0^T e^{-q\beta r^q (\alpha - 1)} \, dr \right)^{\frac{1}{q}} \left( \int_0^T e^{-2\beta \sigma \sigma^2} \, d\sigma \right)^{\frac{1}{2}} \| DX^n \|_s.
\]

This inequality proves (3.7).

Finally, we show that there exists unique \( Q \in \mathcal{H}^p \) such that \( Q \) satisfies (3.6) and \( DX^n \rightarrow Q \) in \( \mathcal{H}^p \) as \( n \rightarrow \infty \). By the closedness of the operator \( D \), one has that \( X(t) \) belongs to \( D^{1,2}(H) \) and there exists a version of \( DX \) such that \( DX = Q \) satisfying (3.6). \( \square \)

Similarly to [24, Corollary 2.7], by a standard truncation procedure, we obtain the following result.

**Corollary 3.5.** Assume Hypothesis 3.1 holds true, \( X \) is the solution of (3.1) and \( f \in \mathcal{G}^1(C; R) \) satisfies

\[
|\nabla_x f (x) h| \leq C |h|_C (1 + |x|_C)^m, \quad h, x \in C,
\]

for some \( C > 0 \) and \( m \geq 0 \). Then for every \( t \in [0, T] \), \( f(X_t) \in D^{1,2} \) and for a.e. \( s \in [0, T] \),

\[
D_s (f(X_t)) = \nabla_x f(X_t) D_s (X_t).
\]

**4. The forward–backward system**

In this section we consider the system of stochastic differential equations

\[
\begin{cases}
    dX(s) = AX(s) \, ds + F(s, X_s) \, ds + G(s, X_s) \, dW(s), & s \in [t, T], \\
    X_t = x, \\
    dY(s) = -\psi (s, X_s, Y(s), Z(s)) \, ds + Z(s) \, dW(s), \\
    Y(T) = \phi(X_T)
\end{cases}
\]

(4.1)
for $s$ varying on the time interval $[t, T] \subset [0, T]$. As in Section 3 we extend the domain of the solution setting $X(s, t, x) = x((s - t) \lor (-\tau))$ for $s \in [-\tau, t)$.

We make the following assumptions.

**Hypothesis 4.1.** (i) The mapping $\psi : [0, T] \times C \times R \times L_2(\mathcal{S}, R) \to R$ is Borel measurable such that, for all $t \in [0, T]$, $\psi(t, \cdot, \cdot, \cdot) : C \times R \times L_2(\mathcal{S}, R) \to R$ is continuous and for some $L > 0$ and $m \geq 1$,

\[
\left|\psi(s, x, y_1, z_1) - \psi(s, x, y_2, z_2)\right| \leq L \left( |y_1 - y_2| + |z_1 - z_2| \right),
\]

\[
\left|\psi(s, x, y, z)\right| \leq L \left( 1 + |x|^m + |y| + |z| \right),
\]

for every $s \in [0, T], x \in C, y, y_1, y_2 \in R, z, z_1, z_2 \in L_2(\mathcal{S}, R)$.
(ii) The map $\psi(t, \cdot, \cdot, \cdot) \in \mathcal{G}^1(\mathcal{C} \times R \times L_2(\mathcal{S}, R); R)$ and there exist $L > 0$ and $m \geq 1$ such that

\[
\left|\nabla_x \psi(t, x, y, z)h\right| \leq L |h| \left( 1 + |x| + |y| \right)^m \left( 1 + |z| \right)
\]

for every $t \in [0, T], x, h \in C, y \in R$ and $z \in L_2(\mathcal{S}, R)$.
(iii) The map $\phi \in \mathcal{G}^1(\mathcal{C}; R)$ and there exists $L > 0$ such that, for every $x_1, x_2 \in C$,

\[
\left|\phi(x_1) - \phi(x_2)\right| \leq L |x_1 - x_2| \mathcal{C}.
\]

Firstly, we prove a technical result, Theorem 4.2, which will be used in the rest of this section.

**Theorem 4.2.** Assume that $u : [0, T] \times C \to R$ is a Borel measurable function such that $u(t, \cdot) \in \mathcal{G}^1(C, R)$ and

\[
|u(t, x)| + |\nabla_x u(t, x)| \leq C \left( 1 + |x| \right)^m,
\]

for some $C > 0, m \geq 0$ and for every $t \in [0, T], x \in C$. Then the joint quadratic variation on $[t, T']$

\[
\left\{u(\cdot, X_s), W e_i\right\}_{[t, T']} = \int_t^{T'} \nabla_0 u(s, X_s) G(s, X_s) e_i ds,
\]

for every $x \in C, i = 1, 2, \ldots, d$ and $0 < T' < T$.

**Proof.** We apply the techniques introduced in Theorem 3.1 of [24]. We only have to prove that

\[
C^\epsilon := C^\epsilon_{[0, T']} \left( u(\cdot, X_s), W e_i \right) = \frac{1}{\epsilon} \int_0^{T'} \left( u(t + \epsilon, X_{t+\epsilon}) - u(t, X_t) \right) \left( W(t + \epsilon) e_i - W(t) e_i \right) dt
\]

\[
\to \int_0^{T'} \nabla_0 u(t, X_t) G(t, X_t) e_i dt,
\]

in probability as $\epsilon \to 0$. 
First we notice that from the rules of Malliavin calculus it follows that, for a.e. \( t \in [0, T'] \),

\[
\left( u(t + \epsilon, X_{t+\epsilon}) - u(t, X_t) \right)(W(t + \epsilon)e_i - W(t)e_i) \\
= \left( u(t + \epsilon, X_{t+\epsilon}) - u(t, X_t) \right) \int_t^{t+\epsilon} e_i \, dW(s) \\
= \int_t^{t+\epsilon} D_s(u(t + \epsilon, X_{t+\epsilon}) - u(t, X_t)) e_i \, ds + \int_t^{t+\epsilon} \left( u(t + \epsilon, X_{t+\epsilon}) - u(t, X_t) \right) e_i \, d\hat{W}(s),
\]

where the symbol \( d\hat{W} \) denotes the Skorohod integral.

Next we recall that \( D_s(u(t, X_t)) = 0 \) for \( s > t \) by the adaptedness, and from (3.8) and (4.3) we deduce that

\[
C^\epsilon = -\frac{1}{\epsilon} \left( 1 - T' \right) \int_0^{t+\epsilon} \sum_{i=1}^n \nabla X_i u(t + \epsilon, X_{t+\epsilon}) D_s (X_{t+\epsilon}) e_i \, ds \, dt \\
+ \frac{1}{\epsilon} \int_0^{T'+\epsilon} \sum_{i=1}^n (u(t + \epsilon, X_{t+\epsilon}) - u(t, X_t)) dte_i \, d\hat{W}(s) \\
= \frac{1}{\epsilon} \int_0^{T'} \sum_{i=1}^n \nabla X_i u(t + \epsilon, X_{t+\epsilon}) e^{(t+\epsilon + \theta - s)} A G(s, X_s) e_i \, ds \, dt \\
+ \frac{1}{\epsilon} \int_0^{T'} \sum_{i=1}^n \nabla X_i u(t + \epsilon, X_{t+\epsilon}) e^{(t+\epsilon + \theta)} \int_s^{t+\epsilon} D_s (X_{t+\epsilon}) dW(\sigma) e_i \, ds \, dt \\
+ \frac{1}{\epsilon} \int_0^{T'} \sum_{i=1}^n \nabla X_i u(t + \epsilon, X_{t+\epsilon}) e^{(t+\epsilon + \theta)} \int_s^{t+\epsilon} D_s F(\sigma, X_s) \, d\sigma e_i \, ds \, dt \\
+ \frac{1}{\epsilon} \int_0^{T'} \sum_{i=1}^n (u(t + \epsilon, X_{t+\epsilon}) - u(t, X_t)) dte_i \, d\hat{W}(s) \\
= I_1^\epsilon + I_2^\epsilon + I_3^\epsilon + I_4^\epsilon.
\]

By the similar procedure in Theorem 3.1 of [24], we can show that \( I_1^\epsilon + I_2^\epsilon + I_3^\epsilon \rightarrow 0 \) in probability as \( \epsilon \rightarrow 0 \). Now let us compute the limit of \( I_1^\epsilon \).

\[
I_1^\epsilon = -\frac{1}{\epsilon} \left( 1 - T' \right) \int_0^{t+\epsilon} \sum_{i=1}^n \nabla X_i u(t + \epsilon, X_{t+\epsilon}) e^{(t+\epsilon + \theta - s)} A G(s, X_s) e_i \, ds \, dt \\
= \frac{1}{\epsilon} \int_{t-\epsilon}^{T'+\epsilon} \sum_{i=1}^n \nabla X_i u(t, X_t) 1_{[s-t,0]}(\theta) e^{(t+\theta - s)} A G(s, X_s) e_i \, ds \, dt
\]
Next let us show that $H^\epsilon_1 \to 0$, P-a.s. as $\epsilon \to 0$.

$$|H^\epsilon_1| \leq C \left( 1 + \sup_{t \in [0,T]} |X_t| \right) \frac{1}{\epsilon} \int_{t-\epsilon}^{t+\epsilon} \int \sup_{s-t \leq \theta \leq 0} \left| e^{(t+\theta-s)A} G(s, X_s) e_t - G(t, X_t) e_t \right| ds \, dt$$

$$\leq C \left( 1 + \sup_{t \in [0,T]} |X_t| \right) \frac{1}{\epsilon} \int_{t-\epsilon}^{t+\epsilon} \int \sup_{s-t \leq \theta \leq 0} \left| e^{(t+\theta-s)A} G(s, X_s) e_t - G(t, X_t) e_t \right| ds \, dt$$

$$\leq CM \left( 1 + \sup_{t \in [0,T]} |X_t| \right) \frac{1}{\epsilon} \int_{t-\epsilon}^{t+\epsilon} \sup_{0 \leq s \leq t} \left| G(s, X_s) e_t - G(t, X_t) e_t \right| ds \, dt$$

$$+ C \left( 1 + \sup_{t \in [0,T]} |X_t| \right) \frac{1}{\epsilon} \int_{t-\epsilon}^{t+\epsilon} \sup_{0 \leq s \leq t} \left| e^{sA} G(t, X_t) e_t - G(t, X_t) e_t \right| ds \, dt$$

$$\to 0 \quad \text{P-a.s. as } \epsilon \to 0.$$

Now it remains to consider $H^\epsilon_2$. By (2.6),

$$H^\epsilon_2 = \frac{1}{\epsilon} \int_{t-\epsilon}^{t+\epsilon} \int \nabla_x u(t, X_t) 1_{[s-t,0]}(\theta) G(t, X_t) e_t ds \, dt$$

$$= \frac{1}{\epsilon} \int_{t-\epsilon}^{t+\epsilon} \int \left( \int_0^{t+\theta} G(t, X_t) e_t d\nabla_x u(t, X_t)(\theta) \right) ds \, dt$$

$$= \frac{1}{\epsilon} \int_{t-\epsilon}^{t+\epsilon} \int (\epsilon + \theta) G(t, X_t) e_t d\nabla_x u(t, X_t)(\theta) ds \, dt$$

$$= \int \int_0^{t+\epsilon} \left( 1 + \frac{\theta}{\epsilon} \right) G(t, X_t) e_t d\nabla_x u(t, X_t)(\theta) ds \, dt.$$
By (2.5) we have that, P-a.s.,

\[
\int_{-\tau}^{0} \left( 1 + \frac{\theta}{\epsilon} \right)^+ G(t, X_t)e_i \frac{d\bar{X}}{\bar{X}}(\theta)
\]

\[
\rightarrow \int_{-\tau}^{0} 1_{[0]}(\theta)G(t, X_t)e_i \frac{d\bar{X}}{\bar{X}}(\theta) = \bar{v}_0 G(t, X_t)e_i,
\]

and by the dominated convergence theorem we obtain that

\[
H_2^c \rightarrow \int_{0}^{T'} \bar{v}_0 G(t, X_t)e_i dt.
\]

This shows that \(C^c\) converges in probability and its limit is

\[
[u(\cdot, X), W e_i]_{[0, T']} = \int_{0}^{T'} \bar{v}_0 G(s, X_s)e_i ds. \quad \square
\]

For \(p \geq 2\) we denote: \(K_p(t) = L_p^p(\Omega; C([t, T]; R)) \times L_p^p(\Omega; L^2([t, T], \Xi))\) endowed with the natural norm. For simplicity, we denote \(K_p(h)\) by \(K_p\).

For the backward–forward system (4.1) we have the following basic result (see Proposition 5.2 in [22]).

**Theorem 4.3.** Assume Hypotheses 3.1 and 4.1 hold, then for every \(p \in [2, \infty)\), there exists a unique solution in \(L_p^m(\Omega, \mathbb{R}) \times C\) of Eqs. (4.1) that will be denoted by \((X(\cdot, t, x), Y(\cdot, t, x), Z(\cdot, t, x))\) and the map \((t, x) \rightarrow (Y(\cdot, t, x), Z(\cdot, t, x))\) belongs to \(G_{0,1}([0, T] \times \mathbb{R}; \mathbb{R})\). Moreover, for every \(h \in C\), there exists a suitable constant \(c > 0\) such that

\[
\left( E \sup_{s \in [0, T]} |\nabla Y(s, t, x) h|^p \right)^{1/p} \leq c |h| (1 + |x|^2). \]

**Corollary 4.4.** Assume Hypotheses 3.1 and 4.1 hold. Then the function \(v(t, x) = Y(t, t, x)\) belongs to \(G_{0,1}([0, T] \times \mathbb{R}; \mathbb{R})\) and there exists a constant \(C > 0\) such that \(|\nabla v(t, x) h| \leq C|h| (1 + |x|^2)\) for all \(t \in [0, T], x \in \mathbb{R}\) and \(h \in C\).

Moreover, for every \(t \in [0, T]\) and \(x \in \mathbb{R}\), we have that

\[
Y(s, t, x) = v(s, X_s(t, x)), \quad P\text{-a.s., for all } s \in [t, T],
\]

\[
Z(s, t, x) = \nabla v(s, X_s(t, x)) G(s, X_s(t, x)), \quad P\text{-a.s., for a.e. } s \in [t, T]. \tag{4.4}
\]

**Proof.** By the similar procedure of Proposition 5.7 in [22], we can show that \(v \in G_{0,1}([0, T] \times \mathbb{R}; \mathbb{R})\), and its property is a direct consequence of Theorem 4.3. From uniqueness of the solution of Eq. (4.1) it follows that \(Y(s, t, x) = v(s, X_s(t, x))\), P-a.s. for all \(s \in [t, T]\).
Now we consider the joint quadratic variation of \( Y(\cdot, t, x) \) and the Wiener process \( W(t) \) on an interval \([t, T']\). By the backward stochastic differential equation we get that

\[
\{ Y(\cdot, t, x), W(t) \}_{[t, T']} = \int_t^{T'} Z(s, t, x)e_i \, ds.
\]

From Theorem 4.2 it follows that

\[
\{ \nu(\cdot, X(t, x)), W(t) \}_{[t, T']} = \int_t^{T'} \nabla_0 \nu(s, X_x(t, x)) G(s, X_x(t, x)) e_i \, ds.
\]

Therefore, we obtain that

\[
Z(s, t, x) = \nabla_0 \nu(s, X_x(t, x)) G(s, X_x(t, x)), \quad \text{P-a.s., for a.e. } s \in [t, T]. \quad \square
\]

5. The Kolmogorov equation

5.1. Classical solution of the Kolmogorov equation

Some authors considered the classical solution of the Kolmogorov equation associated with stochastic evolution equations (see [33]) and with stochastic delay differential equations (see [28,34]). In this subsection, we study the classical solution of the Kolmogorov equation associated with stochastic delay evolution equations.

Let \( X(s, t, x) \) denote the mild solution of the stochastic delay evolution equation

\[
\begin{aligned}
&dX(s) = AX(s) \, ds + F(s, X_x) \, ds + G(s, X_x) \, dW(s), \quad s \in [t, T], \\
&X_t = x \in C
\end{aligned}
\]

and \( X_x(t, x)(\theta) = X(s + \theta, t, x), \theta \in [-\tau, 0]. \)

Let us introduce \( \mathcal{L}_t \). For a Borel measurable function \( f : C \to R \), we define

\[
S(f)(x) = \lim_{h \to 0^+} \frac{1}{h} \left[ f(\tilde{x}_h) - f(x) \right], \quad x \in C,
\]

where \( \tilde{x} : [-\tau, T] \to H \) is an extension of \( x \) defined by

\[
\tilde{x}(s) = \begin{cases} x(s), & s \in [-\tau, 0), \\ x(0), & s \geq 0, \end{cases}
\]

and \( \tilde{x}_\tau \) is defined by

\[
\tilde{x}_{\tau}(\theta) = \tilde{x}(s + \theta), \quad \theta \in [-\tau, 0].
\]

We denote by \( \widehat{D}(S) \) the domain of the operator \( S \), being the set of \( f : C \to R \) such that the above limit exists for all \( x \in C \). We note that \( \widehat{D}(S) \) does not contain \( C_0^1(C; R) \). We define \( D(S) \) as the set of all functions \( \Phi : [0, T] \times C \to R \) such that \( \Phi(t, \cdot) \in \widehat{D}(S) \) for all \( t \in [0, T] \).

We define \( \mathcal{L}_t \) by

\[
\mathcal{L}_t[\phi](x) = S(\phi)(x) + \left[ A\phi(0)1_0 + F(t, x)1_0, \nabla_x \phi(x) \right] + \frac{1}{2} \sum_{i=1}^{\infty} \nabla^2_x \phi(x) (G(t, x)e_i 1_0, G(t, x)e_i 1_0).
\]
where \( \{e_i\}_{i=1}^{\infty} \) is a basis of \( E \), \( \nabla_x \Phi \), \( \nabla^2_x \Phi \) are first and second Fréchet derivatives of \( \Phi \), \( \nabla_x \Phi \) and \( \nabla^2_x \Phi \) denote the extension of \( \nabla_x \Phi \) and \( \nabla^2_x \Phi \), respectively (see Lemma 2.1 and Lemma 2.2).

In this subsection, we will show that, if the function \( v(t, x) = E\Phi(X_T(t, x)) \) is sufficiently regular, then \( v(t, x) \) is a classical solution of the Kolmogorov equation:

\[
\begin{aligned}
\frac{\partial v(t, x)}{\partial t} + L_t[v(t, \cdot)](x) &= 0, \quad t \in [0, T], \quad x \in C, \quad x(0) \in D(A), \\
v(T, x) &= \Phi(x).
\end{aligned}
\tag{5.1.1}
\]

Let \( C^* \) and \( C^1 \) be the space of bounded linear functionals \( f : C \to R \) and bounded bilinear functionals \( \Phi : C \times C \to R \), of the space \( C \), respectively. We denote by \( C_{lip}^{1,2}(0, T) \times C \) the space of functions \( \Phi : [0, T] \times C \to R \) such that \( \frac{\partial \Phi}{\partial t} : [0, T] \times C \to R \) and \( \nabla^2_x \Phi : [0, T] \times C \to C^1 \) are continuous and satisfy, for a suitable constant \( K > 0 \),

\[
\|\nabla^2_x \Phi(t, x) - \nabla^2_x \Phi(t, y)\|_{C^1} \leq K \|x - y\|. \quad t \in [0, T], \quad x, y \in C.
\]

We say \( v(t, x) \) is a classical solution of Eq. (5.1.1) if \( v \in C_{lip}^{1,2}(0, T) \times C \cap D(S) \) and (5.1.1) is fulfilled for any \( t \in [0, T] \) and \( x \in C \), where \( x(0) \in D(A) \).

We define \( \hat{x} \) by

\[
\hat{x}(s) = \begin{cases} 
  x(s), & s \in [-\tau, 0), \\
  e^{sA}x(0), & s \geq 0
\end{cases}
\]

and \( \hat{x}_s \) is defined by

\[
\hat{x}_s(\theta) = \hat{x}(s + \theta), \quad \theta \in [-\tau, 0].
\]

**Lemma 5.1.1.** There exists a constant \( K > 0 \) independent of \( s \) such that

\[
\left\| \frac{1}{s - t} E(X_s(t, x) - \hat{x}_{s-t}) \right\|_{C} \leq K, \quad s \in (t, T], \quad x \in C.
\tag{5.1.2}
\]

Moreover, if \( f \in C^* \) we have that

\[
\lim_{s \to t^+} \frac{1}{s - t} E f(X_s(t, x) - \hat{x}_{s-t}) = \int f(t, x) 1_{\Theta}, \quad x \in C.
\]

**Proof.** By the definitions of \( X_s(t, x) \) and \( \hat{x}_s \), we have that, for every \( s \in [t, T] \),

\[
X_s(t, x)(\theta) - \hat{x}_{s-t}(\theta) = \begin{cases} 
  \int_t^{s+\theta} e^{(s+\theta-\sigma)A} F(\sigma, X_\sigma) d\sigma + \int_t^{s+\theta} e^{(s+\theta-\sigma)A} G(\sigma, X_\sigma) dW(\sigma), & s + \theta \geq t, \\
  0, & s + \theta < t.
\end{cases}
\]

Hence, we get that

\[
\left[ E\left( \frac{1}{s - t} (X_s(t, x) - \hat{x}_{s-t}) \right) \right](\theta) = E\left[ \frac{1}{s - t} (X_s(t, x)(\theta) - \hat{x}_{s-t}(\theta)) \right] = \begin{cases} 
  \frac{1}{s - t} \int_t^{s+\theta} E\left[ e^{(s+\theta-\sigma)A} F(\sigma, X_\sigma) \right] d\sigma, & s + \theta \geq t, \\
  0, & s + \theta < t.
\end{cases}
\]
Thus
\[
\lim_{s \to t^+} \left[ E \left( \frac{1}{s-t} (X_s(t, x) - \hat{X}_{s-t}) \right) \right](\theta) = \begin{cases} 
\lim_{s \to t^+} \frac{1}{s-t} \int_t^s E[e^{(s-\sigma)A} F(\sigma, X_\sigma)] d\sigma, & \theta = 0, \\
0, & -\tau \leq \theta < 0
\end{cases}
= F(t, x)1_{[-\tau, 0]}(\theta).
\]

Now let us prove (5.1.2). By Theorem 3.2, we get that
\[
\left| E \left( \frac{1}{s-t} (X_s(t, x) - \hat{X}_{s-t}) \right) \right| (\theta) \leq \frac{M}{s-t} \int_t^s \left| E(1 + |X_\sigma|) \right| d\sigma \leq M \sup_{s \in [t, T]} E(1 + |X_s|)
\]
\[
\leq M + C \left( 1 + \sup_{s \in [t, T]} |x| \right) =: K.
\]

Therefore, we have that \( \| \frac{1}{s-t} E(X_s(t, x) - \hat{X}_{s-t}) \|_C \leq K \) for all \( s \in (t, T] \).

Let us define \( \bar{X} \) by
\[
\bar{X}_s(t, x)(\theta) = \begin{cases}
\int_t^{s+\theta} F(t, x) d\sigma, & s + \theta \geq t, \\
0, & s + \theta < t.
\end{cases}
\]

Then
\[
\lim_{s \to t^+} \sup_{\theta \in [-\tau, 0]} \left| E \left( \frac{1}{s-t} (X_s(t, x) - \hat{X}_{s-t} - \bar{X}_s) \right) \right|(\theta)
\]
\[
= \lim_{s \to t^+} \sup_{\theta \in [-\tau, 0]} \frac{1}{s-t} \int_t^{s+\theta} \left[ e^{(s+\theta-\sigma)A} E\left[ F(\sigma, X_\sigma) - F(t, x) \right] + \left( e^{(s+\theta-\sigma)A} F(t, x) - F(t, x) \right) \right] d\sigma
\]
\[
\leq \lim_{s \to t^+} \sup_{\theta \in [-\tau, 0]} \frac{1}{s-t} \int_t^s \left| e^{(s+\theta-\sigma)A} E\left[ F(\sigma, X_\sigma) - F(t, x) \right] \right| d\sigma
\]
\[
+ \lim_{s \to t^+} \sup_{\sigma \in [0, s-t]} \left| e^{\sigma A} E\left[ F(t, x) - F(t, x) \right] \right|
\]
\[
\leq \lim_{s \to t^+} \sup_{\theta \in [-\tau, 0]} \frac{1}{s-t} \int_t^s \left| e^{(s+\theta-\sigma)A} E\left[ F(\sigma, X_\sigma) - F(\sigma, x) \right] \right| d\sigma
\]
\[
+ \int_t^s \left| e^{(s+\theta-\sigma)A} \left[ F(\sigma, x) - F(t, x) \right] \right| d\sigma
\]
\[
\leq ML \lim_{s \to t^+} \sup_{\sigma \in [t, s]} \left| X_\sigma - x \right|_C + M \lim_{s \to t^+} \sup_{\sigma \in [t, s]} \left| F(\sigma, x) - F(t, x) \right|
\]
\[
= 0,
\]
and
\[
\lim_{s \to t^+} \left[ \frac{1}{s-t} \bar{X}_s(t, x)(\theta) - F(t, x)1_{[-\tau, 0]}(\theta) \right] = -F(t, x)1_{[-\tau, 0]}(\theta).
\]
It is easy to see that \( \sup_{\theta \in [-\tau, 0]} |(1/s-t)X_s(t, x)(\theta) - F(t, x)1_{[-\tau, 0]}(\theta), h_i) | \leq |(F(t, x), h_i)| \) for all \( i \in \mathbb{N} \). If \( f \in C^* \), we get that

\[
\frac{1}{s-t}Ef (X_s(t, x) - \hat{X}_{s-t}) = f \left[ \frac{1}{s-t}E (X_s(t, x) - \hat{X}_{s-t}) \right] \\
= f \left[ E \left( \frac{1}{s-t} (X_s(t, x) - \hat{X}_{s-t} - \bar{X}_s) \right) \right] + f \left( \frac{1}{s-t} \bar{X}_s \right).
\]

By the property (V), we finally obtain that

\[
\lim_{s \to t^+} \frac{1}{s-t}Ef (X_s(t, x) - \hat{X}_{s-t}) = f(F(t, x)1_0), \quad x \in C. \quad \square
\]

For every \( \epsilon > 0 \) and a.e. \( \omega \in \Omega \), we define \( W_{t+\epsilon}^* \in C([-\tau, 0], \Xi) \) by

\[
W_{t+\epsilon}^*(\omega)(\theta) = \begin{cases} 
\frac{1}{\sqrt{\epsilon}} [W(\omega)(t+\epsilon + \theta) - W(\omega)(t)], & \theta \in (-\epsilon, 0], \\
0, & \theta \in [-\tau, -\epsilon].
\end{cases}
\]

**Lemma 5.1.2.** Let \( \beta \) be a continuous bilinear functional on \( C \). Then

\[
\lim_{\epsilon \to 0^+} \left[ \frac{1}{\epsilon} E \beta(X_{t+\epsilon} - \hat{X}_\epsilon, X_{t+\epsilon} - \hat{X}_\epsilon) - E \beta(G(t, x)W_{t+\epsilon}^*, G(t, x)W_{t+\epsilon}^*) \right] = 0. \quad (5.13)
\]

**Proof.** Firstly, we prove that

\[
\lim_{\epsilon \to 0^+} E \left\| \frac{1}{\sqrt{\epsilon}} (X_{t+\epsilon} - \hat{X}_\epsilon) - G(t, x)W_{t+\epsilon}^* \right\|_C^2 = 0, \quad (5.14)
\]

where

\[
G(t, x)W_{t+\epsilon}^*(\omega)(\theta) = \begin{cases} 
\frac{1}{\sqrt{\epsilon}} \int_t^{t+\epsilon+\theta} e^{(t+\epsilon+\theta-\sigma)} A G(t, x) dW(\sigma), & \theta \in (-\epsilon, 0], \\
0, & \theta \in [-\tau, -\epsilon].
\end{cases}
\]

Observe that

\[
\frac{1}{\sqrt{\epsilon}} (X_{t+\epsilon} - \hat{X}_\epsilon)(\theta) - G(t, x)W_{t+\epsilon}^*(\theta) \]

\[
= \begin{cases} 
\frac{1}{\sqrt{\epsilon}} \int_t^{t+\epsilon+\theta} e^{(t+\epsilon+\theta-\sigma)} A F(\sigma, X_\sigma) d\sigma \\
+ \frac{1}{\sqrt{\epsilon}} \int_t^{t+\epsilon+\theta} e^{(t+\epsilon+\theta-\sigma)} A G(\sigma, X_\sigma) dW(\sigma) \\
- \frac{1}{\sqrt{\epsilon}} \int_t^{t+\epsilon+\theta} e^{(t+\epsilon+\theta-\sigma)} A G(t, x) dW(\sigma), & \theta \in (-\epsilon, 0], \\
0, & \theta \in [-\tau, -\epsilon].
\end{cases}
\]

Therefore, for a suitable constant \( C > 0 \) that can be different in different places...
\[
\lim_{\epsilon \to 0^+} \left( E \left\| \frac{1}{\sqrt{\epsilon}} (X_{t+\epsilon} - \hat{x}_\epsilon) - G(t, x) \overline{W^*}_{t+\epsilon} \right\|^2 \right) \leq \lim_{\epsilon \to 0^+} E \sup_{\theta \in [-\tau, 0]} \left| \frac{1}{\sqrt{\epsilon}} (X_{t+\epsilon} - \hat{x}_\epsilon)(\theta) - G(t, x) \overline{W^*}_{t+\epsilon}(\theta) \right|^4 \\
\leq \frac{C}{\epsilon^2} \lim_{\epsilon \to 0^+} E \sup_{\theta \in [-\epsilon, 0]} \left| \int_t^{t+\epsilon+\theta} e^{(t+\epsilon+\theta-\sigma)A} F(\sigma, X_\sigma) \, d\sigma \right|^4 \\
+ \frac{C}{\epsilon^2} \lim_{\epsilon \to 0^+} E \sup_{\theta \in [-\epsilon, 0]} \left| \int_t^{t+\epsilon+\theta} e^{(t+\epsilon+\theta-\sigma)A} (G(\sigma, X_\sigma) - G(t, x)) \, dW(\sigma) \right|^4 \\
\leq C \lim_{\epsilon \to 0^+} \epsilon^{-1} \left( \int_t^t E |X_{\sigma} - x|^4 \, d\sigma + \int_t^{t+\epsilon} |G(\sigma, x) - G(t, x)|^4_{L_2(\mathbb{Z}, \mathcal{H})} \, d\sigma \right) = 0.
\]

Since \( \beta \) is continuous bilinear functional, for all \( \epsilon > 0 \),
\[
\left| \frac{1}{\epsilon} E \beta(X_{t+\epsilon} - \hat{x}_\epsilon, X_{t+\epsilon} - \hat{x}_\epsilon) - E \beta(G(t, x) \overline{W^*}_{t+\epsilon}, G(t, x) \overline{W^*}_{t+\epsilon}) \right| = \left| E \beta \left( \frac{1}{\sqrt{\epsilon}} (X_{t+\epsilon} - \hat{x}_\epsilon) - G(t, x) \overline{W^*}_{t+\epsilon}, \frac{1}{\sqrt{\epsilon}} (X_{t+\epsilon} - \hat{x}_\epsilon) - G(t, x) \overline{W^*}_{t+\epsilon} \right) \right| \\
+ E \beta \left( \frac{1}{\sqrt{\epsilon}} (X_{t+\epsilon} - \hat{x}_\epsilon) - G(t, x) \overline{W^*}_{t+\epsilon}, G(t, x) \overline{W^*}_{t+\epsilon} \right) \\
+ E \beta \left( G(t, x) \overline{W^*}_{t+\epsilon}, \frac{1}{\sqrt{\epsilon}} (X_{t+\epsilon} - \hat{x}_\epsilon) - G(t, x) \overline{W^*}_{t+\epsilon} \right) \\
\leq \| \beta \| E \left\| \frac{1}{\sqrt{\epsilon}} (X_{t+\epsilon} - \hat{x}_\epsilon) - G(t, x) \overline{W^*}_{t+\epsilon} \right\|^2 \\
+ 2 \| \beta \| \left[ E \left\| \frac{1}{\sqrt{\epsilon}} (X_{t+\epsilon} - \hat{x}_\epsilon) - G(t, x) \overline{W^*}_{t+\epsilon} \right\|^2 \right]^{\frac{1}{2}} \left[ E \| G(t, x) \overline{W^*}_{t+\epsilon} \|^4 \right]^{\frac{1}{2}}. \quad (5.1.5)
\]

On the other hand,
\[
E \left\| G(t, x) \overline{W^*}_{t+\epsilon} \right\|^4 = E \sup_{\theta \in [-\epsilon, 0]} \left| \int_t^{t+\epsilon+\theta} \frac{1}{\sqrt{\epsilon}} e^{(t+\epsilon+\theta-\sigma)A} G(t, x) \, dW(\sigma) \right|^4 \\
\leq \frac{M^4}{\epsilon^2} \left( \int_t^t \left\| G(t, x) \right\|^2 \, d\sigma \right)^2 = M^4 \left\| G(t, x) \right\|^4. \quad (5.1.6)
\]

Combining (5.1.5) and (5.1.6) and letting \( \epsilon \to 0^+ \) give that
\[
\lim_{\epsilon \to 0^+} \left[ \frac{1}{\epsilon} E \beta(X_{t+\epsilon} - \hat{x}_\epsilon, X_{t+\epsilon} - \hat{x}_\epsilon) - E \beta(G(t, x) \overline{W^*}_{t+\epsilon}, G(t, x) \overline{W^*}_{t+\epsilon}) \right] = 0. \quad (5.1.7)
\]
In order to prove (5.1.3), one needs to show that
\[
\lim_{\epsilon \to 0^+} \left[ E \beta \left( G(t, x) W_{t+\epsilon}^v, G(t, x) W_{t+\epsilon}^v \right) - E \beta \left( G(t, x) W_{t+\epsilon}^v, G(t, x) W_{t+\epsilon}^v \right) \right] = 0. \tag{5.1.8}
\]
We define \( G^n(t, x) \in L_2(\mathcal{E}, H) \) by
\[
G^n(t, x) y = \sum_{i=1}^{n} G(t, x) (y, e_i) e_i, \quad y \in \mathcal{E},
\]
and \( \overline{G^n(t, x)} \in L_2(\mathcal{E}, H) \) by
\[
\overline{G^n}(t, x) y = \sum_{i=n+1}^{\infty} G(t, x) (y, e_i) e_i, \quad y \in \mathcal{E}.
\]
We assume that \( G^n(t, x) y \in D(A) \) for all \( y \in \mathcal{E} \), then
\[
\lim_{\epsilon \to 0^+} E \left\| G^n(t, x) W_{t+\epsilon}^v - G^n(t, x) W_{t+\epsilon}^v \right\|_C^2
\]
\[
= \lim_{\epsilon \to 0^+} \epsilon^{-1} E \sup_{\theta \in [-\epsilon, 0]} \int_0^{\epsilon+\theta} \int_0^{t+\theta+\epsilon+\theta-\sigma} e^{rA} \overline{G^n}(t, x) dr dW(\sigma) \]
\[
= \lim_{\epsilon \to 0^+} \epsilon^{-1} E \sup_{\theta \in [-\epsilon, 0]} \int_0^{\epsilon+\theta} \int_0^{t+\theta+\epsilon+\theta-\sigma} e^{rA} \overline{G^n}(t, x) dr dW(\sigma) \]
\[
= \lim_{\epsilon \to 0^+} \epsilon^{-1} E \sup_{\theta \in [-\epsilon, 0]} \int_0^{\epsilon+\theta} e^{rA} \overline{G^n}(t, x) \left( W(t + \epsilon + \theta - r) - W(t) \right) dr \]
\[
\leq \lim_{\epsilon \to 0^+} M^2 \epsilon^{-1} \left( \int_0^{\epsilon} \sup_{\theta \in [-\epsilon, 0]} \left| AG^n(t, x) \left( W(t + \epsilon + \theta - r) - W(t) \right) \right| dr \right)^2 \]
\[
\leq \lim_{\epsilon \to 0^+} M^2 \int_0^{\epsilon} E \sup_{\theta \in [-\epsilon, 0]} \left| AG^n(t, x) \left( W(t + \epsilon + \theta - r) - W(t) \right) \right|^2 dr \]
\[
= \lim_{\epsilon \to 0^+} M^2 \int_0^{\epsilon} E \sup_{\theta \in [-\epsilon, 0]} \int_0^{t+\theta+\epsilon+\theta-r} \left| AG^n(t, x) dW(\sigma) \right|^2 dr \]
\[
\leq \lim_{\epsilon \to 0^+} \frac{1}{2} M^2 \left| AG^n(t, x) \right|_{L_2(\mathcal{E}, H)}^2 \epsilon^2 \]
\[
= 0.
\]
For the general $G^n(t, x) \in L_2(\mathcal{E}, H)$, we may assume there exists a sequence $\{x_m\}_{m \geq 1}$ such that $x_m y \in D(A)$ for all $m \geq 1$, $y \in \mathcal{E}$ and $x_m \to G^n(t, x)$ in $L_2(\mathcal{E}, H)$ as $m \to \infty$. Then for a suitable constant $C > 0$,

$$E \|G^n(t, x)W^*_{t+\epsilon} - G^n(t, x)\overline{W}^*_{t+\epsilon} - (x_mW^*_{t+\epsilon} - x_m\overline{W}^*_{t+\epsilon})\|^2_C$$

$$\leq \epsilon^{-1} E \sup_{\theta \in [-\epsilon, 0]} \left| \int_t^{t+\epsilon+\theta} e^{(t+\epsilon+\theta - \sigma)A} - I \right| (G^n(t, x) - x_m) dW(\sigma)$$

$$\leq \epsilon^{-1} CE \int_t^{t+\epsilon} \|G^n(t, x) - x_m\|^2_{L_2(\mathcal{E}, H)} d\sigma$$

$$= C \|G^n(t, x) - x_m\|^2_{L_2(\mathcal{E}, H)}.$$  

Hence

$$E \|G^n(t, x)W^*_{t+\epsilon} - G^n(t, x)\overline{W}^*_{t+\epsilon}\|^2_C \leq 2E \|G^n(t, x)W^*_{t+\epsilon} - G^n(t, x)\overline{W}^*_{t+\epsilon} - (x_mW^*_{t+\epsilon} - x_m\overline{W}^*_{t+\epsilon})\|^2_C$$

$$+ 2\|x_mW^*_{t+\epsilon} - x_m\overline{W}^*_{t+\epsilon}\|^2_C$$

$$\leq 2C \|G^n(t, x) - x_m\|^2_{L_2(\mathcal{E}, H)} + 2\|x_mW^*_{t+\epsilon} - x_m\overline{W}^*_{t+\epsilon}\|^2_C.$$ 

For any $\epsilon > 0$, firstly, letting $m$ be large enough such that $2C \|G^n(t, x) - x_m\|^2_{L_2(\mathcal{E}, H)} < \frac{\epsilon}{2}$, then for the fixed $m$, there exists a constant $\delta > 0$ such that $2\|x_mW^*_{t+\epsilon} - x_m\overline{W}^*_{t+\epsilon}\|^2_C < \frac{\epsilon}{2}$ for all $\epsilon < \delta$. Therefore, we get that

$$\lim_{\epsilon \to 0^+} E \|G^n(t, x)W^*_{t+\epsilon} - G^n(t, x)\overline{W}^*_{t+\epsilon}\|^2_C = 0.$$ 

Finally, we have that

$$E \|G(t, x)W^*_{t+\epsilon} - G(t, x)\overline{W}^*_{t+\epsilon}\|^2_C$$

$$\leq \epsilon^{-1} E \sup_{\theta \in [-\epsilon, 0]} \left| \int_t^{t+\epsilon+\theta} e^{(t+\epsilon+\theta - \sigma)A} G(t, x) - G(t, x) dW(\sigma) \right|^2$$

$$\leq 2E \|G^n(t, x)W^*_{t+\epsilon} - G^n(t, x)\overline{W}^*_{t+\epsilon}\|^2_C$$

$$+ 2\epsilon^{-1} E \sup_{\theta \in [-\epsilon, 0]} \left| \int_t^{t+\epsilon+\theta} e^{(t+\epsilon+\theta - \sigma)A} G^n(t, x) - G^n(t, x) dW(\sigma) \right|^2$$

$$\leq 2E \|G^n(t, x)W^*_{t+\epsilon} - G^n(t, x)\overline{W}^*_{t+\epsilon}\|^2_C + 2(M^2 + 1) \left| G^n(t, x) \overline{G^n(t, x)} \right|^2_{L_2(\mathcal{E}, H)}$$

$$= 2E \|G^n(t, x)W^*_{t+\epsilon} - G^n(t, x)\overline{W}^*_{t+\epsilon}\|^2_C + 2(M^2 + 1) \sum_{i=n+1}^{\infty} \left| G(t, x)e_i \right|^2_H.$$
Since $|G(t, x)|_{L^2(H)} < \infty$, let $n$ be large enough such that $2(M^2 + 1) \sum_{i=n+1}^{\infty} |G(t, x)\varepsilon_i|^2 < \frac{\varepsilon}{2}$, then for the fixed $n$, there exists a constant $\delta > 0$ such that $2\|G^n(t, x)\mathcal{W}^*_{t+\varepsilon} - G^n(t, x)\mathcal{W}^*_{t+\varepsilon}\|_C^2 < \frac{\varepsilon}{2}$ for all $\varepsilon < \delta$. Therefore, we get that

$$\lim_{\varepsilon \to 0^+} E\|G(t, x)\mathcal{W}^*_{t+\varepsilon} - G(t, x)\mathcal{W}^*_{t+\varepsilon}\|_C^2 = 0.$$  

By the same steps as above, we can show that (5.1.8) holds. Then, combining (5.1.7) and (5.1.8) we get (5.1.3). □

**Theorem 5.1.3.** Suppose $\Phi \in C^{1,2}_{lip}([0, T] \times \mathcal{C}) \cap D(S)$. Then, for each $x \in \mathcal{C}$ and $x(0) \in D(A)$, we have that

$$\lim_{\varepsilon \to 0^+} \frac{E[\Phi(t + \varepsilon, X_{t+\varepsilon}(t, x)) - \Phi(t, x)]}{\varepsilon} = \frac{\partial}{\partial t} \Phi(t, x) + S(\Phi)(t, x) + \left\{ A\Phi(0)1_0 + F(t, x)1_0, \nabla_x \Phi(x) \right\}$$

$$+ \frac{1}{2} \sum_{i=1}^{\infty} \nabla_x^2 \Phi(t, x) \left( G(t, x)\varepsilon_i1_0, G(t, x)\varepsilon_i1_0 \right). \quad (5.1.9)$$

**Proof.** Since $\Phi \in C^{1,2}_{lip}([0, T] \times \mathcal{C})$, then by Taylor’s theorem we get that

$$\Phi(t + \varepsilon, X_{t+\varepsilon}(t, x)) - \Phi(t, x) = \Phi(t + \varepsilon, X_{t+\varepsilon}(t, x)) - \Phi(t, X_{t+\varepsilon}(t, x)) + \Phi(t, \hat{x}_\varepsilon) - \Phi(t, x)$$

$$+ \nabla_x \Phi(t, \hat{x}_{t+\varepsilon})(X_{t+\varepsilon} - \hat{x}_\varepsilon) + R_2(\varepsilon), \quad a.s., \quad t > 0,$n

where

$$R_2(\varepsilon) = \int_0^1 (1-s) \nabla_x^2 \Phi[t, \hat{x}_\varepsilon + s(X_{t+\varepsilon}(t, x) - \hat{x}_\varepsilon)](X_{t+\varepsilon}(t, x) - \hat{x}_\varepsilon, X_{t+\varepsilon}(t, x) - \hat{x}_\varepsilon) ds, \quad a.s.$$

Taking expectations, we show that

$$\frac{1}{\varepsilon} E[\Phi(t + \varepsilon, X_{t+\varepsilon}(t, x)) - \Phi(t, x)]$$

$$= \frac{1}{\varepsilon} E[\Phi(t + \varepsilon, X_{t+\varepsilon}(t, x)) - \Phi(t, X_{t+\varepsilon}(t, x))] + \frac{1}{\varepsilon} E[\Phi(t, \hat{x}_\varepsilon) - \Phi(t, x)]$$

$$+ \frac{1}{\varepsilon} E[\nabla_x \Phi(t, \hat{x}_\varepsilon)(X_{t+\varepsilon}(t, x) - \hat{x}_\varepsilon)] + \frac{1}{\varepsilon} ER_2(\varepsilon). \quad (5.1.10)$$

From $\Phi \in C^{1,2}_{lip}([0, T] \times \mathcal{C})$, it follows that

$$\lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} E[\Phi(t + \varepsilon, X_{t+\varepsilon}(t, x)) - \Phi(t, X_{t+\varepsilon}(t, x))] = \frac{\partial}{\partial t} \Phi(t, x).$$

By $\Phi \in D(S)$, we have that

$$\lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \left[ \Phi(t, \hat{x}_\varepsilon) - \Phi(t, x) \right] = S(\Phi)(x).$$
Let us define \( \tilde{x} \) by

\[
\tilde{x}_e (\theta) = \begin{cases} 
\int_0^{\epsilon + \theta} Ax(0) \, d\sigma, & \epsilon + \theta \geq 0, \\
0, & \epsilon + \theta < 0.
\end{cases}
\]

By the definition of \( \hat{x} \) and \( \tilde{x} \),

\[
\hat{x}_e (\theta) - \tilde{x}_e (\theta) = [e^{(\epsilon + \theta)A}x(0) - x(0)]1_{(-\epsilon, 0]}(\theta).
\]

Then we show that

\[
\lim_{\epsilon \to 0^+} \sup_{\theta \in [-\tau, 0]} \left| \frac{1}{\epsilon} [\hat{x}_e - \tilde{x}_e] (\theta) \right| = \lim_{\epsilon \to 0^+} \sup_{\theta \in [-\tau, 0]} \left| \frac{1}{\epsilon} \int_0^{\epsilon + \theta} [e^{A\sigma} Ax(0) - Ax(0)] \, d\sigma \right| \\
\leq \lim_{\epsilon \to 0^+} \sup_{\sigma \in [0, \epsilon]} \left| e^{A\sigma} Ax(0) - Ax(0) \right| = 0.
\]

Using the continuity of \( \nabla^2_x \Phi(t, \cdot) \), for every \( l \in [0, 1] \), there exists a constant \( r \in (0, 1) \) such that

\[
\left| \frac{1}{\epsilon} \nabla_x \Phi(t, \hat{x}_e + l(\hat{x}_e - \tilde{x}_e)) (\hat{x}_e - \tilde{x}_e) - \frac{1}{\epsilon} \nabla_x \Phi(t, x) (\hat{x}_e - \tilde{x}_e) \right| \\
\leq \left\| \nabla_x \Phi(t, \hat{x}_e + l(\hat{x}_e - \tilde{x}_e)) - \nabla_x \Phi(t, x) \right\| \left( 1 \frac{1}{\epsilon} (\hat{x}_e - \tilde{x}_e) \right) \\
\leq \left\| \nabla^2_x \Phi(t, x + r(\hat{x}_e + l(\hat{x}_e - \tilde{x}_e) - x)) \right\| \left\| \hat{x}_e + l(\hat{x}_e - \tilde{x}_e) - x \right\| \left( 1 \frac{1}{\epsilon} (\hat{x}_e - \tilde{x}_e) \right) \\
\leq \left( \left\| \nabla^2_x \Phi(t, x) \right\| + K \left\| \hat{x}_e - x \right\| + K \left\| \hat{x}_e - \tilde{x}_e \right\| \right) \left( \left\| \hat{x}_e - x \right\| + \left\| \hat{x}_e - \tilde{x}_e \right\| \right) \left( 1 \frac{1}{\epsilon} (\hat{x}_e - \tilde{x}_e) \right).
\]

Letting \( \epsilon \to 0^+ \) we get that

\[
\lim_{\epsilon \to 0^+} \frac{1}{\epsilon} \nabla_x \Phi(t, \hat{x}_e + l(\hat{x}_e - \tilde{x}_e)) (\hat{x}_e - \tilde{x}_e) = \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} \nabla_x \Phi(t, x)(\hat{x}_e - \tilde{x}_e)
\]

uniformly in \( l \in [0, 1] \). Therefore,

\[
\lim_{\epsilon \to 0^+} \frac{1}{\epsilon} \left[ \Phi(t, \hat{x}_e) - \Phi(t, x) \right] = \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} \left[ \Phi(t, \hat{x}_e) - \Phi(t, \tilde{x}_e) + \Phi(t, \tilde{x}_e) - \Phi(t, x) \right] \\
= \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} \nabla_x \Phi(t, \hat{x}_e + l(\hat{x}_e - \tilde{x}_e)) (\hat{x}_e - \tilde{x}_e) + S(\Phi)(x) \\
= \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} \nabla_x \Phi(t, x)(\hat{x}_e - \tilde{x}_e) + S(\Phi)(x) \\
= \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} \nabla_x \Phi(t, x)(\hat{x}_e - \tilde{x}_e) + \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} \nabla_x \Phi(t, x)\tilde{x}_e + S(\Phi)(x) \\
= \langle Ax(0)1_0, \nabla_x \Phi(t, x) \rangle + S(\Phi)(x). \quad (5.1.11)
\]
Now let us consider the limit \( \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} E \nabla_x \Phi(t, \tilde{x}_e)(X_{t+\epsilon}(t, x) - \tilde{x}_e) \). By Lemma 5.1.1 and the continuity of \( \nabla_x \Phi(t, \cdot) \), we have

\[
\lim_{\epsilon \to 0^+} \frac{1}{\epsilon} E \nabla_x \Phi(t, \tilde{x}_e)(X_{t+\epsilon}(t, x) - \tilde{x}_e) - \frac{1}{\epsilon} E \nabla_x \Phi(t, x)(X_{t+\epsilon}(t, x) - \tilde{x}_e) = \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} E \nabla_x \Phi(t, \tilde{x}_e)(X_{t+\epsilon}(t, x) - \tilde{x}_e)
\]

Then, by Lemma 5.1.1, we obtain that

\[
\lim_{\epsilon \to 0^+} \frac{1}{\epsilon} E \nabla_x \Phi(t, \tilde{x}_e)(X_{t+\epsilon}(t, x) - \tilde{x}_e) = \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} E \nabla_x \Phi(t, x)(X_{t+\epsilon}(t, x) - \tilde{x}_e) = \nabla_x \Phi(t, x)(F(t, x)1_0).
\]

Finally, we look at the limit \( \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} E R_2(\epsilon) \). By the Burkholder–Davis–Gundy inequalities, for some constant \( C \) that may vary from line to line, we have

\[
E \left\| X_{t+\epsilon}(t, x) - \tilde{x}_e \right\|^4 \leq 8M^4 \epsilon \left( \int_{t}^{t+\epsilon} \left| F(\sigma, X_{\sigma}) \right| d\sigma \right)^4 + 8CE \left( \int_{t}^{t+\epsilon} \left| G(\sigma, X_{\sigma}) \right|_{L_2(\Xi, H)}^2 d\sigma \right)^2 \leq C(\epsilon^2 + \epsilon^4).
\]

Furthermore, for \( l \in [0, 1] \), we have that

\[
\left| \frac{1}{\epsilon} E \nabla_x^2 \Phi(t, \tilde{x}_e + l(X_{t+\epsilon}(t, x) - \tilde{x}_e))(X_{t+\epsilon}(t, x) - \tilde{x}_e, X_{t+\epsilon}(t, x) - \tilde{x}_e) - \frac{1}{\epsilon} E \nabla_x^2 \Phi(t, x)(X_{t+\epsilon}(t, x) - \tilde{x}_e, X_{t+\epsilon}(t, x) - \tilde{x}_e) \right|
\]

\[
\leq \frac{1}{\epsilon} E \left( \left\| \nabla_x^2 \Phi(t, \tilde{x}_e + l(X_{t+\epsilon}(t, x) - \tilde{x}_e)) \right\| X_{t+\epsilon}(t, x) - \tilde{x}_e \right\|^2 \right)
\]

\[
\leq \frac{1}{\epsilon} \left[ E \left\| \nabla_x^2 \Phi(t, \tilde{x}_e + l(X_{t+\epsilon}(t, x) - \tilde{x}_e)) \right\|^2 \left\| X_{t+\epsilon}(t, x) - \tilde{x}_e \right\|^4 \right]^\frac{1}{2}
\]

\[
\leq C \left( 1 + \epsilon^2 \right)^\frac{1}{2} \left[ E \left\| \nabla_x^2 \Phi(t, \tilde{x}_e + l(X_{t+\epsilon}(t, x) - \tilde{x}_e)) \right\|^2 \left\| X_{t+\epsilon}(t, x) - \tilde{x}_e \right\|^4 \right]^\frac{1}{2}.
\]

But
\[
E \| \nabla_x^2 \Phi (t, \hat{x}_e + l(X_{t+\epsilon} (t, x) - \hat{x}_e)) - \nabla_x^2 \Phi (t, x) \|^2 \\
\leq 2 K^2 \big( \| \hat{x}_e - x \|^2_C + (E \| X_{t+\epsilon} (t, x) - \hat{x}_e \|_C^4 \big)^{1/2} \\
\leq 2 K^2 \big( \| \hat{x}_e - x \|^2_C + C^2 \epsilon (1 + \epsilon^2)^{1/2} \big). 
\]

Letting \( \epsilon \to 0^+ \) in (5.1.13) and (5.1.14), we get that
\[
\lim_{\epsilon \to 0^+} \frac{1}{E} E \| \nabla_x^2 \Phi (t, \hat{x}_e + l(X_{t+\epsilon} (t, x) - \hat{x}_e)) (X_{t+\epsilon} (t, x) - \hat{x}_e, X_{t+\epsilon} (t, x) - \hat{x}_e) \\
= \lim_{\epsilon \to 0^+} \frac{1}{E} E \| \nabla_x^2 \Phi (t, x) (X_{t+\epsilon} (t, x) - \hat{x}_e, X_{t+\epsilon} (t, x) - \hat{x}_e) \\
\text{uniformly in } l \in [0, 1].
\]
We define a continuous bilinear functional \( \gamma : C([-\tau, 0]; R^m) \times C([-\tau, 0]; R^m) \to R \) by
\[
\gamma (\tilde{y}, \tilde{z}) = \nabla_x^2 \Phi (t, x) (G(t, x) y, G(t, x) z), \quad y, z \in C([-\tau, 0]; Z),
\]
where \( y = \sum_{i=1}^m y_i e_i, z = \sum_{i=1}^m z_i e_i, y_i, z_i \in C([-\tau, 0], R) \) and \( \tilde{y} = (y_1, y_2, \ldots, y_m), \tilde{z} = (z_1, z_2, \ldots, z_m) \in C([-\tau, 0]; R^m) \). Let \( W^{*, m}_{t+\epsilon} = \sum_{i=1}^m (W^*_{t+\epsilon}, e_i) e_i \) and \( W^{*, m}_{t+\epsilon} = \sum_{i=m+1}^\infty (W^*_{t+\epsilon}, e_i) e_i \) for every \( m \geq 1 \). From Lemma 3.5 on p. 88 of [28], it follows that, for every \( m \geq 1 \),
\[
\lim_{\epsilon \to 0^+} E \gamma (W^{*, m}_{t+\epsilon}, W^{*, m}_{t+\epsilon}) = \sum_{i=1}^m \gamma (k_i 1_0, k_i 1_0),
\]
where \( \{ k_i \}_{i=1}^m \) is the canonical bases of \( R^m \).

Now let us show that \( \gamma (k_1 1_0, k_1 1_0) = \nabla_x^2 \Phi (t, x) (G(t, x) e_1) 1_0, G(t, x) e_1 1_0) \). Take
\[
\chi^\theta (\theta) = \begin{cases} 
  e_i, & \theta \in [-\tau, -\frac{1}{n}], \\
  -n\theta e_i, & \theta \in (-\frac{1}{n}, 0]. 
\end{cases}
\]
Clearly, \( (G(t, x)x_n) \to (G(t, x) e_1) 1_0 \) as \( n \to \infty \) and \( \sup_{n \geq 1} |(G(t, x)x_n, h_j)| \leq |(G(t, x) e_1, h_j)| \). By property (W), we have that
\[
\nabla_x^2 \Phi (t, x) (G(t, x) e_1) 1_0, G(t, x) e_1 1_0) \\
= \nabla_x^2 \Phi (t, x) (G(t, x) e_1) 1_{[-\tau, 0]}, G(t, x) e_1 1_{[-\tau, 0)]} \\
- \nabla_x^2 \Phi (t, x) (G(t, x) e_1) 1_{[-\tau, 0)}, G(t, x) e_1 1_{[-\tau, 0)} \\
+ \nabla_x^2 \Phi (t, x) (G(t, x) x_n, G(t, x) e_1) 1_{[-\tau, 0]} \\
+ \nabla_x^2 \Phi (t, x) (G(t, x) x_n, G(t, x) e_1) 1_{[-\tau, 0)} \\
- \nabla_x^2 \Phi (t, x) (G(t, x) x_n, G(t, x) x_n) \\
= \gamma (k_1 1_{[-\tau, 0]}, k_1 1_{[-\tau, 0]} - \lim_{n \to \infty} \gamma (k_1 1_{[-\tau, 0]}, (x_n, e_i) k_j) \\
+ \gamma (x_n, e_i) k_j, k_i 1_{[-\tau, 0]} - \gamma (x_n, e_i) k_i, (x_n, e_i) k_j) \\
= \gamma (k_1 1_{[-\tau, 0]}, k_1 1_{[-\tau, 0)} - \gamma (k_1 1_{[-\tau, 0]), k_i 1_{[-\tau, 0)} \\
- \gamma (k_1 1_{[-\tau, 0)}, k_1 1_{[-\tau, 0)} + \gamma (k_1 1_{[-\tau, 0), k_1 1_{[-\tau, 0)} \\
= \gamma (k_1 1_0, e_1 1_0).
Thus
\[
\lim_{\epsilon \to 0^+} E\nabla^2 \Phi(t, x)(G(t, x) W^*_t + \epsilon, G(t, x) W^*_t + \epsilon) = \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} E\gamma^m(\tilde{W}^*_t + \epsilon, \tilde{W}^*_t + \epsilon) = \sum_{i=1}^m \nabla^2 \Phi(t, x)(G(t, x)e_i 1_0, G(t, x)e_i 1_0).
\]

Since
\[
\left| E\nabla^2 \Phi(t, x)(G(t, x) W^*_t + \epsilon, G(t, x) W^*_t + \epsilon) - \sum_{i=1}^m \nabla^2 \Phi(t, x)(G(t, x)e_i 1_0, G(t, x)e_i 1_0) \right| \leq \left| E\nabla^2 \Phi(t, x)(G(t, x) W^*_t + \epsilon, G(t, x) W^*_t + \epsilon) - \sum_{i=1}^m \nabla^2 \Phi(t, x)(G(t, x)e_i 1_0, G(t, x)e_i 1_0) \right|
\]
\[
+ \left| E\nabla^2 \Phi(t, x)(G(t, x) W^*_t + \epsilon, G(t, x) W^*_t + \epsilon) - \sum_{i=m+1}^\infty \nabla^2 \Phi(t, x)(G(t, x)e_i 1_0, G(t, x)e_i 1_0) \right|
\]
\[
+ 5 \left\| \nabla^2 \Phi(t, x) \right\| \left( \sum_{i=m+1}^\infty \left| G(t, x)e_i \right|^2 \right),
\]
by the similar procedure as in Lemma 5.1.2, we can show that
\[
\lim_{\epsilon \to 0^+} E\nabla^2 \Phi(t, x)(G(t, x) W^*_t + \epsilon, G(t, x) W^*_t + \epsilon) = \sum_{i=1}^\infty \nabla^2 \Phi(t, x)(G(t, x)e_i 1_0, G(t, x)e_i 1_0).
\]

From this and Lemma 5.1.2, by the dominated convergence theorem, one gets
\[
\lim_{\epsilon \to 0^+} \frac{1}{\epsilon} E\mathcal{R}_2(\epsilon) = \int_0^1 (1-s) \lim_{\epsilon \to 0^+} E\nabla^2 \Phi(t, x)(G(t, x) W^*_t + \epsilon, G(t, x) W^*_t + \epsilon) \, ds
\]
\[
= \frac{1}{2} \sum_{i=1}^\infty \nabla^2 \Phi(t, x)(G(t, x)e_i 1_0, G(t, x)e_i 1_0).
\]

Letting $\epsilon \to 0^+$ in (5.1.10) and putting together the results of (5.1.11), (5.1.12) and (5.1.15), we finally obtain (5.1.9). The proof is finished. \hfill \Box

**Corollary 5.1.4.** If the function $v(t, x) \in C^{1,2}_{\text{lip}}([0, T] \times \mathcal{C}) \cap D(S)$, where $v(t, x) = E\phi(X_t(t, x))$. Then, $v(t, x)$ is a classical solution of (5.1).\hfill \Box

**Proof.** From the definition of $v(t, x)$, it follows that $v(T, x) = \phi(x)$. Moreover, by Corollary 4.4 in this case of $\psi \equiv 0$, we have that
\[
E v(t + \epsilon, X_{t+\epsilon}) = E Y(t + \epsilon, t, x) = E Y(t, t, x) = v(t, x).
\]
Hence the statement of the corollary follows by Theorem 5.1.3. \hfill \Box
5.2. Mild solution of the Kolmogorov equation

Let \( X(·, t, x) \) denote the mild solution of (3.3). In order to study the mild solution of the Kolmogorov equation, we first introduce the Markov property of stochastic process \( X_s(t, x), s \in [t, T] \).

**Theorem 5.2.1.** Assume Hypothesis 3.1(i)–(iii) holds. Then for arbitrary \( f \in B_b(C) \) and \( 0 \leq t \leq l \leq s \leq T \), we have that

\[
E\left[ f(X_s(t, x)) \mid F_l \right] = E\left( f(X_s(l, y)) \right) \bigg|_{y=X_l(t, x)}, \quad \text{P-a.s.}
\]

**Proof.** Without any loss of generality, we assume that \( f \in C_b(C) \). By Lemma 1.1 in [33] there exists a sequence of \( C \)-valued \( F_l \)-measurable simple functions

\[
f_m : \Omega \to C, \quad f_m = \sum_{k=1}^{N_m} h_k^{(m)} I_{\{f_m = h_k^{(m)}\}}, \quad N_m \in N,
\]

where \( h_1^{(m)}, \ldots, h_m^{(m)} \) are pairwise distinct and \( \Omega = \bigcup_{k=1}^{N_m} \{f_m = h_k^{(m)}\} \), such that

\[
|f_m(\omega) - X_l(t, x)(\omega)| \downarrow 0 \quad \text{for all } \omega \in \Omega \text{ as } n \to \infty.
\]

For any \( B \in F_l \), we have

\[
\int_B f(X_s(t, x)) \, dP = EI_B f(X_s(l, X_l(t, x)))
\]

\[
= \lim_{m \to \infty} EI_B f(X_s(l, f_m(\omega))) = \lim_{m \to \infty} \sum_{k=1}^{N_m} E( I_B I_{\{f_m = h_k^{(m)}\}} f(X_s(l, h_k^{(m)})))
\]

\[
= \lim_{m \to \infty} \sum_{k=1}^{N_m} \left[ E(I_B I_{\{f_m = h_k^{(m)}\}} f(X_s(l, h_k^{(m)}))) \right]
\]

\[
= \lim_{m \to \infty} E \left[ I_B \sum_{k=1}^{N_m} I_{\{f_m = h_k^{(m)}\}} f(X_s(l, h_k^{(m)})) \right]
\]

\[
= \lim_{m \to \infty} E[I_B f(X_s(l, y)) \bigg|_{y=f_m}] = \int_B Ef(X_s(l, y)) \bigg|_{y=X_l(t, x)} \, dP,
\]

and we get that

\[
E\left[ f(X_s(t, x)) \mid F_l \right] = Ef(X_s(l, y)) \bigg|_{y=X_l(t, x)}, \quad \text{P-a.s.}
\]

The proof is finished. \( \Box \)

We denote by \( B(C) \) the set of measurable functions \( \phi : C \to R \) with polynomial growth. The transition semigroup \( P_{t,s} \) is defined for arbitrary \( \phi \in B(C) \) by the formula

\[
P_{t,s} \phi(x) = E\phi(X_s(t, x)), \quad x \in C,
\]
where $X(\cdot, t, x)$ denotes the solution of (3.3). The estimate $E \sup_{t \in [t-\tau, T]} |X(s, t, x)|^p \leq C (1 + |x|)^p$ shows that $P_{t,s}$ is well defined from $\mathcal{B}(C)$ to itself. Clearly, $P_{t,s} P_{s,l} = P_{t,l}$, $t \leq s \leq l$.

We study a generalization Kolmogorov equation of the following form:

$$\frac{\partial v(t, x)}{\partial t} + \mathcal{L}_t[v(t, \cdot)](x) = \psi(t, x, v(t, x), \overline{\nu_0} v(t, x) G(t, x)), \quad t \in [0, T], x \in C, x(0) \in D(A),$$

$$v(T, x) = \phi(x). \tag{5.2.1}$$

**Definition 5.2.2.** We say a function $v : [0, T] \times C \to R$ is a mild solution of the nonlinear Kolmogorov equation (5.2.1) if $v \in \mathcal{G}^{0,1}([0, T] \times C; R)$, there exist some constants $C > 0$, $q \geq 0$ such that

$$|\nabla_x v(t, x) h| \leq C |h| (1 + |x|)^q, \quad t \in [0, T], x, h \in C, \tag{5.2.2}$$

and the following equality holds true,

$$v(t, x) = \int_t^T P_{t,s}[\psi(s, \cdot, v(s, \cdot), \overline{\nu_0} v(s, \cdot) G(s, \cdot))](x) \, ds + P_{t,T}[\phi](x), \quad t \in [0, T], x \in C. \tag{5.2.3}$$

We note that the formula (5.2.3) is meaningful if $\nabla_0 v$ is well defined and provided $\psi(t, \cdot, \cdot, \cdot)$ and $v(t, \cdot)$ satisfy polynomial growth and measurability conditions.

**Theorem 5.2.3.** Assume that Hypotheses 3.1 and 4.1 hold. Then there exists a unique mild solution $v$ of the nonlinear Kolmogorov equation (5.2.1). The function $v$ coincides with the one introduced in Corollary 4.4.

**Proof.** (Existence) Let $v$ be the function defined in Corollary 4.4. Then, $v$ has the regularity properties stated in Definition 5.2.1. It remains to prove that equality (5.2.3) holds true. It follows from Corollary 4.4 that

$$P_{t,s}[\psi(s, \cdot, v(s, \cdot), \overline{\nu_0} v(s, \cdot) G(s, \cdot))](x)$$

$$= E[\psi(s, X_s(t, x), v(s, X_s(t, x)), \overline{\nu_0} v(s, X_s(t, x)) G(s, X_s(t, x)))]$$

$$= E \psi(s, X_s(t, x), Y(s, t, x), Z(s, t, x)).$$

Hence we get that

$$\int_t^T P_{t,s}[\psi(s, \cdot, v(s, \cdot), \overline{\nu_0} v(s, \cdot) G(s, \cdot))](x) \, ds = E \int_t^T \psi(s, X_s(t, x), Y(s, t, x), Z(s, t, x)) \, ds. \tag{5.2.4}$$

The backward equation of system (4.1) is

$$Y(t, t, x) + \int_t^T Z(s, t, x) \, dW(s) = \int_t^T \psi(s, X_s(t, x), Y(s, t, x), Z(s, t, x)) \, ds + \phi(X_T(t, x)).$$

Taking expectation and applying (5.2.4) we obtain the equality (5.2.3).
(Uniqueness) Let \( v \) be a mild solution of \((5.2.1)\), by \((5.2.3)\) we have, for every \( s \in [t, T] \) and \( x \in \mathcal{C} \),

\[
v(s, x) = E\phi(X_T(s, x)) + E \int_s^T \left[ \psi(\sigma, X_{\sigma}(s, x), v(\sigma, X_{\sigma}(s, x)), \nabla_0 v(\sigma, X_{\sigma}(s, x)) G(\sigma, X_{\sigma}(s, x)) \right] \, d\sigma.
\]

Since \( X_{\sigma}(s, x) \) is independent of \( \mathcal{F}_s \), we can replace the expectation occurring in the last formula by conditional expectation given by \( \mathcal{F}_s \). By \( X_{\sigma}(s, x) \) is \( \mathcal{F}_s \)-measurable, we can replace \( x \) by \( X_s(t, x) \). Using the identity \( X_{\sigma}(s, X_s(t, x)) = X_{\sigma}(t, x) \), we obtain that

\[
v(s, X_s(t, x)) = E[\phi(X_T(t, x))|\mathcal{F}_s]
\]

\[+ E \left[ \int_s^T \left[ \psi(\sigma, X_{\sigma}(t, x), v(\sigma, X_{\sigma}(t, x)), \nabla_0 v(\sigma, X_{\sigma}(t, x)) G(\sigma, X_{\sigma}(t, x)) \right] \, d\sigma \bigg| \mathcal{F}_s \right]
\]

\[= E[\xi|\mathcal{F}_s] - E \left[ \int_s^T \left[ \psi(\sigma, X_{\sigma}(t, x), v(\sigma, X_{\sigma}(t, x)), \nabla_0 v(\sigma, X_{\sigma}(t, x)) G(\sigma, X_{\sigma}(t, x)) \right] \, d\sigma \bigg| \mathcal{F}_s \right].
\]

where

\[\xi = \phi(X_T(t, x)) + \int_t^T \left[ \psi(\sigma, X_{\sigma}(t, x), v(\sigma, X_{\sigma}(t, x)), \nabla_0 v(\sigma, X_{\sigma}(t, x)) G(\sigma, X_{\sigma}(t, x)) \right] \, d\sigma.
\]

By the representation theorem (see Proposition 4.1 in \([22]\)), there exists \( \tilde{Z} \in L_\mathcal{P}^2(\Omega \times [0, T], \mathcal{F}) \) such that \( E[\xi|\mathcal{F}_s] = \nu(t, x) + \int_t^s \tilde{Z}(\sigma) \, dW(\sigma) \), \( s \in [t, T] \). Therefore,

\[
v(s, X_s(t, x)) = \nu(t, x) + \int_t^s \tilde{Z}(\sigma) \, dW(\sigma)
\]

\[+ \int_t^s \psi(\sigma, X_{\sigma}(t, x), v(\sigma, X_{\sigma}(t, x)), \nabla_0 v(\sigma, X_{\sigma}(t, x)) G(\sigma, X_{\sigma}(t, x)) \, d\sigma. \quad (5.2.5)
\]

By Theorem 4.2 we have that \( \langle \nu(\cdot, X_{\sigma}(t, x)), \omega \rangle_{\mathcal{H}^s[0, T]} = \int_t^T \nabla_0 v(\sigma, X_{\sigma}(t, x)) G(\sigma, X_{\sigma}(t, x)) d\omega, \quad (5.2.18\text{a}) \). Hence, \( \tilde{Z}(\sigma) = \nabla_0 v(\sigma, X_{\sigma}(t, x)) G(\sigma, X_{\sigma}(t, x)) \), and by \( \nu(T, X_T(t, x)) = \phi(X_T(t, x)) \) equality \((5.2.5)\) can be rewritten as

\[
v(s, X_s(t, x)) = \nu(t, x) + \int_t^s \nabla_0 v(\sigma, X_{\sigma}(t, x)) G(\sigma, X_{\sigma}(t, x)) \, dW(\sigma)
\]

\[+ \int_t^s \psi(\sigma, X_{\sigma}(t, x), v(\sigma, X_{\sigma}(t, x)), \nabla_0 v(\sigma, X_{\sigma}(t, x)) G(\sigma, X_{\sigma}(t, x)) \, d\sigma
\]
\[
\begin{align*}
\phi(X_T(t, x)) - \int_0^T \nabla_0 \mathbf{v}(\sigma, X_T(t, x)) G(\sigma, X_T(t, x)) \, dW(\sigma) \\
+ \int_0^T \psi(\sigma, X_T(t, x)), v(\sigma, X_T(t, x)), \nabla_0 \mathbf{v}(\sigma, X_T(t, x)) G(\sigma, X_T(t, x)) \, d\sigma.
\end{align*}
\]

Comparing with the backward equation of system (4.1) we see that the pairs \((Y(s, t, x), Z(s, t, x))\) and \((\mathbf{v}(s, X_T(t, x)), \nabla_0 \mathbf{v}(s, X_T(t, x)) G(s, X_T(t, x)))\), \(s \in [t, T]\), solve the same equation. By uniqueness, we have \(Y(s, t, x) = \mathbf{v}(s, X_T(t, x)), s \in [t, T]\). Setting \(s = t\) we show \(Y(t, t, x) = \mathbf{v}(t, x)\). \(\square\)

6. Application to optimal control

In this section we consider the controlled state equation:

\[
\begin{cases}
    dX^u(s) = AX^u(s) \, ds + F(s, X^u(s)) \, ds + G(s, X^u(s)) R(s, X^u(s)) \, ds \\
    \quad + G(s, X^u(s)) \, dW(s), \quad s \in [t, T], \\
    X^u_t = x.
\end{cases}
\]

The solution of this equation will be denoted by \(X^u(s, t, x)\) or simply by \(X^u(s)\). We consider a cost functional of the form:

\[
J(t, x, u) = E \int_t^T q(s, X^u(s)) \, ds + E\phi(X_T^u).
\]

We formulate the optimal control problem in the weak sense following the approach of [32]. By an admissible control system we mean \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P, W, u, X^u)\), where \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)\) is a filtered probability space satisfying the usual conditions, \(W\) is a cylindrical \(P\)-Wiener process with values in \(\Xi\), with respect to the filtration \((\mathcal{F}_t)_{t \geq 0}\). \(u\) is an \(\mathcal{F}_t\)-predictable process with values in \(U\), \(X^u\) is a mild solution of (6.1). An admissible control system will be briefly denoted by \((W, u, X^u)\) in the following.

Our purpose is to minimize the function \(J\) over all admissible control systems.

We define in a classical way the Hamiltonian function relative to the above problem: for all \(t \in [0, T], x \in \mathcal{C}, z \in \Xi\),

\[
\psi(t, x, z) = \inf \{ q(t, x, u) + zR(t, x, u) : u \in U \}
\]

and the corresponding, possibly empty, set of minimizers

\[
\Gamma(t, x, z) = \{ u \in U, \ q(t, x, u) + zR(t, x, u) = \psi(t, x, z) \}.
\]

We remark that by the Filippov theorem, see e.g. [35, Theorem 8.2.10, p. 316], if \(\Gamma\) takes non-empty values there exists a Borel measurable map \(\Gamma_0 : [0, T] \times \mathcal{C} \times \Xi \to U\) such that \(\Gamma_0(t, x, z) \in \Gamma(t, x, z)\) for \(t \in [0, T], x \in \mathcal{C}, z \in \Xi\).

We are now ready to formulate the assumptions we need.

**Hypothesis 6.1.** (i) \(A, F,\text{ and } G\) verify Hypothesis 3.1.

(ii) \((U, \mathcal{U})\) is a measurable space, \(q : [0, T] \times \mathcal{C} \times U \to R\) is measurable, \(R : [0, T] \times \mathcal{C} \times U \to \Xi\) is measurable and bounded.

(iii) \(\phi\) satisfies Hypothesis 4.1(iii).
(iv) The Hamiltonian $\psi$ defined in (6.3) satisfies the requirements of point (i) and point (ii) of Hypothesis 4.1.

We are in a position to prove the main result of this section:

**Theorem 6.2.** We assume Hypothesis 6.1 holds true and assume that the set-valued map $\Gamma$ has non-empty values and it admits a measurable selection $\Gamma_0 : [0, T] \times C \times S \rightarrow U$. Let $\nu$ denote the function in the statement of Corollary 4.4. Then for all admissible control system we have $J(t, x, u) \geq \nu(t, x)$ and the equality holds if and only if

$$u(s) = \Gamma_0(s, X_s^u, \overline{\nu}(s, X_s)), \quad P\text{-a.s. for almost every } s \in [t, T].$$

Moreover, the closed loop equation

$$\begin{cases}
    dX(s) = AX(s) \, ds + F(s, X_s) \, ds \\
    + G(s, X_s)(R(s, X_s, \Gamma_0(s, X_s, \overline{\nu}(s, X_s))G(s, X_s)) \, ds + dW(s), & s \in [t, T].
\end{cases}$$

admits a weak solution $(\Omega, F, \{F_t\}_{t \geq 0}, P, W, X)$ which is unique in law and setting

$$u(s) = \Gamma_0(s, X_s, \overline{\nu}(s, X_s)G(s, X_s)),$$

we obtain an optimal admissible control system $(W, u, X)$.

**Proof.** For all admissible control system $(\Omega, F, \{F_t\}_{t \geq 0}, P, W, u, X^u)$, let us define

$$W^u(s) = W(s) + \int_{t \land s}^s R(\sigma, X^u_{\sigma}, u(\sigma)) \, d\sigma, \quad s \in [0, T],$$

and

$$\rho = \exp\left(\int_t^T -R^*(s, X^u_s, u(s)) \, dW(s) - \frac{1}{2} \int_t^T |R(s, X^u_s, u(s))|^2 \, ds\right).$$

the Novikov condition implies that $E \rho = 1$. Setting $dP^u = \rho dP$, by the Girsanov theorem $W^u$ is a Wiener process under $P^u$. Let us denote by $\{F^u_t\}_{t \geq 0}$ the filtration generated by $W^u$ and completed in the usual way. Relatively to $W^u$ Eq. (6.1) can be rewritten as

$$\begin{cases}
    dX^u(s) = AX^u(s) \, ds + F(t, X^u_s) \, ds + G(s, X^u_s) \, dW(s), & s \in [t, T], \\
    X^u_t = x.
\end{cases}$$

In the space $(\Omega, F, \{F^u_t\}_{t \geq 0}, P^u)$, we consider the system of forward–backward equations

$$\begin{cases}
    dX^u(s) = AX^u(s) \, ds + F(t, X^u_s) \, ds + G(s, X^u_s) \, dW^u(s), & s \in [t, T], \\
    X^u_t = x, \\
    dY^u(s) = -\psi(s, X^u_s, Z^u(s)) \, ds + Z^u(s) \, dW^u(s), \\
    Y^u(T) = \phi(X^u_T).
\end{cases}$$

(6.5)
Writing the backward equation in (6.5) with respect to the process $W$ we get

$$
Y_t^u(s) + \int_s^T Z_t^u(\sigma) \, dW(\sigma) = \phi(X_t^u) + \int_s^T \left( \psi(\sigma, X_t^u, Z_t^u(\sigma)) - Z_t^u(\sigma) R(\sigma, X_t^u, u(\sigma)) \right) \, d\sigma. \tag{6.6}
$$

Recalling that $R$ is bounded we get, for some constant $C$,

$$
E^u[\rho^{-2}] = E^u\left[ \exp 2 \int_t^T R(s, X_t^u, u(s)) \, dW^u(s) - \frac{1}{2} \int_t^T |R(s, X_t^u, u(s))|^2 \, ds \right]
$$

$$
= E^u\left[ \exp \left( \int_t^T 2R(s, X_t^u, u(s)) \, dW^u(s) - \frac{1}{2} \int_t^T 4|R(s, X_t^u, u(s))|^2 \, ds \right) \times \exp \int_t^T |R(s, X_t^u, u(s))|^2 \, ds \right]
$$

$$
\leq C E^u \exp \left( \int_t^T 2R(s, X_t^u, u(s)) \, dW^u(s) - \frac{1}{2} \int_t^T 4|R(s, X_t^u, u(s))|^2 \, ds \right)
$$

$$
= C.
$$

It follows that

$$
E\left( \int_t^T |Z_t^u(s)|^2 \, ds \right)^{\frac{1}{2}} = E^u\left[ \left( \int_t^T |Z_t^u(s)|^2 \, ds \right)^{\frac{1}{2}} \rho^{-1} \right] \leq E^u \left( \int_t^T |Z_t^u(s)|^2 \, ds \right)^{\frac{1}{2}} \left( E^u \rho^{-2} \right)^{\frac{1}{2}} < \infty.
$$

We conclude that the stochastic integral in (6.6) has zero expectation. If we set $s = t$ in (6.6) and we take expectation with respect to $P$, we obtain

$$
EY_t^u(t) = \phi(X_t^u) + E \int_t^T \left[ \psi(\sigma, X_t^u, Z_t^u(\sigma)) - Z_t^u(\sigma) R(\sigma, X_t^u, u(\sigma)) \right] \, d\sigma.
$$

Since $Y_t^u(t, x) = \nu(t, x)$ and $Z_t^u(s, t, x) = \tilde{\nu}_0 v(s, x) G(s, x)$ P-a.s. for a.a. $s \in [t, T]$ we have that

$$
J(t, x, u) = \nu(t, x) + E \int_t^T \left[ -\psi(s, X_s^u, \tilde{\nu}_0 v(s, x) G(s, x)) \right]
$$

$$
+ \tilde{\nu}_0 v(s, x) G(s, x) R(s, X_s^u, u(s)) + q(s, X_s^u, u(s)) \right] \, ds.
$$

It implies that $J(t, x, u) \geq \nu(t, x)$ and that the equality holds if and only if (6.4) holds.
(Existence) Let $(\Omega, \mathcal{F}, P)$ be a given complete probability space with a filtration $\{\mathcal{F}_t\}_{t \geq 0}$ satisfying the usual conditions. $\{W(t), t \geq 0\}$ is a cylindrical Wiener process in $\mathcal{S}$ with respect to $\{\mathcal{F}_t\}_{t \geq 0}$. Let $X(\cdot)$ be the mild solution of

$$
\begin{aligned}
d &X(s) = A X(s) \, ds + F(s, X_s) \, ds + G(s, X_s) \, dW(s), \quad s \in [t, T], \\
X_t & = x,
\end{aligned}
$$

and by the Girsanov theorem, let $P^1$ be the probability on $\Omega$ under which

$$
W^1(s) = W(s) - \int_{t \wedge s}^s R(\sigma, X_\sigma, \Gamma_0(\sigma, X_\sigma, \nabla \varphi(s, X_\sigma) G(s, X_s))) \, d\sigma
$$

is a Wiener process (notice that $R$ is bounded). Then $X$ is the mild solution of the closed loop equation relatively to the probability $P^1$ and the Wiener process $W^1$.

(Uniqueness) Let $X$ be a weak solution of the closed loop equation in an $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P, W)$. We define

$$
\rho = \exp \left( \int_{t}^{T} -R^*(\sigma, X_\sigma, \Gamma_0(\sigma, X_\sigma, \nabla \varphi(s, X_\sigma) G(s, X_s))) \, dW(\sigma) \\
- \frac{1}{2} \int_{t}^{T} \left| R(\sigma, X_\sigma, \Gamma_0(\sigma, X_\sigma, \nabla \varphi(s, X_\sigma) G(s, X_s))) \right|^2 \, d\sigma \right),
$$

and

$$
W^0(s) = W(s) + \int_{t \wedge s}^s R(\sigma, X_\sigma, \Gamma_0(\sigma, X_\sigma, \nabla \varphi(s, X_\sigma) G(s, X_s))) \, d\sigma, \quad s \in [0, T].
$$

From the fact that $R$ is bounded, it follows that $E \rho = 1$. Setting $dP^0 = \rho \, dP$, by the Girsanov theorem $W^0$ is a Wiener process under probability measure $P^0$. Let us denote by $\{\mathcal{F}_t^0\}_{t \geq 0}$ the filtration generated by $W^0$ and completed in the usual way. In $(\Omega, \mathcal{F}, \{\mathcal{F}_t^0\}_{t \geq 0}, P^0, W)$, $X$ is a mild solution of

$$
\begin{aligned}
d &X(s) = A X(s) \, ds + F(t, X_s) \, ds + G(s, X_s) \, dW^0(s), \quad s \in [t, T], \\
X_t & = x
\end{aligned}
$$

and

$$
\rho = \exp \left( \int_{t}^{T} -R^*(\sigma, X_\sigma, \Gamma_0(\sigma, X_\sigma, \zeta(\sigma, X_\sigma))) \, dW^0(\sigma) \\
+ \frac{1}{2} \int_{t}^{T} \left| R(\sigma, X_\sigma, \Gamma_0(\sigma, X_\sigma, \zeta(\sigma, X_\sigma))) \right|^2 \, d\sigma \right).
$$

By Hypothesis 6.1, the law of $(X, W^0)$ is uniquely determined by $A, F, G$ and $x$. Taking into account the last displayed formula, we conclude that the law of $(X, W^0, \rho)$ under $P^0$ is also uniquely determined, and consequently so is the law of $X$ under $P$. The proof is finished. □
Example 6.2.1. As an application of our results, we consider, for \( t \in [0, T] \) and \( \xi \in [0, 1] \), the following equation:

\[
\begin{align*}
\frac{dz^u(t, \xi)}{dt} &= \Delta z^u(t, \xi) + f(t, \int_{-1}^{0} z^u(t+\theta, \xi) \, d\theta) + \sigma(t, \int_{-1}^{0} z^u(t+\theta, \xi) \, d\theta)
\end{align*}
\]

Moreover, we consider the cost functional

\[
J(t_0, x_0, u) = E \int_{0}^{T} \left( t, \int_{-1}^{0} z^u(t+\theta, \xi) \, d\theta, u(t) \right) d\xi \, dt + E \int_{0}^{1} \rho(\xi, \int_{-1}^{0} z^u(t+\theta, \xi) \, d\theta) d\xi.
\]

(6.8)

To satisfy our Hypothesis 6.1 we assume the following.

Hypothesis 6.3. (1) The initial condition \( x \in C([-1, 0], L^2(0, 1)) \).

(2) The mappings \( f, \sigma : [0, T] \times R \rightarrow R \) are measurable, and \( f, \sigma \in C^{0,1}([0, T] \times R; R) \) and satisfy, for some constant \( L > 0 \),

\[
|f(t, 0)| + |f'(t, x)| \leq L, \quad |\sigma(t, 0)| + |\sigma'(t, x)| \leq L, \quad t \in [0, T], \; x \in R.
\]

(3) We say \( u \) is an admissible control if it is a real valued predictable process taking values in a bounded closed subset \( U \) of \( R \).

(4) \( \rho : [0, 1] \times R \rightarrow R \) is continuous and bounded, and \( \rho(\xi, \cdot) \in C_1^1(R) \). Moreover, there exists a function \( c_1 \) continuous on \([0, 1]\) such that \( |\nabla_x \rho(\xi, x)| \leq c_1(\xi) \).

(5) For every \( t \in [0, T] \) and \( \xi \in [0, 1] \), \( l(t, \xi, \cdot, \cdot) : R \times U \rightarrow R \) is continuous, and there is a constant \( L \) such that, for every \( t \in [0, T] \), \( \xi \in [0, 1] \), \( y_1, y_2 \in R \), \( u \in U \),

\[
|l(t, \xi, y_1, u) - l(t, \xi, y_2, u)| \leq L(1 + |y_1| + |y_2|)|y_1 - y_2|.
\]

and for every \( t \in [0, T] \)

\[
\frac{1}{0} \sup_{u \in U} |l(t, \xi, 0, u)| \, d\xi \leq L.
\]

To rewrite the problem in an abstract way we set \( H = L^2([0, 1]) \), \( \mathcal{E} = R \), and let \( A \) denote the Laplace operator \( \Delta \) in \( H \) with domain \( W^{2,2}([0, 1]) \cap W^{1,2}_0([0, 1]) \).
We set
\[
F(t, x)(\xi) = f\left(t, \int_{-1}^{0} x(\theta, \xi) \, d\theta \right), \quad \left(G(t, x)u(t)\right)(\xi) = \sigma\left(t, \int_{-1}^{0} x(\theta, \xi) \, d\theta \right) r(\xi)u(t),
\]
\[
q(t, x, u) = \int_{0}^{1} \left( t, \int_{-1}^{0} x(\theta, \xi) \, d\theta, u(t) \right) d\xi, \quad \phi(x) = \int_{0}^{1} \rho \left( \frac{\xi}{t}, \int_{-1}^{0} x(\theta, \xi) \, d\theta \right) d\xi,
\]
for \( t \in [0, T], x \in C([-1, 0]; \mathbb{R}), u \in U. \)

Under the previous assumptions we know that the assumptions in Hypothesis 6.1(i), (ii), (iii) are satisfied.

We define the Hamiltonian:
\[
\psi(t, x, z) = \inf_{u \in U} \left\{ zu + q(t, x, u) \right\}, \quad t \in [0, T], \; x \in C([-1, 0]; \mathbb{R}), \; z \in \mathbb{R}.
\]

We note that, for all \( t \in [0, T], x \in C([-1, 0]; \mathbb{R}), q(t, x, \cdot) \) is continuous on the compact set \( U \), so the above infimum is attained. Therefore if we define
\[
\Gamma(t, x, z) = \left\{ u \in U : zu + q(t, x, u) = \psi(s, x, z) \right\},
\]
then \( \Gamma(s, x, z) \neq \emptyset \) for all \( t \in [0, T], x \in C([-1, 0]; \mathbb{R}) \) and \( z \in \mathbb{R} \). By the Filippov theorem, see e.g. [35], \( \Gamma \) admits a measurable selection, i.e. there exists a measurable function \( \Gamma_{0} : [0, T] \times C([-1, 0]; \mathbb{R}) \times \mathbb{R} \to U \) with \( \Gamma_{0}(t, x, z) \in \Gamma(t, x, z) \) for all \( t \in [0, T], x \in C([-1, 0]; \mathbb{R}) \) and \( z \in \mathbb{R} \).

Assume that for every \( t \in [0, T] \) the map \( \psi(t, \cdot, \cdot) \) is differentiable with bounded derivatives, then the Hamiltonian defined in (6.10) satisfies Hypothesis 6.1(iv). Therefore, the closed loop equation admits a weak solution \( (\Omega, \mathcal{F}, \{\mathcal{F}_{t}\}_{t \geq 0}, P, W, u, z(\cdot)) \) and setting
\[
u(s) = \Gamma_{0}(s, z_{s}(\cdot), \nabla_{0} \psi(s, z_{s}(\cdot))G(s, z_{s}(\cdot))),
\]
we obtain an optimal admissible control system \((W, u, z(\cdot))\).

References