Finite-dimensional representations for controlled diffusions with delay

Salvatore Federico* Peter Tankov†

Abstract

We study stochastic delay differential equations (SDDE) where the coefficients depend on the moving averages of the state process. As a first contribution, we provide sufficient conditions under which the solution of the SDDE and a linear path functional of it admit a finite-dimensional Markovian representation. As a second contribution, we show how approximate finite-dimensional Markovian representations may be constructed when these conditions are not satisfied, and provide an estimate of the error corresponding to these approximations. These results are applied to optimal control and optimal stopping problems for stochastic systems with delay.

Key words: Stochastic delay differential equation (SDDE), Markovian representation, Laguerre polynomials, stochastic control, optimal stopping.

MSC2010: 60H10, 60G40, 93E20.

1 Introduction

In this paper we study a class of controlled stochastic differential equations with memory, where the coefficients of the equation depend on the moving average of the past values of the solution process (so called distributed delay):

$$dS_t = b\left(S_t, \int_{\mathbb{R}^-} \tilde{\alpha}_1(\xi) S_{t+\xi} d\xi, u_t\right) dt + \sigma\left(S_t, \int_{\mathbb{R}^-} \tilde{\beta}_1(\xi) S_{t+\xi} d\xi, u_t\right) dW_t, \tag{1}$$

where b, σ, $\tilde{\alpha}_1, \tilde{\beta}_1$ are given functions and $u = (u_t)_{t\geq 0}$ is a control process. Equations of this type appear in a variety of domains such as economics [23, 24] and finance [1, 4, 5, 19, 25], as well as in physical sciences [29]. In general this equation is infinite-dimensional, which means that it can be formulated as evolution equation in an infinite-dimensional space of the form $\mathbb{R} \times H_1$, where H_1 is a Hilbert space, for the process $\mathbf{X}_t = (S_t, (S_{t+\xi})_{\xi\leq 0})$, but cannot be represented via a finite-dimensional controlled Markov process. This makes solving stochastic control and optimal stopping problems associated to such systems notoriously difficult.

For this reason we are interested in finding exact - when possible - or approximate finite dimensional Markovian representations for S_t and also for linear path functionals of the form $Z_t = \langle \gamma, \mathbf{X}_t \rangle$, where γ is fixed. Indeed, the latter functional may represent for example the reward process of a control problem and make the problem non-Markovian even when the state process S is Markovian. We say that the process (S, Z) admits a finite-dimensional Markovian representation if there exists a finite-dimensional subspace V of the space $\mathbb{R} \times H_1$

*Università di Milano. E-mail: salvatore.federico@unimi.it
†Université Paris Diderot. E-mail: tankov@math.univ-paris-diderot.fr

such that: 1) V contains the vector $(1,0) \in \mathbb{R} \times H_1$; 2) the projection of $\mathbf{X}_t$ on this subspace, call it $\mathbf{X}_t^V$, satisfies a finite dimensional stochastic differential equation; 3) the processes S and Z can be written respectively as $S_t = \langle (1,0), \mathbf{X}_t^V \rangle$ and $Z_t = \langle \gamma, \mathbf{X}_t^V \rangle$. On the other hand, to find an approximate finite dimensional Markovian representation for (S, Z), we need to find a sequence of processes $(\mathbf{X}_t^n)$ and a sequence of subspaces (V_n), such that for every n, the projection $\mathbf{X}^{n,V_n}$ satisfies a finite-dimensional (controlled Markovian) SDE, and such that for a sequence (γ_n) to be determined, $S_t^n = \langle (1,0), \mathbf{X}_t^{n,V} \rangle$ and $Z_t^n = \langle \gamma_n, \mathbf{X}_t^{n,V} \rangle$ converge respectively to S_t and Z_t as $n \longrightarrow \infty$.

Our approach is different from most existing studies of invariance for stochastic equations on Hilbert spaces (see e.g. [15, 16]), which require that the entire solution stays on a finite-dimensional submanifold of the original space. Instead, we require that a projection of the solution or an approximation thereof evolves on a finite-dimensional space. This projection only contains partial information about the solution, but if the reward function of the control problem only depends on this projection, this information is sufficient to solve the control problem.

Optimal control problems for stochastic systems with memory have been considered by many authors starting with [26]. Solving the problem in the infinite-dimensional setting being very difficult, some recent contributions focus on special cases where the problem reduces to a finite-dimensional one [2, 9, 28, 30]. In the general case, [27] extends the Markov chain approximation method to stochastic equations with delay. A similar method is developed in [32], and [17] establish convergence rates for an approximation of this kind. The infinite-dimensional Hilbertian approach to controlled deterministic and stochastic systems with delays in the state variable was employed in some papers. For the deterministic case we can quote [12, 13], which perform a study of the Hamilton-Jacobi-Bellman (HJB) equation in infinite dimension[1]; for the stochastic case we can quote [11, 23, 24] with some partial results on the solution of the control problem (in [23, 24] the delay is considered also in the control variable, but the diffusion term is just additive). We should also mention the Banach space approach employed by [20]: the problem is embedded in the space of continuous functions and the HJB equation is approached using the concept of mild solutions. Optimal stopping problems for stochastic systems with delay can be treated with methods similar to those used for optimal control. [14] and [21] discuss special cases where the infinite-dimensional problem reduces to a finite-dimensional one. In the specific context of American options written on the moving average of the asset price, [4] propose a method based on Laguerre polynomial approximation, which is extended and refined in the present paper.

Let us now briefly summarize the contents of the paper. In section 2 we define the stochastic delay differential equation, state the assumptions on the coefficients and introduce the main notation. In Section 3 we introduce and study an alternative representation for this equation, as an evolution equation in an infinite-dimensional Hilbert space. Section 4 contains the main results of the paper. First, we provide sufficient conditions for existence of an exact finite dimensional Markovian representation for the output process, namely that the coefficients belong to a certain exponential-polynomial family (sum of exponential functions multiplied by polynomials). Second, we describe a method for constructing an approximate finite-dimensional representation using a specific exponential-polynomial family based on Laguerre polynomials. The error of the approximation is also analyzed here (Proposition 4.8). Finally, Section 5 briefly discusses the applications of our method to the solution of optimal control and optimal stopping problems for stochastic systems with delay. Detailed analysis of these applications and numerical examples is left for further research.

[1] When the delay appears also in the control variable the infinite-dimensional representation is more involved. We refer to [3, Part II, Ch. 4], where a general theory is developed based on the paper [34].

2 The controlled stochastic delay differential equation

Let $(\Omega, \mathbb{F}, \mathbb{P})$ be a complete probability space endowed with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions. We assume defined on this filtered probability space a Brownian motion $W = (W_t)_{t\geq 0}$.

Let $\mathbb{R}^- := (-\infty, 0]$. To distinguish, deterministic functions of time will be denoted with the time index in parentheses, while time-continuous stochastic processes will be denoted with the time index as subscript.

Let $S = (S_t)_{t\geq 0}$ be a controlled diffusion on this space solving a stochastic delay differential equation (SDDE)

$$dS_t \;=\; b\left(S_t, \int_{\mathbb{R}^-} \tilde{\alpha}_1(\xi) S_{t+\xi} d\xi, u_t\right) dt + \sigma\left(S_t, \int_{\mathbb{R}^-} \tilde{\beta}_1(\xi) S_{t+\xi} d\xi, u_t\right) dW_t, \tag{2}$$

where $b, \sigma, \tilde{\alpha}_1, \tilde{\beta}_1$ are given functions and $u = (u_t)_{t\geq 0}$ is an adapted control process. Due to the dependence on the past, S is not a controlled Markov diffusion. Moreover, in order to define the process S, one needs to specify an initial condition not only at $t = 0$, but also for all $t < 0$. In other terms, (2) has to be completed (in general) with an initial condition of the form

$$S_0 \;=\; s_0 \;\in\; \mathbb{R}; \qquad S_\xi \;=\; s_1(\xi), \quad \xi < 0, \tag{3}$$

where s_1 is a given function. So, *the initial datum is a function.* From what we have said it is clear that, even if the process S is one-dimensional, it may not in general be represented as a finite-dimensional controlled Markov diffusion.[2] Moreover, even when the process S is a controlled Markov diffusion, i.e. $\alpha_1 \equiv 0, \beta_1 \equiv 0$ in (2), the control problem may still not be a finite-dimensional Markovian one, as one may need to consider also the process[3]

$$(Z_t)_{t\geq 0} \;=\; \left(\int_{\mathbb{R}^-} \tilde{\gamma}_1(\xi) S_{t+\xi} d\xi\right)_{t\geq 0}, \tag{4}$$

where $\tilde{\gamma}_1$ is a function. Also in this case it is clear that in general the Markovian representation of the system must be infinite-dimensional. In this paper we deal with the problem of rewriting the above system in an exact or approximate way in terms of a Markov controlled finite-dimensional diffusion when at least one among the functions $\tilde{\alpha}_1, \tilde{\beta}_1, \tilde{\gamma}_1$ is not identically equal to 0.

Remark 2.1. We stress that, altough we take one-dimensional processes W, S, Z, the argument can be easily generalized to the case of multi-dimensional processes. Also we have taken an autonomous equation for S, i.e. there is no explicit time dependence in the coefficients b, σ; this is done just for notational convenience: all computations can be performed also in the non-autonomous case.

In the sequel, we are going to reformulate equation (2) as an evolution equation in a Hilbert space. To allow a set of initial data possibly containing the constant functions, we work with weighted spaces. We consider on $\mathbb{R}^-$ a weight function w and make the following standing assumption.

Assumption 2.2. *$w \in C^1(\mathbb{R}^-; \mathbb{R})$, $w > 0$, w'/w is bounded.*

Moreover, without any loss of generality, we also suppose that $w(0) = 1$. Denote

$$L^2_w \;:=\; L^2(\mathbb{R}^-, w(\xi)d\xi\,; \mathbb{R}). \tag{5}$$

[2] Nevertheless there are examples where a finite-dimensional Markovian representation can be obtained. We will study this kind of situation in Section 4.2, giving sufficient conditions for a finite-dimensional Markovian representation.

[3] For example this process could appear in the cost functional of a control problem.

When $w \equiv 1$ we simply denote the space above by L^2. Throughout the paper, we shall work under the following assumptions on the model, guaranteeing existence and uniqueness for the solution to (2) and good properties for the problem we aim to study.

Assumption 2.3.

1. *The control process u takes values in a Borel set $U \subset \mathbb{R}^d$;*
2. *$u \in \mathcal{U}$, where*
$$\mathcal{U} \;=\; \{(u_t)_{t\geq 0} \text{ adapted process belonging to } L^2_{loc}(\mathbb{R}^+; L^2(\Omega; U))\};$$
3. *$b, \sigma : \mathbb{R}^2 \times U \to \mathbb{R}$ are such that there exist constants $C_1, C_2 \geq 0$ with*
$$\begin{aligned} |b(x,y,u) - b(x',y',u)| \;+\; |\sigma(x,y,u) - \sigma(x',y',u)| & \\ \leq\; C_1\,(|x-x'| \;+\; |y-y'|), & \quad \forall x,x',y,y' \in \mathbb{R},\ \forall u \in U; \end{aligned}$$
$$|b(x,y,u)| \;+\; |\sigma(x,y,u)| \;\leq\; C_2\,(1+|x|+|y|), \qquad \forall x,y, \in \mathbb{R},\ \forall u \in U;$$
4. *There exists w satisfying Assumption 2.2 such that the functions $\tilde{\alpha}_1 w^{-1/2}, \tilde{\beta}_1 w^{-1/2}, \tilde{\gamma}_1 w^{-1/2}$ belong to L^2.*

Remark 2.4. Typical weights are the exponential ones: $w(\xi) = e^{\lambda\xi}$, $\lambda \in \mathbb{R}$. However, in some cases it may be necessary to use other weight functions. For example, let $\tilde{\gamma}_1(\xi) = \frac{1}{1+|\xi|^p}$ with $p > 2$. Then, taking $w(\xi) = \frac{1}{1+|\xi|^{\frac{p}{2}}}$, we ensure that simultaneously $\tilde{\gamma}_1 w^{-1/2} \in L^2$ and the constant functions belong to L^2_w. These two properties cannot hold simultaneously with an exponential weight function.

Define
$$\alpha_1 \;=\; \tilde{\alpha}_1 w^{-1} \qquad \beta_1 \;=\; \tilde{\beta}_1 w^{-1}, \qquad \gamma_1 \;=\; \tilde{\gamma}_1 w^{-1}.$$
Then, due to Assumption 2.3(4), we have $\alpha_1, \beta_1, \gamma_1 \in L^2_w$. Moreover, (2)-(3) can be rewritten as

$$\begin{cases} dS_t = b\left(S_t, \int_{\mathbb{R}^-} \alpha_1(\xi) S_{t+\xi} w(\xi) d\xi, u_t\right) dt + \sigma\left(S_t, \int_{\mathbb{R}^-} \beta_1(\xi) S_{t+\xi} w(\xi) d\xi, u_t\right) dW_t, \\ \\ S_0 = s_0, \qquad S_\xi = s_1(\xi), \quad \xi < 0, \end{cases} \tag{6}$$

and (4) can be rewritten as

$$Z \;=\; (Z_t)_{t\geq 0} \;=\; \left(\int_{\mathbb{R}^-} \gamma_1(\xi) S_{t+\xi} w(\xi) d\xi\right)_{t\geq 0}. \tag{7}$$

Proposition 2.5. *For every $\mathbf{s} := (s_0, s_1(\cdot)) \in \mathbb{R} \times L^2_w$ and $u \in \mathcal{U}$, (6) admits a unique up to indistinguishabiliy strong solution $S^{\mathbf{s},u}$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and this solution admits a version with continuous paths.*

Proof. This result is an easy corollary of Theorem IX.2.1 in [33] (note that without making any changes to the proof, this theorem can be extended to the case when the coefficients depend on a random adapted control). Let $T > 0$ and define the maps $\tilde{B}, \tilde{\Sigma} : [0,T] \times C([0,T]; \mathbb{R}) \times \Omega \;\longrightarrow \mathbb{R}$

$$\tilde{B}(t, z(\cdot), \omega) \;:=\; b\Big(z(t), \int_{-\infty}^{-t} \alpha_1(\xi) s_1(t+\xi) w(\xi) d\xi + \int_{-t}^{0} z(t+\xi) \alpha_1(\xi) w(\xi) d\xi, u_t(\omega)\Big),$$

$$\tilde{\Sigma}(t, z(\cdot), \omega) \;:=\; \sigma\Big(z(t), \int_{-\infty}^{-t} \beta_1(\xi) s_1(t+\xi) w(\xi) d\xi + \int_{-t}^{0} z(t+\xi) \beta_1(\xi) w(\xi) d\xi, u_t(\omega)\Big).$$

By localizing in time, to use the aforementioned result we need to check that:

1. $t \mapsto \tilde{B}(t, z(\cdot), \omega)$, $t \mapsto \tilde{\Sigma}(t, z(\cdot), \omega)$ are bounded for every constant function $z(\cdot) \equiv z_0$ uniformly in $\omega \in \Omega$;
2. $\tilde{B}(t, \cdot, \omega), \tilde{\Sigma}(t, \cdot, \omega)$ are Lipschitz continuous, with respect to the uniform norm on $C([0,T];\mathbb{R})$, uniformly in $t \in [0,T]$, $\omega \in \Omega$.

Let us focus on $\tilde{B}$, the proof for $\tilde{\Sigma}$ being the same. We first check the local boundedness on constant functions. Let $z(\cdot) \equiv z_0$. By the linear growth assumption,

$$\begin{aligned}|\tilde{B}(t, z(\cdot), \omega)| &= |b(z_0, \int_{-\infty}^{-t} \alpha_1(\xi) s_1(t+\xi) w(\xi) d\xi + z_0 \int_{-t}^{0} \alpha_1(\xi) w(\xi) d\xi, u_t(\omega))| \\ &\leq C_2 \left(1 + |z_0| + \left| \int_{-\infty}^{-t} \alpha_1(\xi) s_1(t+\xi) w(\xi) d\xi \right| + |z_0| \left| \int_{-t}^{0} \alpha_1(\xi) w(\xi) d\xi \right| \right).\end{aligned}$$

Let us denote the upper bound of $|w'|/w$ by K. Then, by Gronwall's inequality, for all $a, b \leq 0$,

$$w(a) \ \leq \ w(b) e^{K|b-a|}. \tag{8}$$

The term involving s_1 then satisfies

$$\begin{aligned}\left| \int_{-\infty}^{-t} \alpha_1(\xi) s_1(t+\xi) w(\xi) d\xi \right| &\leq \left(\int_{-\infty}^{-t} \alpha_1^2(\xi) w(\xi) d\xi \right)^{\frac{1}{2}} \left(\int_{-\infty}^{-t} s_1^2(t+\xi) w(\xi) d\xi \right)^{\frac{1}{2}} \\ &\leq \left(\int_{-\infty}^{0} \alpha_1^2(\xi) w(\xi) d\xi \right)^{\frac{1}{2}} \left(\int_{-\infty}^{0} s_1^2(\xi) w(\xi - t) d\xi \right)^{\frac{1}{2}} \\ &\leq e^{\frac{Kt}{2}} \left(\int_{-\infty}^{0} \alpha_1^2(\xi) w(\xi) d\xi \right)^{\frac{1}{2}} \left(\int_{-\infty}^{0} s_1^2(\xi) w(\xi) d\xi \right)^{\frac{1}{2}},\end{aligned}$$

which is bounded in $[0,T]$ since $\alpha_1, s_1 \in L^2_w$. Similarly,

$$\left| \int_{-t}^{0} \alpha_1(\xi) w(\xi) d\xi \right| \ \leq \ \int_{-t}^{0} |\alpha_1(\xi)| w(\xi) d\xi \ \leq \ \left(\int_{-t}^{0} \alpha_1^2(\xi) w(\xi) d\xi \right)^{\frac{1}{2}} \left(\int_{-t}^{0} w(\xi) d\xi \right)^{\frac{1}{2}}, \tag{9}$$

which is bounded in $[0,T]$ as well since $\alpha_1 \in L^2_w$ and $w \in C^1(\mathbb{R}^-;\mathbb{R})$.

Let us now check the Lipschitz condition. By the Lipschitz property of b,

$$\begin{aligned}|\tilde{B}(t, z(\cdot), \omega) - \tilde{B}(t, z'(\cdot), \omega)| &\leq C_1 \Big(|z(t) - z'(t)| + \int_{-t}^{0} |z(t+\xi) - z'(t+\xi)| |\alpha_1(\xi)| w(\xi) d\xi \Big) \\ &\leq C_1 \Big(1 + \int_{-T}^{0} |\alpha_1(\xi)| w(\xi) d\xi \Big) \cdot \max_{0 \leq u \leq T} |z(u) - z'(u)|,\end{aligned}$$

and the functional Lipschitz property follows from inequality (9). □

3 Product space infinite-dimensional representation

In this section we provide an infinite-dimensional representation of SDDE (6) in the product Hilbert space

$$H_w \ := \ \mathbb{R} \times L^2_w.$$

When $w \equiv 1$ we simply denote the space above by H. We denote by $\mathbf{x} = (x_0, x_1)$ the generic element of H_w, noting that the second component is a function. The norm and the inner product of H_w, defined in the usual way from the norm and the inner products of the two components, will be denoted, respectively, by $\|\cdot\|_w$, $\langle \cdot, \cdot \rangle_w$. Also, when $w \equiv 1$ we simply denote the norm and the inner product above by $\|\cdot\|$, $\langle \cdot, \cdot \rangle$.

3.1 Preliminaries

Let us introduce the weighted Sobolev spaces on $\mathbb{R}^-$ as follows (we refer to [6, Ch. VIII] for an introduction to Sobolev spaces under the Lebesgue measure on intervals). Given $f \in L^2_{loc}(\mathbb{R}^-;\mathbb{R})$, we say that f admits weak derivative $g \in L^2_{loc}(\mathbb{R}^-;\mathbb{R})$ if

$$\int_{\mathbb{R}} f(\xi)\varphi'(\xi)d\xi \;=\; -\int_{\mathbb{R}} g(\xi)\varphi(\xi)d\xi, \quad \forall \varphi \in C^1_c((-\infty,0);\mathbb{R}).$$

It is well known that, if such a function g exists, it is unique. Moreover it coincides with the classical derivative f' when $f \in C^1(\mathbb{R}^-;\mathbb{R})$. By extension, the function g is denoted by f' in general, i.e., also when $f \notin C^1(\mathbb{R}^-;\mathbb{R})$. We denote the space of functions of $L^2_{loc}(\mathbb{R}^-;\mathbb{R})$ admitting weak derivative in $L^2_{loc}(\mathbb{R}^-;\mathbb{R})$ by $W^{1,2}_{loc}(\mathbb{R}^-;\mathbb{R})$. [6, Th. VIII.2] states that for every $f \in W^{1,2}_{loc}(\mathbb{R}^-;\mathbb{R})$ there exists a locally absolutely continuous version of f on $\mathbb{R}^-$, so that it holds

$$f(\xi) - f(\xi_0) \;=\; \int_{\xi_0}^{\xi} f'(r)dr, \quad \forall\, \xi_0 \le \xi \le 0. \tag{10}$$

Given $f \in W^{1,2}_{loc}(\mathbb{R}^-;\mathbb{R})$, we shall always refer to its absolutely continuous version. By [6, Cor. VIII.10], if $f, g \in W^{1,2}_{loc}(\mathbb{R}^-;\mathbb{R})$, then $fg \in W^{1,2}_{loc}(\mathbb{R}^-;\mathbb{R})$ and

$$(fg)' \;=\; f'g + fg', \tag{11}$$

so the integration by parts formula

$$\int_a^b f'(\xi)g(\xi)d\xi \;=\; f(b)g(b) - f(a)g(a) - \int_a^b f(\xi)g'(\xi)d\xi, \quad \forall\, a \le b \le 0,$$

holds true for all $f, g \in W^{1,2}_{loc}(\mathbb{R}^-;\mathbb{R})$. On the elements of the space $W^{1,2}_{loc}(\mathbb{R}^-;\mathbb{R})$ we define the norm

$$\|f\|_{W^{1,2}_w} \;:=\; \int_{\mathbb{R}^-} (|f(\xi)|^2 + |f'(\xi)|^2) w(\xi) d\xi,$$

and, moreover, we define the space

$$W^{1,2}_w \;:=\; \left\{ f \in W^{1,2}_{loc} \mid \|f\|_{W^{1,2}_w} < \infty \right\}.$$

Clearly $W^{1,2}_w \subset L^2_w$. The linear maps

$$\begin{array}{ccc} (W^{1,2}_w, \|\cdot\|_{W^{1,2}_w}) & \longrightarrow & (L^2_w, \|\cdot\|_{L^2_w}) \times (L^2_w, \|\cdot\|_{L^2_w}), \\ f & \longmapsto & (f, f'), \end{array}$$

and

$$\begin{array}{ccc} (L^2_w, \|\cdot\|_{L^2_w}) & \longrightarrow & (L^2, \|\cdot\|_{L^2}), \\ f & \longmapsto & f w^{\frac{1}{2}}, \end{array}$$

are isometries, so, since L^2 is a separable Banach space, we deduce that $W^{1,2}_w$ is a separable Hilbert space when endowed with the inner product

$$\langle f, g\rangle_{W^{1,2}_w} \;:=\; \int_{\mathbb{R}^-} (f(\xi)g(\xi) + f'(\xi)g'(\xi)) w(\xi) d\xi.$$

By the assumption that w'/w bounded, denoting the upper bound of $|w'|/w$ by K, we see that if $f \in W^{1,2}_w$, then

$$\begin{aligned} \|fw^{1/2}\|^2_{W^{1,2}} &= \int_{\mathbb{R}^-} \left\{ f^2(\xi)w(\xi) + \left(f'(\xi)\sqrt{w(\xi)} + f(\xi)\frac{w'(\xi)}{2\sqrt{w(\xi)}} \right)^2 \right\} d\xi \\ &\le \int_{\mathbb{R}^-} \left(f^2(\xi) + 2f'(\xi)^2 + \frac{K^2}{2} f^2(\xi) \right) w(\xi) d\xi \le \left(2 + \frac{K^2}{2} \right) \|f\|^2_{W^{1,2}_w}. \end{aligned} \tag{12}$$

Thus, if $f \in W^{1,2}_w$, then $fw^{1/2} \in W^{1,2}$. Hence, Corollary [6, Cor. VIII.8] applied to $fw^{1/2}$ yields

$$\lim_{\xi \to -\infty} f(\xi) w^{\frac{1}{2}}(\xi) = 0, \quad \forall f \in W^{1,2}_w.$$

Recalling again our assumptions on w, we see that the following weighted integration by parts formula holds for all $f, g \in W^{1,2}_w$:

$$\int_{\mathbb{R}^-} f'(\xi) g(\xi) w(\xi) d\xi = f(0)g(0) - \int_{\mathbb{R}^-} f(\xi) \left(g'(\xi) + g(\xi) \frac{w'(\xi)}{w(\xi)} \right) w(\xi) d\xi. \tag{13}$$

Now, consider on the space H_w the family of linear bounded operators $(T(t))_{t\geq 0}$ acting as follows:

$$T(t)\mathbf{x} = ([T(t)\mathbf{x}]_0, [T(t)\mathbf{x}]_1) = \left(x_0,\ x_0 \mathbf{1}_{(0,t]}(t+\cdot) + x_1(t+\cdot)\mathbf{1}_{\mathbb{R}^-}(t+\cdot)\right). \tag{14}$$

Simple computations show that

$$\|T(t)\|_{\mathcal{L}(H_w)} \leq 1 + t, \qquad \forall t \geq 0. \tag{15}$$

We are going to study the semigroup properties of $(T(t))_{t\geq 0}$. For basic facts about the theory of semigroups we refer to the classical monographs [8, 10].

Proposition 3.1. *The family of linear operators $(T(t))_{t\geq 0}$ defined in* (14) *is a strongly continuous semigroup on the space H_w, generated by the closed unbounded operator $\mathcal{A}$ defined on*

$$D(\mathcal{A}) = \{\mathbf{x} = (x_0, x_1) \in H_w \mid x_1 \in W^{1,2}_w,\ x_0 = x_1(0)\} \tag{16}$$

by

$$\mathcal{A}\mathbf{x} = (0,\ x_1'). \tag{17}$$

Proof. The fact that $(T(t))_{t\geq 0}$ is a semigroup is immediate by the definition. The fact that it is strongly continuous follows by the continuity of translations in L^2_w, which can be proved, e.g., starting from the continuity of translation in L^2 and exploiting (8).

Now let us show that $(T(t))_{t\geq 0}$ is generated by $\mathcal{A}$. Set

$$\mathcal{D} := \{\mathbf{x} = (x_0, x_1) \in H_w \mid x_1 \in W^{1,2}_w,\ x_0 = x_1(0)\}$$

and take $\mathbf{x} \in \mathcal{D}$. Since $x_1 \in W^{1,2}_w$, it is absolutely continuous. So, extending x_1 to $\mathbb{R}$ by setting $x_1(\xi) = x_1(0)$ for $\xi > 0$, we can write

$$x_1(t+\xi) - x_1(\xi) = t \int_0^1 x_1'(\xi + \lambda t) d\lambda, \quad \forall \xi \in \mathbb{R}^-,\ \forall t \geq 0.$$

Hence, taking into account that $x_1(0) = x_0$, we have

$$\begin{aligned} \left\| \frac{T(t)\mathbf{x} - \mathbf{x}}{t} - (0, x_1') \right\|_w^2 &= \int_{\mathbb{R}^-} w(\xi) d\xi \left| \int_0^1 (x_1'(\xi + \lambda t) - x_1'(\xi)) d\lambda \right|^2 \\ &\leq \int_0^1 d\lambda \int_{\mathbb{R}^-} |x_1'(\xi + \lambda t) - x_1'(\xi)|^2 w(\xi) d\xi \\ &= \int_0^1 d\lambda \int_{\mathbb{R}^-} \|T(\lambda t)(0, x_1') - (0, x_1')\|_{H_w}^2 d\xi. \end{aligned}$$

By (15) and from the inequality above, we can conclude by dominated convergence that $\mathcal{D} \subset D(\mathcal{A})$ and that $\mathcal{A}$ acts as stated in (17) on the elements of $\mathcal{D}$.

We need now to show that $\mathcal{D} = D(\mathcal{A})$. For that, we notice that $\mathcal{D}$ is clearly dense in H_w and that $T(t)\mathcal{D} \subset \mathcal{D}$ for any $t \geq 0$. Hence, by [10, Ch. II, Prop. 1.7, p. 53], $\mathcal{D}$ is a core for

$D(\mathcal{A})$ (i.e. is dense in $D(\mathcal{A})$ endowed with the graph norm $\|\cdot\|_{D(\mathcal{A})}$). Hence, it just remains to show that $\mathcal{D}$ is closed with respect to the graph norm to conclude $\mathcal{D} = D(\mathcal{A})$. So, take a sequence $(\mathbf{x}^n) = (x_0^n, x_1^n) \subset \mathcal{D}$ converging with respect to $\|\cdot\|_{D(\mathcal{A})}$ to some $\mathbf{x} = (x_0, x_1) \in H_w$. Then we have

$$x_0^n \to x_0 \quad \text{in } \mathbb{R}; \qquad x_1^n \to x_1 \quad \text{in } W_w^{1,2}. \tag{18}$$

We immediately deduce that $x_1 \in W_w^{1,2}$. By (12), the linear map

$$\mathcal{L}: \ W_w^{1,2} \ \to \ W^{1,2}, \quad f \ \mapsto \ f w^{1/2},$$

is continuous. Since we have (see, e.g. [6, Th. 8.8]) the Sobolev continuous embedding $\iota: \ W^{1,2} \hookrightarrow L^\infty(\mathbb{R}^-;\mathbb{R})$, the map $\iota \circ \mathcal{L}$ is continuous. Taking into account also that x_1 is absolutely continuous, we deduce from the second convergence in (18)

$$x_1^n(0) \to x_1(0) \quad \text{in } \mathbb{R}.$$

Since $x_0^n = x_1^n(0)$, we conclude $x_1(0) = x_0$, and the proof is complete. □

3.2 Infinite-dimensional representation

Define the elements of H_w

$$\mathbf{e}^0 \ := \ (1,0), \qquad \boldsymbol{\alpha} \ := \ (0,\alpha_1), \qquad \boldsymbol{\beta} \ := \ (0,\beta_1),$$

and the Lipschitz continuous nonlinear operators

$$B: \ H_w \times U \ \to \ H_w, \qquad B(\mathbf{x},u) \ := \ b(\langle \mathbf{e}^0, \mathbf{x}\rangle_w, \langle \boldsymbol{\alpha}, \mathbf{x}\rangle_w, u)\, \mathbf{e}^0;$$

$$\Sigma: \ H_w \times U \ \to \ H_w, \qquad \Sigma(\mathbf{x},u) \ := \ \sigma(\langle \mathbf{e}^0, \mathbf{x}\rangle_w, \langle \boldsymbol{\beta}, \mathbf{x}\rangle_w, u)\, \mathbf{e}^0.$$

Given $\mathbf{x} \in H_w$ and $u \in \mathcal{U}$, we consider the following stochastic evolution equation in the space H_w:

$$\begin{cases} d\mathbf{X}_t \ = \ \mathcal{A}\mathbf{X}_t dt + B(\mathbf{X}_t, u_t)dt + \Sigma(\mathbf{X}_t, u_t) dW_t, \\ \mathbf{X}_0 \ = \ \mathbf{x}, \end{cases} \tag{19}$$

At least formally (19) should represent (6) in H_w: if there exists a unique solution (in some sense) $\mathbf{X}$ to (19), we expect that

$$\mathbf{X}_t \ = \ \big(S_t, (S_{t+\xi})_{\xi \in \mathbb{R}^-}\big), \qquad \forall t \geq 0,$$

where S is the solution to (6) with $\mathbf{s} = \mathbf{x}$. We notice that (19) is an equation in infinite dimension, but the noise is one-dimensional[4].

We are going to introduce two concepts of solution to (19), which in this case coincide with each other. Before that we introduce the operator $\mathcal{A}^*$ adjoint of $\mathcal{A}$.

Proposition 3.2. *The adjoint $\mathcal{A}^*$ of the operator $\mathcal{A}$ is defined on*

$$D(\mathcal{A}^*) \ = \ \{\mathbf{x} \in H_w \mid x_1 \in W_w^{1,2}\}, \tag{20}$$

by

$$\mathcal{A}^*\mathbf{x} \ = \ \left(x_1(0), \, -x_1' - x_1 \frac{w'}{w} \right). \tag{21}$$

[4] In the usual language of stochastic integration in infinite-dimension (see [7, 22, 31]), $\Sigma(\mathbf{x},u) \in H_w$ should be seen as a Hilbert-Schmidt operator from $\mathbb{R}$ to H_w.

Proof. Let $\mathcal{D} := \{\mathbf{x} \in H_w \mid x_1 \in W^{1,2}_w\}$, $\mathbf{x} \in \mathcal{D}$ and $\mathbf{y} \in D(\mathcal{A})$. Using (13) and the fact that $y_1(0) = y_0$, we can write

$$\begin{aligned}\langle \mathcal{A}\mathbf{y}, \mathbf{x}\rangle_w &= \int_{\mathbb{R}^-} y_1'(\xi)x_1(\xi)w(\xi)d\xi \\ &= y_1(0)x_1(0) - \int_{\mathbb{R}^-} y_1(\xi)\left(x_1'(\xi) + x_1(\xi)\frac{w'(\xi)}{w(\xi)}\right)w(\xi)d\xi \\ &= y_0x_1(0) - \int_{\mathbb{R}^-} y_1(\xi)\left(x_1'(\xi) + x_1(\xi)\frac{w'(\xi)}{w(\xi)}\right)w(\xi)d\xi.\end{aligned}$$

So, we can conclude that $\mathcal{D} \subset D(\mathcal{A}^*)$, and that $\mathcal{A}^*$ acts as in (21) on $\mathcal{D}$.

Now let us show that actually $D(\mathcal{A}^*) = \mathcal{D}$. Simple computations shows that the expression of the adjoint semigroup of $T(\cdot)$ in H_w is

$$T^*(t) = \left(x_0 + \int_{-t}^0 x_1(\xi),\ x_1(\cdot - t)\frac{w(\cdot - t)}{w(\cdot)}\right), \qquad \mathbf{x} \in H_w.$$

The set $\mathcal{D}$ is clearly dense in H_w and $T^*(t)\mathcal{D} \subset \mathcal{D}$ for any $t \geq 0$. Hence, by [10, Ch. II, Prop. 1.7, p. 53], $\mathcal{D}$ is a core for $D(\mathcal{A}^*)$. On the other hand, in analogy with the proof of Proposition 3.1, one can show that $\mathcal{D}$ is closed with respect to the graph norm, so we conclude that $\mathcal{D} = D(\mathcal{A}^*)$. □

Definition 3.3. (i) Let $\mathbf{x} \in H_w$, and let $u \in \mathcal{U}$. An adapted process $\mathbf{X} = \mathbf{X}^{\mathbf{x},u} \in L^2_{loc}(\mathbb{R}^+; L^2(\Omega; H_w))$ is called mild solution to (19) if for every $t \geq 0$

$$\mathbf{X}_t = T(t)\mathbf{x} + \int_0^t T(t-r)B(\mathbf{X}_r, u(r))dr + \int_0^t T(t-r)\Sigma(\mathbf{X}_r, u(r))dW_r. \quad (22)$$

(ii) Let $\mathbf{x} \in H_w$, $u \in \mathcal{U}$. An adapted process $\mathbf{X} = \mathbf{X}^{\mathbf{x},u} \in L^2_{loc}(\mathbb{R}^+; L^2(\Omega; H_w))$ is called weak solution to (19) if for each $\boldsymbol{\varphi} \in D(\mathcal{A}^*)$ and every $t \geq 0$

$$\begin{aligned}\langle \mathbf{X}_t, \boldsymbol{\varphi}\rangle_w &= \langle \mathbf{x}, \boldsymbol{\varphi}\rangle_w + \int_0^t \langle \mathbf{X}_s, \mathcal{A}^*\boldsymbol{\varphi}\rangle_w ds \\ &\qquad + \int_0^t \langle B(\mathbf{X}_s, u_s), \boldsymbol{\varphi}\rangle_w ds + \int_0^t \langle \Sigma(\mathbf{X}_s, u_s), \boldsymbol{\varphi}\rangle_w dW_s.\end{aligned}$$

Theorem 3.4. *For each $\mathbf{x} \in H_w$ and $u \in \mathcal{U}$, the SDE* (19) *admits a unique (up to indistinguishabiliy) continuous-paths mild solution $\mathbf{X} = \mathbf{X}^{\mathbf{x},u}$ which coincides with the unique weak solution.*

Moreover, we have the equality in $L^2(\Omega, \mathcal{F}, \mathbb{P}; H_w)$

$$\mathbf{X}_t = \left(S_t,\, (S_{t+\xi})_{\xi\in\mathbb{R}^-}\right), \qquad \forall t \geq 0,$$

where S is the solution to (6) *under the control u and with initial datum $\mathbf{s} = \mathbf{x}$.*

Proof. Due to our assumptions, the existence and uniqueness of the continuous-paths mild solution, as well as the fact that it coincides with the (unique) weak solution, is a straightforward application of the theory of infinite-dimensional stochastic differential equations (see, e.g., [7, Th. 7.4] or [22, Th. 3.3] for the existence and uniqueness of mild solutions and [22, Th. 3.2] for the equivalence of weak and mild solutions).

For the second part of the claim, let S be the solution of (6) and define

$$\widetilde{\mathbf{X}}_t := \left(S_t,\, (S_{t+\xi})_{\xi\in\mathbb{R}^-}\right), \quad \mathbf{x} := (s_0, (s_1(\xi))_{\xi\in\mathbb{R}^-}).$$

Then,

$$
\begin{aligned}
S_t &= s_0 + \int_0^t b(S_r, \int_{\mathbb{R}^-} S_{r+\xi}\alpha_1(\xi)w(\xi)d\xi, u(r))dr + \int_0^t \sigma(S_r, \int_{\mathbb{R}^-} S_{r+\xi}\beta_1(\xi)w(\xi)d\xi, u(r))dW_r \\
&= s_0 + \int_0^t b(\langle \mathbf{e}^0, \widetilde{\mathbf{X}}_r\rangle_w, \langle \alpha, \widetilde{\mathbf{X}}_r\rangle_w, u(r))dr + \int_0^t \sigma(\langle \mathbf{e}^0, \widetilde{\mathbf{X}}_r\rangle_w, \langle \beta, \widetilde{\mathbf{X}}_r\rangle_w, u(r))dW_r \\
&= [T(t)\mathbf{x}]_0 + \int_0^t [T(t-r)B(\widetilde{\mathbf{X}}_r, u(r))]_0 dr + \int_0^t [T(t-r)\Sigma(\widetilde{\mathbf{X}}_r, u(r))]_0 dW_r
\end{aligned}
$$

and for every $\xi \in \mathbb{R}^-$,

$$
\begin{aligned}
S_{t+\xi} &= \mathbf{1}_{t+\xi<0} s_1(t+\xi) + \mathbf{1}_{t+\xi\geq 0}\Big\{ s_0 + \int_0^{t+\xi} b(S_r, \int_{\mathbb{R}^-} S_{r+\eta}\alpha_1(\eta)w(\eta)d\eta, u(r))dr \\
&\qquad + \int_0^{t+\xi} \sigma(S_r, \int_{\mathbb{R}^-} S_{r+\eta}\beta_1(\eta)w(\eta)d\eta, u(r))dW_r \Big\} \\
&= \mathbf{1}_{t+\xi<0} s_1(t+\xi) + \mathbf{1}_{t+\xi\geq 0}\Big\{ s_0 + \int_0^{t+\xi} b(\langle \mathbf{e}^0, \widetilde{\mathbf{X}}_r\rangle_w, \langle \alpha, \widetilde{\mathbf{X}}_r\rangle_w, u(r))dr \\
&\qquad + \int_0^{t+\xi} \sigma(\langle \mathbf{e}^0, \widetilde{\mathbf{X}}_r\rangle_w, \langle \beta, \widetilde{\mathbf{X}}_r\rangle_w, u(r))dW_r \Big\} \\
&= s_0\mathbf{1}_{t+\xi\geq 0} + s_1(t+\xi)\mathbf{1}_{t+\xi<0} + \int_0^t \mathbf{1}_{t-r+\xi\geq 0}\, b(\langle \mathbf{e}^0, \widetilde{\mathbf{X}}_r\rangle_w, \langle \alpha, \widetilde{\mathbf{X}}_r\rangle_w, u(r))dr \\
&\qquad + \int_0^t \mathbf{1}_{t-r+\xi\geq 0}\, \sigma(\langle \mathbf{e}^0, \widetilde{\mathbf{X}}_r\rangle_w, \langle \beta, \widetilde{\mathbf{X}}_r\rangle_w, u(r))dW_r \\
&= \Big\{ [T(t)\mathbf{x}]_1 + \int_0^t [T(t-r)B(\widetilde{\mathbf{X}}_r, u(r))]_1 dr + \int_0^t [T(t-r)\Sigma(\widetilde{\mathbf{X}}_r, u(r))]_1 dW_r \Big\}\Big|_\xi
\end{aligned}
$$

which shows that $\widetilde{\mathbf{X}}$ satisfies (22) and therefore coincides with the unique mild solution. □

Since the two concepts of solutions coincide each other in this case, from now on we just say solution to refer to the mild or weak solution. The following technical result will be used in the following section.

Proposition 3.5. *Let $\mathbf{X} = \mathbf{X}^{\mathbf{x},u}$ be the solution to* (19)*. Then*

$$
\mathbb{E}\left[\sup_{0\leq t\leq T} \|\mathbf{X}_t\|_w^2\right] \quad \leq \quad p_1(T)\|\mathbf{x}\|_w^2 + p_2(T), \qquad \forall T \geq 0
$$

where p_1, p_2 are locally bounded functions.

Proof. We notice that

$$
\|B(\mathbf{x}, u)\|_w \quad \leq \quad C_{b,\boldsymbol{\alpha}}(1 + \|\mathbf{x}\|_w), \qquad \|\Sigma(\mathbf{x}, u)\|_w \quad \leq \quad C_{\sigma,\boldsymbol{\beta}}(1 + \|\mathbf{x}\|_w), \tag{23}
$$

where

$$
C_{b,\boldsymbol{\alpha}} \quad = \quad C_2(1 + \|\boldsymbol{\alpha}\|_w), \qquad C_{\sigma,\boldsymbol{\beta}} \quad = \quad C_2(1 + \|\boldsymbol{\beta}\|_w).
$$

Let $T > 0$. Using Definition 3.3-(i) and (15), we have

$$
\begin{aligned}
\mathbb{E}\left[\sup_{0\leq t\leq T} \|\mathbf{X}_t\|_w^2\right] &\leq 3(1+T)^2\|\mathbf{x}\|_w^2 + 3\mathbb{E}\left[\sup_{0\leq t\leq T}\left\| \int_0^t T(t-r)B(\mathbf{X}_r, u(r))dr \right\|_w^2\right] \\
&\qquad + 3\mathbb{E}\left[\sup_{0\leq t\leq T}\left\| \int_0^t T(t-r)\Sigma(\mathbf{X}_r, u(r))dW_r \right\|_w^2\right].
\end{aligned} \tag{24}
$$

By Jensen's inequality (using the convexity of $\|\cdot\|_w^2$) and by the estimates (23) and (15), we deduce

$$\left\| \int_0^t T(t-r)B(\mathbf{X}_r, u(r))dr \right\|_w^2 \leq t \int_0^t \|T(t-r)B(\mathbf{X}_r, u(r))\|_w^2 dr$$
$$\leq t \int_0^t (1+t-r)^2 C_{b,\boldsymbol{\alpha}}^2 (1+\|\mathbf{X}_r\|_w)^2 dr$$
$$\leq 2T(1+T)^2 C_{b,\boldsymbol{\alpha}}^2 \Big(\int_0^T (1+\|\mathbf{X}_r\|_w^2) dr \Big). \quad (25)$$

On the other hand, the estimates (23) and (15), Doob's inequality and Itô's isometry in infinite dimension (see, e.g., [22, Ch, 2] or [31, Ch. 2]) also yield

$$\mathbb{E}\left[\sup_{0\leq t\leq T} \left\| \int_0^t T(t-r)\Sigma(\mathbf{X}_r, u(r))dW_r \right\|_w^2 \right] \ \leq \ 8(1+T)^2 C_{\sigma,\boldsymbol{\beta}}^2 \int_0^T (1+\mathbb{E}\|\mathbf{X}_r\|_w^2) dr.$$

The claim follows from Gronwall's inequality. □

4 Markovian representations

In this section we give sufficient conditions for the existence of exact representations and provide a method to construct approximate representations for the process (S, Z) as a deterministic function of the current state of a finite-dimensional controlled Markov diffusion.

4.1 Preliminaries

The first step is to characterize the finite-dimensional subspaces of H_w which are stable with respect to the operator $\mathcal{A}^*$, which generates the infinite-dimensional structure of our delay equation.

Set

$$\lambda^* \ := \ \inf\{\lambda \in \mathbb{R} \mid e^{\lambda\xi} w^{-\frac{1}{2}}(\xi) \in L^2\}. \quad (26)$$

We introduce the following assumption, which will hold from now on.

Assumption 4.1. *We assume that the infimum in* (26) *is not attained, i.e.*

$$\int_{\mathbb{R}^-} e^{2\lambda^*\xi} w^{-1}(\xi) d\xi \ = \ \infty.$$

Recalling (21) we introduce the differential operator

$$\mathcal{D}_w : \ W_w^{1,2} \ \longrightarrow \ H_w, \qquad v \ \longmapsto \ -\frac{(vw)'}{w} \ = \ -v' - v\frac{w'}{w}.$$

Definition 4.2. We say that a finite dimensional subspace $\mathcal{V}$ of L_w^2 is $\mathcal{D}_w$-stable if $\mathcal{V} \subset W_w^{1,2}$ and $\mathcal{D}_w\mathcal{V} \subset \mathcal{V}$.

We have the following characterization of $\mathcal{D}_w$-stable subspaces.

Proposition 4.3. *$\mathcal{V}$ is an n-dimensional $\mathcal{D}_w$-stable subspace of L_w^2 if and only if*

$$\mathcal{V} \ = \ Span\,\{v_1 w^{-1}, ..., v_n w^{-1}\}, \quad (27)$$

where $\{v_1, ..., v_n\}$ is a set of linearly independent functions such that the vector function $\mathbf{v} = (v_1, ..., v_n)$ is solution of the vector-valued ODE

$$\mathbf{v}'(\xi) \ = \ \mathbf{M}\mathbf{v}(\xi), \quad \xi \leq 0, \quad (28)$$

for some matrix $\mathbf{M} \in \mathbb{R}^{n\times n}$ whose eigenvalues have real part strictly greater than λ^*.

Proof. Let $\mathcal{V}$ be in the form (27) with $\{v_1, ..., v_n\}$ linearly independent functions such that $\mathbf{v} = (v_1, ..., v_n)$ is solution of (28) for some matrix $\mathbf{M} \in \mathbb{R}^{n\times n}$ with eigenvalues $(\lambda_p)_{p=1,...,k}$ such that $Re(\lambda_p) > \lambda^*$ for all $p = 1, ..., k$. Clearly $\dim(V) = n$. Since v solves (28), the v^i's are linear combination of functions of the form

$$\xi^j e^{Re(\lambda_p)\xi} \cos(Im(\lambda_p)\xi), \quad \xi^j e^{Re(\lambda_p)\xi} \sin(Im(\lambda_p)\xi). \tag{29}$$

Then, since $Re(\lambda_p) > \lambda^*$ for all $p = 1, ..., k$, by definition of λ^* we see that

$$w^{-1}v_i \ \in \ W_w^{1,2}, \quad \forall i = 1, ..., n,$$

and therefore $\mathcal{V} \subset W_w^{1,2}$. On the other hand, given $f \in \mathcal{V}$, we have

$$f \ = \ w^{-1}\sum_{i=1}^{n}\mu_i v_i \ = \ w^{-1}\boldsymbol{\mu}^T\mathbf{v}, \quad \text{for some } \boldsymbol{\mu} = (\mu_i)_{i=1,...,n} \in \mathbb{R}^n.$$

Hence, since $\mathbf{v}$ solves (28), we see that

$$\mathcal{D}_w f \ = \ -\frac{(ww^{-1}\boldsymbol{\mu}^T\mathbf{v})'}{w} \ = \ -w^{-1}\boldsymbol{\mu}^T\mathbf{v}' \ = \ -w^{-1}\boldsymbol{\mu}^T\mathbf{M}\mathbf{v} \ \in \ \mathcal{V},$$

showing the $\mathcal{D}_w$-stability of $\mathcal{V}$.

Conversely, let us suppose that $\mathcal{V}$ is an n-dimensional $\mathcal{D}_w$-stable subspace, and let $\{\tilde{v}_1, ..., \tilde{v}_n\}$ be a basis of $\mathcal{V}$. Then $\{\tilde{v}_1 w, ..., \tilde{v}_n w\}$ is a set of linearly independent functions, and, for each $i = 1, ..., n$, there exists $(\tilde{m}_{ij})_{j=1,...n}$, such that

$$-\frac{(\tilde{v}^i w)'}{w} \ = \ \sum_{j=1}^{n}\tilde{m}_{ij}\tilde{v}^j.$$

It follows that $\mathbf{v} = (v_1, ..., v_n) := (\tilde{v}_1 w, ..., \tilde{v}_n w)$ solves (28) with $M = (m_{ij})$, $m_{ij} = -\tilde{m}_{ij}$. Moreover, since $\mathbf{v}$ solves (28), the v^i's must be linear combination of functions of the form (29), where the λ_p's are the eigenvalues of M. So, we also deduce that the eigenvalues of M must have real part strictly greater than λ^*, as $\tilde{v}_i \in L_w^2$ for all $i = 1, ..., n$, and that actually $\tilde{v}_i \in W_w^{1,2}$. □

In view of Proposition 4.3, we see that the n-dimensional $\mathcal{D}_w$-stable subspaces $\mathcal{V}$ of L_w^2 are of the form

$$\mathcal{V} \ = \ \text{Span}\Big\{w(\xi)^{-1}\xi^j e^{Re(\lambda_p)\xi}\cos(Im(\lambda_p)\xi), \ \ w(\xi)^{-1}\xi^j e^{Re(\lambda_p)\xi}\sin(Im(\lambda_p)\xi), \\ 0 \ \le \ j \ \le \ n_p - 1, \ \ 1 \ \le \ p \ \le \ k\Big\}, \tag{30}$$

for some $k \ge 1$, where

$$(n_1, ..., n_k) \ \in \ \mathbb{N}^k \ \text{ s.t. } \ n_1 + ... + n_k \ = \ n, \tag{31}$$

is the vector of multiplicities associated to the (vector of) eigenvalues

$$(\lambda_1, ..., \lambda_k) \ \in \ \mathcal{C}^k, \tag{32}$$

with $\mathcal{C}^k$ defined by

$$\mathcal{C}^k \ = \ \{z = (z_1, ..., z_k) \in \mathbb{C}^k \mid z_i \ne z_j, \ \forall i \ne j; \ \ Re(z_j) > \lambda^*, \ \ \forall j = 1, ..., k; \\ \forall j \in \{1, ..., k\} \ \exists i \in \{1, ..., k\} \text{ s.t. } \bar{z}_i = z_j\}.$$

Conversely, all the subspaces $\mathcal{V}$ of the form (30) above, with $(n_1, ..., n_k)$ and $(\lambda_1, ..., \lambda_p)$ satisfying (31)-(32), are n-dimensional $\mathcal{D}_w$-stable subspaces of L_w^2.

Now, given an n-dimensional subspace $\mathcal{V} \subset L_w^2$, denote

$$\bar{\mathcal{V}} \ := \ \{\mathbf{x} \in H_w \mid x_0 = 0, \ x_1 \in \mathcal{V}\}.$$

Definition 4.4. We say that an $(n+1)$-dimensional subspace $V \subset H_w$ is $\mathcal{A}^*$- stable if $\mathbf{e}^0 \in V$, $V \subset D(\mathcal{A}^*)$ and $\mathcal{A}^* V \subset V$.

Noticing that $\mathcal{A}^*\mathbf{e}^0 = \mathbf{0}$, we immediately get the following corollary.

Corollary 4.5. *An $(n+1)$-dimensional subspace $V \subset H_w$ is $\mathcal{A}^*$- stable if and only if*

$$V \;\; = \;\; Span\ \{\mathbf{e}^0, \bar{\mathcal{V}}\},$$

with $\mathcal{V}$ being some n-dimensional $\mathcal{D}_w$-stable subspace of L^2_w.

4.2 Exact finite-dimensional representation

Suppose that $\boldsymbol{\alpha}, \boldsymbol{\beta} \in V$, where V is an $(n+1)$-dimensional $\mathcal{A}^*$-stable subspace of H_w. Let $\{\mathbf{e}^0, \mathbf{e}^1, ..., \mathbf{e}^n\}$ be an orthonormal basis of $(V, \langle \cdot, \cdot \rangle_w)$ and define

$$X_t^k := \langle \mathbf{e}^k, \mathbf{X}_t \rangle_w, \quad \alpha^k := \langle \mathbf{e}^k, \boldsymbol{\alpha} \rangle_w, \quad \beta^k := \langle \mathbf{e}^k, \boldsymbol{\beta} \rangle_w, \qquad k = 0, ..., n. \tag{33}$$

Then, since the projection of $\boldsymbol{\alpha}, \boldsymbol{\beta}$ onto $V^\perp$ is the null vector and taking into account that $\alpha^0 = \beta^0 = 0$, we have

$$\langle \boldsymbol{\alpha},\ \mathbf{X}_t \rangle = \sum_{k=1}^n \alpha^k X_t^k, \qquad \langle \boldsymbol{\beta},\ \mathbf{X}_t \rangle = \sum_{k=1}^n \beta^k X_t^k,$$

and so from Theorem 3.4 we see that we can write the dynamics of $S_t = X_t^0$ as

$$dS_t \;\; = \;\; b\Big(S_t, \sum_{k=1}^n \alpha^k X_t^k, u_t\Big)dt + \sigma\Big(S_t, \sum_{k=1}^n \beta^k X_t^k, u_t\Big)dW_t. \tag{34}$$

By $\mathcal{A}^*$-stability of V we have the existence of a vector $\mathbf{q} = (q_{k0})_{k=1,\dots,n} \in \mathbb{R}^n$ and of a matrix $\mathbf{Q} = (q_{kh})_{h,k=1,\dots,n} \in \mathbb{R}^{n\times n}$ such that

$$\mathcal{A}^*\mathbf{e}^k \;\; = \;\; q_{k0}\mathbf{e}^0 + \sum_{h=1}^n q_{kh}\mathbf{e}^h, \quad k = 1, .., n. \tag{35}$$

The dynamics of S_t involves the processes X_t^k, $k = 1, ..., n$, whose dynamics, plugging $\mathbf{e}^1, ..., \mathbf{e}^n$ in place of $\boldsymbol{\varphi}$ in the definition of weak solution of (19), can be expressed in terms of themselves and of S by means of the vector $\mathbf{q}$ and of the matrix $\mathbf{Q}$ as

$$dX_t^k \;\; = \;\; \Big(q_{k0}S_t + \sum_{h=1}^n q_{kh}X_t^h\Big)dt, \qquad k = 1, ..., n. \tag{36}$$

The system of $n+1$ equations (34) and (36) provides an $(n+1)$-dimensional Markovian representation of (6), i.e. of S_t, with initial datum $\mathbf{s} = (s_0, s_1(\cdot))$, the corresponding initial data being

$$(x^0, x^1, ..., x^n) \;\; = \;\; \big(\langle \mathbf{e}^0, \mathbf{s}\rangle_w, \langle \mathbf{e}^1, \mathbf{s}\rangle_w, ..., \langle \mathbf{e}^n, \mathbf{s}\rangle_w\big)\,. \tag{37}$$

If also

$$\boldsymbol{\gamma} \;:=\; (0, \gamma_1) \in V,$$

then the projection of $\boldsymbol{\gamma}$ onto $V^\perp$ is the null vector and we can write

$$Z_t \;\; = \;\; \sum_{k=0}^n \gamma^k X_t^k,$$

where

$$\gamma^k \;:=\; \langle \mathbf{e}^k, \boldsymbol{\gamma}\rangle_w, \qquad k = 0, ..., n,$$

obtaining a representation of the process Z in terms of the $(n+1)$-dimensional (controlled) Markov diffusion $(S_t = X^0, X^1, ..., X^n)$.

So we see that in this case the process (S, Z) admits a finite-dimensional Markovian representation in the sense of our definition given in the introduction: there exists a finite-dimensional subspace $V \subset H_w$ such that:

1. V contains the vector $\mathbf{e}^0$;
2. the projection of $\mathbf{X}_t$ on this subspace, i.e., in this case, the vector $(S_t = X^0, X^1, ..., X^n)$ defined above, satisfies a finite dimensional stochastic differential equation;
3. the processes S and Z can be written respectively as $S_t = \langle \mathbf{e}^0, \mathbf{X}_t^V \rangle$ and $Z_t = \langle \gamma, \mathbf{X}_t^V \rangle$.

Remark 4.6. From Proposition 4.3, it is clear that $\boldsymbol{\alpha}, \boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ belong to an $(n+1)$-dimensional $\mathcal{A}_w^*$-stable subspace of H_w (with $\mathcal{A}_w^*$ being the operator $\mathcal{A}^*$ realized with the weight function w) if and only if the original coefficients $\tilde{\boldsymbol{\alpha}}, \tilde{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\gamma}}$ belong to an $(n+1)$-dimensional $\mathcal{A}_1^*$-stable subspace of H (with $\mathcal{A}_1^*$ being the operator $\mathcal{A}^*$ realized with the weight function identically equal to 1). Therefore, the property of having a finite-dimensional Markovian representation does not depend on the choice of the weight function.

4.3 Countable representation

In general $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ fail to lie in an $\mathcal{A}^*$-stable finite-dimensional subspace of H_w. However one can consider an increasing sequence of $\mathcal{A}^*$-stable subspaces of H_w and expand the problem along this sequence. In order to construct such an increasing sequence, we consider specific subclasses of the general representation (30). The simplest case to consider consists in taking $k = 1$ in (30), i.e. considering a sequence of subspaces of L_w^2 of the form

$$\mathcal{V}^n = \text{Span}\left\{ w(\xi)^{-1} \xi^j e^{\lambda \xi}, \quad j = 0, ..., n-1 \right\}, \quad \lambda > \lambda^*. \tag{38}$$

To simplify the orthogonalization procedure of the basis of such subspaces, we restrict our analysis to the case of exponential weights, i.e. $w(\xi) = e^{p\xi}$, $p \in \mathbb{R}$, for which the construction of an orthonormal basis can be reduced to a known case, as we will see below. In this case clearly $\lambda^* = p/2$ and we can choose, e.g., $\lambda > \max\{p, p/2\}$ to satisfy the constraint on λ in (38). With this choice, setting $p_0 := \lambda - p$ we have $p_0 > 0$, and the subspaces in (38) are rewritten as

$$\mathcal{V}^n = \text{Span}\left\{ e^{p_0 \xi} \xi^j, \quad j = 0, ..., n-1 \right\} \tag{39}$$

or equivalently

$$\mathcal{V}^n = \text{Span}\left\{ e^{p_0 \xi} (2p_0 \xi)^j, \quad j = 0, ..., n-1 \right\}. \tag{40}$$

The subspaces (40) for different values of n are the sequence of subspaces we shall consider. An orthogonal basis with respect to the inner product $\langle \cdot, \cdot \rangle_{L_w^2}$ for the subspaces above can be constructed from the Laguerre polynomials as follows. Define, for $k \geq 0$, the Laguerre poynomials

$$\tilde{P}_k(\xi) := \sum_{i=0}^{k} \binom{k}{i} \frac{(-1)^i}{i!} \xi^i, \quad \xi \geq 0.$$

Since we are working with $\mathbb{R}^-$ instead of $\mathbb{R}^+$ the sign of the argument is inverted so we consider Laguerre's polynomials on $\mathbb{R}^-$ defined as

$$P_k(\xi) = \tilde{P}_k(-\xi), \quad \xi \leq 0.$$

From the definition of P_k's, we have

$$P_k(0) = 1, \quad k \geq 0, \tag{41}$$

and moreover, using an induction argument, we get

$$P_k' \;=\; \sum_{i=0}^{k-1} P_i, \quad k \geq 1. \tag{42}$$

As is well known, the Laguerre functions

$$l_{k,p_0}(\xi) \;=\; \sqrt{2p_0}\; P_k(2p_0\xi)\; e^{p_0\xi}$$

are an orthonormal basis for L^2. So the functions

$$L_{k,p_0,p}(\xi) \;=\; e^{-\frac{p}{2}\xi}\; l_{k,p_0}(\xi), \qquad k=0,...,n-1, \tag{43}$$

are an orthonormal basis with respect to the inner product $\langle \cdot, \cdot \rangle_{L^2_w}$ for $\mathcal{V}^n$ defined in (40) and the sequence of functions

$$(L_{k,p_0,p})_{k\geq 0}$$

is an orthonormal basis for L^2_w.

Consider the system of vectors $(\mathbf{e}_k)_{k\geq 0}$ in H_w where

$$\mathbf{e}^0 \;=\; (1,0); \qquad \mathbf{e}^k \;=\; (0, L_{k-1,p_0,p}), \quad k\geq 1.$$

Then, from the argument above, this system is an orthonormal basis in H_w.

Using (41)–(43), we have for $k \geq 1$ (with the convention that $\sum_{i=1}^{0} = 0$)

$$\begin{aligned}
L'_{k-1,p_0,p}(\xi) &= \sqrt{2p_0}\frac{d}{d\xi}(e^{(p_0-p/2)\xi}P_{k-1}(2p_0\xi)) \\
&= (p_0-p/2)L_{k-1,p_0,p}(\xi) + (2p_0)^{\frac{3}{2}} e^{(p_0-p/2)\xi}\sum_{i=0}^{k-2} P_i(2p_0\xi) \\
&= (p_0-p/2)L_{k-1,p_0,p}(\xi) + 2p_0\sum_{i=0}^{k-2} L_{i,p_0,p}(\xi)
\end{aligned}$$

So, by (21)

$$\mathcal{A}^*\mathbf{e}^0 \;=\; \mathbf{0}; \qquad \mathcal{A}^*\mathbf{e}^k \;=\; \mathbf{e}^0\sqrt{2p_0} - 2p_0\sum_{i=1}^{k-1}\mathbf{e}^i - (p_0-p/2)\,\mathbf{e}^k, \quad k\geq 1. \tag{44}$$

Remark 4.7. From (44) we see that, setting

$$V^n \;:=\; \text{Span}\left\{\mathbf{e}^0, ..., \mathbf{e}^n\right\}, \qquad n\geq 0,$$

we have the $\mathcal{A}^*$-stability of V^n for each $n \geq 0$.

Setting

$$X^k(t) \;=\; \langle \mathbf{e}^k, \mathbf{X}_t\rangle_w, \qquad \alpha^k \;=\; \langle \mathbf{e}^k, \boldsymbol{\alpha}\rangle_w, \qquad \beta^k \;=\; \langle \mathbf{e}^k, \boldsymbol{\beta}\rangle_w, \quad k\geq 0, \tag{45}$$

and taking into account that $\alpha^0 = \beta^0 = 0$, we have the Fourier series expansions in H_w

$$\mathbf{X}_t \;=\; \sum_{k=0}^{\infty} X^k(t)\mathbf{e}^k, \qquad \boldsymbol{\alpha} \;=\; \sum_{k=1}^{\infty}\alpha^k\mathbf{e}^k, \qquad \boldsymbol{\beta} \;=\;=\; \sum_{k=1}^{\infty}\beta^k\mathbf{e}^k.$$

Then we can rewrite the dynamics of $S_t = X_t^0$ as

$$dS_t \;=\; b\Big(S_t, \sum_{k=1}^{\infty}\alpha^k X_t^k, u_t\Big)dt \;+\; \sigma\Big(S_t, \sum_{k=1}^{\infty}\beta^k X_t^k, u_t\Big)dW_t. \tag{46}$$

Using the definition of weak solution Definition 3.3-(ii) and considering also (44) we have

$$dX^k(t) \;=\; \Big(\sqrt{2p_0}S_t - 2p_0\sum_{i=1}^{k-1} X_t^i - (p_0 - p/2)X_t^k\Big)dt, \qquad k \geq 1. \tag{47}$$

Setting the initial data

$$\big(s_0, (x^k)_{k\geq 1}\big) \;=\; \big(\langle \mathbf{e}^0, \mathbf{s}\rangle_w, (\langle \mathbf{e}^k, \mathbf{s}\rangle_w)_{k\geq 1}\big), \tag{48}$$

equations (46)–(47) provide a countable Markovian representation of our original system (6). Moreover, setting

$$\gamma^k \;:=\; \langle \mathbf{e}^k, \boldsymbol{\gamma}\rangle_w, \qquad k \geq 0, \tag{49}$$

we have the Fourier series expansion for $\boldsymbol{\gamma}$ (note that $\gamma^0 = 0$)

$$\boldsymbol{\gamma} = \sum_{k=1}^{\infty} \gamma^k \mathbf{e}^k,$$

so we also have the representation of the process (4) as

$$Z_t \;=\; \sum_{k=1}^{\infty} \gamma^k X^k(t). \tag{50}$$

4.4 Approximate finite-dimensional representation

When $\boldsymbol{\alpha}, \boldsymbol{\beta}$ belong to some finite dimensional subspace V^n, equations (46)-(47)-(48) provide a finite-dimensional representation of (6) in the spirit of the previous subsection. In this case the dynamics of S requires only the knowledge of $(X^k)_{k=0,\dots,n}$ and the dynamics of these variables is given also in terms of themselves. Finally, when also $\boldsymbol{\gamma}$ belongs to V^n, then (4) can be written in terms of the finite-dimensional Markov process $(X^k)_{k=0,\dots,n}$ and we fall into an exact finite-dimensional representation of the problem. When some of the above conditions fail to be true (i.e. there is no finite dimensional subspace V^n such that $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma} \in V^n$), then we need to truncate the Fourier series for $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ and work with an approximate finite-dimensional representation of the problem. In this case, setting for $n \geq 0$

$$\boldsymbol{\alpha}^n \;:=\; \sum_{k=1}^{n} \alpha^k \mathbf{e}^k, \qquad \boldsymbol{\beta}^n \;:=\; \sum_{k=1}^{n} \beta^k \mathbf{e}^k, \qquad \boldsymbol{\gamma}^n \;:=\; \sum_{k=1}^{n} \gamma^k \mathbf{e}^k,$$

we have the following estimate for the error.

Proposition 4.8. *For each $n \geq 0$, let $(S^n, (X^{n,k})_{k=1,\dots,n})$ be the finite dimensional Markov diffusion solving*

$$dS_t^n \;=\; b\Big(S_t^n, \sum_{k=1}^{n} \alpha^k X_t^{n,k}, u_t\Big)dt \;+\; \sigma\Big(S_t^n, \sum_{k=1}^{n} \beta^k X_t^{n,k}, u_t\Big)dW_t. \tag{51}$$

$$dX_t^{n,k} \;=\; \Big(\sqrt{2p_0}S_t^n - 2p_0\sum_{i=1}^{k-1} X_t^{n,i} - (p_0 - p/2)X_t^{n,k}\Big)dt, \qquad k = 1, \dots, n, \tag{52}$$

with initial data

$$\big(s_0, (x^{n,k})_{k=1,\dots,n}\big) \;=\; \big(\langle \mathbf{e}^0, \mathbf{s}\rangle_w, (\langle \mathbf{e}^k, \mathbf{s}\rangle_w)_{k=1,\dots,n}\big). \tag{53}$$

Then for every $T > 0$, there exists $C = C_{T,\|\mathbf{s}\|_w,\|\boldsymbol{\alpha}\|_w,\|\boldsymbol{\beta}\|_w,\|\boldsymbol{\gamma}\|_w} < \infty$ such that, uniformly on $u \in \mathcal{U}$,

$$\mathbb{E}\Bigg[\sup_{0\leq t\leq T}\Big|S_t - S_t^n\Big|^2\Bigg] \;\leq\; C(\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^n\|_w^2 + \|\boldsymbol{\beta} - \boldsymbol{\beta}^n\|_w^2) \tag{54}$$

and

$$\mathbb{E}\left[\sup_{0\le t\le T}\left|Z_t-\sum_{k=1}^n\gamma^k X_t^{n,k}\right|^2\right] \;\le\; C(\|\boldsymbol{\alpha}-\boldsymbol{\alpha}^n\|_w^2+\|\boldsymbol{\beta}-\boldsymbol{\beta}^n\|_w^2+\|\boldsymbol{\gamma}-\boldsymbol{\gamma}^n\|_w^2). \qquad (55)$$

Proof. From Proposition 3.5 it follows that S_t, S_t^n, X_t^k and $X_t^{n,k}$ are square integrable for all $t\ge 0$, all $n\ge 1$ and all $k\in\{1,...,n\}$. Using standard tools such as Doob's inequality and Itô's isometry, one can show firstly that

$$\begin{aligned}\mathbb{E}[\sup_{0\le t\le T}(S_t-S_t^n)^2] &\le 2T\int_0^T \mathbb{E}[(b(S_r,\sum_{k=1}^\infty\alpha^k X_r^k,u_r)-b(S_r^n,\sum_{k=1}^n\alpha^k X_r^{n,k},u_r))^2]\\ &\quad+8\int_0^T\mathbb{E}[(\sigma(S_r,\sum_{k=1}^\infty\beta^kX_r^k,u_r)-\sigma(S_r^n,\sum_{k=1}^n\beta^kX_r^{n,k},u_r))^2]\\ &\le 48(T+1)C_1^2\int_0^T\mathbb{E}\left\{(S_r-S_r^n)^2+\left(\sum_{k=n+1}^\infty\alpha^kX_r^k\right)^2+\left(\sum_{k=n+1}^\infty\beta^kX_r^k\right)^2+\left(\sum_{k=1}^n\alpha^k(X_r^k-X_r^{n,k})\right)^2\right\}dr<\infty\end{aligned}$$

and similarly,

$$\mathbb{E}[\sup_{0\le t\le T}(X_t^k-X_t^{n,k})^2]<\infty,\quad k=1,\dots,n.$$

Then, let us introduce the quantity

$$M_T \;:=\; \mathbb{E}[\sup_{0\le t\le T}(S_t-S_t^n)^2]+\sum_{k=1}^n\mathbb{E}[\sup_{0\le t\le T}(X_t^k-X_t^{n,k})^2] \;<\; \infty.$$

From the above estimates, we then get

$$\mathbb{E}[\sup_{0\le t\le T}(S_t-S_t^n)^2] \;\le\; 48(T+1)C_1^2\int_0^T\{(1+\|\boldsymbol{\alpha}^n\|^2)M_r+(\|\boldsymbol{\alpha}-\boldsymbol{\alpha}^n\|_w^2+\|\boldsymbol{\beta}-\boldsymbol{\beta}^n\|_w^2)\mathbb{E}[\|\mathbf{X}_r\|_w^2]\}dr$$

and also for $k=1,\dots,n$,

$$\mathbb{E}[\sup_{0\le t\le T}(X_t^k-X_t^{n,k})^2] \;\le\; (k+1)(4p^2+2p)T\int_0^T M_r\,dr,$$

so that for some constant C depending on p, n, T and α,

$$M_T \;\le\; C\int_0^T\{M_r+(\|\boldsymbol{\alpha}-\boldsymbol{\alpha}^n\|_w^2+\|\boldsymbol{\beta}-\boldsymbol{\beta}^n\|_w^2)\mathbb{E}[\|\mathbf{X}_r\|_w^2]\}dr.$$

From Gronwall's inequality and Proposition 3.5 it follows that there exists another constant, also denoted by C, depending on p, n, T, α, β and the initial condition, such that

$$M_T \;\le\; C(\|\boldsymbol{\alpha}-\boldsymbol{\alpha}^n\|_w^2+\|\boldsymbol{\beta}-\boldsymbol{\beta}^n\|_w^2),$$

from which we deduce (54). Finally,

$$\begin{aligned}\mathbb{E}\left[\sup_{0\le t\le T}\left|Z_t-\sum_{k=1}^n\gamma^kX_t^{n,k}\right|^2\right] &\le 2\mathbb{E}\left[\sup_{0\le t\le T}\left|\sum_{k=n+1}^\infty\gamma^kX_t^k\right|^2\right]+\mathbb{E}\left[\sup_{0\le t\le T}\left|\sum_{k=1}^n\gamma^k(X_t^k-X_t^{n,k})\right|^2\right]\\ &\le 2\|\boldsymbol{\gamma}-\boldsymbol{\gamma}^n\|_w^2\mathbb{E}[\sup_{0\le t\le T}\|\mathbf{X}_t\|_w^2]+\|\boldsymbol{\gamma}^n\|_w^2M_T,\end{aligned}$$

Combining this with the bound on M_T obtained above and Proposition 3.5, we get (55) and the proof is complete. □

This proposition shows that the error of approximating the process S with S^n and the process Z with the linear combination $\sum_{k=1}^n \gamma^k X_t^{n,k}$ of components of the finite-dimensional Markovian diffusion $(X^0, X^1, ..., X^n)$ depends on the error of approximating the coefficients $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ with the corresponding truncated Fourier-Laguerre series. The actual convergence rate as $n \to \infty$ will depend on the regularity of the functions $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. For example, from [4, Lemma A.4] it follows that if these functions are constant in the neighborhood of zero, have compact support and finite variation (this is the case e.g., for uniformly weighted moving averages) and $w \equiv 1$ then $\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^n\|_w^2 + \|\boldsymbol{\beta} - \boldsymbol{\beta}^n\|_w^2 + \|\boldsymbol{\gamma} - \boldsymbol{\gamma}^n\|_w^2 \leq Cn^{-3/2}$ for some constant C and n sufficiently large. For C^∞ functions, on the other hand, the convergence rates are faster than polynomial.

5 Application to optimal control and stopping

In this last section we show how the results of the previous one can be implemented to treat optimal control or optimal stopping problems. Within this section it is assumed that S solves (2)-(3) and Z is the process defined in (4). Moreover the coefficients $\alpha^k, \beta^k, \gamma^k$ are the ones defined in (45) and (49).

5.1 Optimal control problems

Let $T > 0$. Given measurable functions $f : [0,T] \times \mathbb{R}^2 \times U \to \mathbb{R}$, $\phi : \mathbb{R}^2 \to \mathbb{R}$, we consider the following optimal control problem:

$$V(\mathbf{s}) \;\; := \;\; \inf_{u \in \mathcal{U}} J(\mathbf{s}; u), \quad \mathbf{s} = (s_0, s_1) \in H_w,$$

where

$$J(\mathbf{s}; u) \; := \; \mathbb{E}\left[\int_0^T f(t, S_t, Z_t, u_t)dt + \phi(S_T, Z_T)\right].$$

This problem cannot be solved by dynamic programming in finite dimension due to the lack of markovianity. However, given $n \geq 0$, we can consider the problem in $\mathbb{R}^{n+1}$

$$V^n(\mathbf{x}^n) \;\; := \;\; \inf_{u \in \mathcal{U}} J^n(\mathbf{x}^n; u), \qquad \mathbf{x}^n = (x_0^n, x_1^n, ..., x_n^n) \in \mathbb{R}^{n+1},$$

where

$$J^n(\mathbf{x}^n; u) \; := \; \mathbb{E}\left[\int_0^T f(t, X_t^{n,0}, Z_t^n, u_t)dt + \phi(Z_T^n)\right],$$

and

(i) the "output" process Z_T^n is

$$Z_t^n \;\; := \;\; \sum_{k=1}^n \gamma_k X_t^{n,k}; \tag{56}$$

(ii) the state equation for the $(n+1)$-dimensional process $(X^{n,k})_{k=0,...,n}$ is

$$\begin{cases} dX_t^{n,0} \; = \; b\left(X_t^{n,0}, \, \sum_{k=1}^n \alpha^k X_t^{n,k}, \, u_t\right)dt + \sigma\left(X_t^{n,0}, \, \sum_{k=1}^n \beta^k X_t^{n,k}, \, u_t\right)dW_t, \\ dX_t^{n,k} \; = \; \left(\sqrt{2p_0} X_t^{n,0} - 2p_0 \sum_{i=1}^{k-1} X_t^{n,i} - (p_0 - p/2)X_t^{n,k}\right)dt, \qquad k = 1, ..., n, \end{cases} \tag{57}$$

with initial data

$$X_0^{n,k} \;\; = \;\; x_k^n, \qquad k = 0, ..., n. \tag{58}$$

This finite-dimensional problem can be solved via the corresponding Hamilton-Jacobi-Bellman equation (e.g., [18]). The following proposition provides an error estimate for the value function.

Proposition 5.1. *Suppose that $f(t,\cdot,\cdot,u)$ is Lipschitz continuous uniformly in $t \in [0,T]$ and $u \in U$, and that ϕ is Lipschitz continuous. Set*

$$\mathbf{x}^n(\mathbf{s}) \quad := \quad (\langle \mathbf{s}, \mathbf{e}^k \rangle_w)_{k=0,\dots,n}, \quad n \geq 0.$$

Then there exists $K = K_{T,\|\mathbf{s}\|_w,\|\boldsymbol{\alpha}\|_w,\|\boldsymbol{\beta}\|_w,\|\boldsymbol{\gamma}\|_w}$ such that

$$|V(\mathbf{s}) - V^n(\mathbf{x}^n(\mathbf{s}))|^2 \quad \leq \quad K(\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^n\|^2 + \|\boldsymbol{\beta} - \boldsymbol{\beta}^n\|^2 + \|\boldsymbol{\gamma} - \boldsymbol{\gamma}^n\|^2), \qquad \forall n \geq 0.$$

Proof. We shall use the notation $S^{u,\mathbf{s}}$, $Z^{u,\mathbf{s}}$ and $X^{n,0,u,\mathbf{x}^n(\mathbf{s})}$, $Z^{n,u,\mathbf{x}^n(\mathbf{s})}$ to make the dependence on the initial condition and the control explicit. The common Lipschitz constant of $f(t,\cdot,\cdot,u)$ and ϕ shall be denoted by K_0. We have

$$\begin{aligned}
|V(\mathbf{s}) - V^n(\mathbf{x}^n(\mathbf{s}))| &\leq \sup_{u\in\mathcal{U}} |J(\mathbf{s};u) - \inf_{u\in\mathcal{U}} J^n(\mathbf{x}^n(\mathbf{s});u)| \\
&\leq \sup_{u\in\mathcal{U}} \mathbb{E}\left[\int_0^T |f(t,S^{u,\mathbf{s}},Z_t^{u,\mathbf{s}},u_t) - f(t,X^{n,0,u,\mathbf{x}^n(\mathbf{s})},Z_t^{n,u,\mathbf{x}^n(\mathbf{s})},u_t)|dt\right] \\
&\quad + \sup_{u\in\mathcal{U}} \mathbb{E}\left[|\phi(S^{u,\mathbf{s}},Z_T^{u,\mathbf{s}}) - \phi(X^{n,0,u,\mathbf{x}^n(\mathbf{s})},Z_T^{n,u,\mathbf{x}^n(\mathbf{s})})|\right] \\
&\leq K_0 \sup_{u\in\mathcal{U}} \mathbb{E}\left[\int_0^T |S_t^{u,\mathbf{s}} - X_t^{n,0,u,\mathbf{x}^n(\mathbf{s})}|dt + |S_T^{u,\mathbf{s}} - X_T^{n,0,u,\mathbf{x}^n(\mathbf{s})}|\right] \\
&\quad + K_0 \sup_{u\in\mathcal{U}} \mathbb{E}\left[\int_0^T |Z_t^{u,\mathbf{s}} - Z_t^{n,u,\mathbf{x}^n(\mathbf{s})}|dt + |Z_T^{u,\mathbf{s}} - Z_T^{n,u,\mathbf{x}^n(\mathbf{s})}|\right] \\
&\leq K_0 \sup_{u\in\mathcal{U}} \left\{\int_0^T \mathbb{E}[|S_t^{u,\mathbf{s}} - X_t^{n,0,u,\mathbf{x}^n(\mathbf{s})}|^2]^{\frac{1}{2}} dt + \mathbb{E}[|S_T^{u,\mathbf{s}} - X_T^{n,0,u,\mathbf{x}^n(\mathbf{s})}|^2]^{\frac{1}{2}}\right\} \\
&\quad + K_0 \sup_{u\in\mathcal{U}} \left\{\int_0^T \mathbb{E}[|Z_t^{u,\mathbf{s}} - Z_t^{n,u,\mathbf{x}^n(\mathbf{s})}|^2]^{\frac{1}{2}} dt + \mathbb{E}[|Z_T^{u,\mathbf{s}} - Z_T^{n,u,\mathbf{x}^n(\mathbf{s})}|^2]^{\frac{1}{2}}\right\} \\
&\leq 2K_0 C_T^{\frac{1}{2}}(T+1)(\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^n\|_w^2 + \|\boldsymbol{\beta} - \boldsymbol{\beta}^n\|_w^2 + \|\boldsymbol{\gamma} - \boldsymbol{\gamma}^n\|_w^2)^{\frac{1}{2}},
\end{aligned}$$

where the last inequality follows from Proposition 4.8, and C_T is the bound on the constant C of that proposition over $t \in [0,T]$. □

The result above can be applied, for instance, to the problem investigated in [23, 24] or to generalizations of the examples shown in [9].

5.2 Optimal stopping problems

Let $T > 0$ and consider (6) when b, σ do not depend on u (so that the diffusion is actually uncontrolled). Letting $\mathcal{T}$ be the set of all stopping times with respect to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ and taking values in the interval $[0,T]$, and given a measurable function $\phi : [0,T] \times \mathbb{R}^2 \to \mathbb{R}$, we consider following optimal stopping problem:

$$V(\mathbf{s}) \quad := \quad \inf_{\tau\in\mathcal{T}} \mathbb{E}\left[\phi(\tau, S_\tau, Z_\tau)\right].$$

Also in this case the problem cannot be solved by dynamic programming in finite dimension due to the lack of markovianity. However, to approximate its solution, given $n \geq 0$, we can consider the problem in $\mathbb{R}^{n+1}$

$$V^n(\mathbf{x}^n) \quad := \quad \inf_{\tau\in\mathcal{T}} \mathbb{E}\left[\phi(\tau, X_\tau^{n,0}, Z_\tau^n)\right], \qquad \mathbf{x}^n = (x_0^n, x_1^n, \dots, x_n^n) \in \mathbb{R}^{n+1},$$

where the "output" process Z_t^n and the state equation for the $(n+1)$-dimensional process $(X^{n,k})_{k=0,\dots,n}$ are given, respectively, by (56) and (57). The following proposition provides an error estimate.

Proposition 5.2. *Suppose that $\phi(t,\cdot,\cdot)$ is Lipschitz continuous uniformly in $t \in [0,T]$. Set*

$$\mathbf{x}^n(\mathbf{s}) \quad := \quad (\langle \mathbf{s}, \mathbf{e}^k \rangle_w)_{k=0,\dots,n}, \quad n \geq 0.$$

Then there exists $K = K_{T, \|\mathbf{s}\|_w, \|\boldsymbol{\alpha}\|_w, \|\boldsymbol{\beta}\|_w, \|\boldsymbol{\gamma}\|_w}$ such that

$$|V(\mathbf{s}) - V^n(\mathbf{x}^n(\mathbf{s}))|^2 \quad \leq \quad K(\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^n\|^2 + \|\boldsymbol{\beta} - \boldsymbol{\beta}^n\|^2 + \|\boldsymbol{\gamma} - \boldsymbol{\gamma}^n\|^2), \qquad \forall n \geq 0.$$

Proof. Similar to the proof of Proposition 5.1. □

This result can be applied, for example, to the problem of pricing American options written on the moving average of the asset price. This problem was studied in [4] by approximating the dynamics by a finite-dimensional Markovian one using Laguerre polynomials, but without passing through the infinite-dimensional representation of the system. Let us briefly recall the problem. Let $T > 0$ and let S be the price of a stock index and consider the financial problem of pricing an American option whose payoff at the exercise time $t \in [0,T]$ depends on a past average of the stock price, i.e.

$$\phi\left(\frac{1}{\delta}\int_{t-\delta}^{t} S_\xi d\xi\right), \quad \delta > 0.$$

Suppose that the price S is a Markov diffusion solving the SDE

$$dS_t \quad = \quad b\,(S_t)\,dt + \sigma\,(S_t)\,dW_t, \qquad S_0 \;=\; s_0 \;>\; 0, \tag{59}$$

and set

$$Z_t \quad = \quad \frac{1}{\delta}\int_{t-\delta}^{t} S(r)dr \quad = \quad \frac{1}{\delta}\int_{-\delta}^{0} S_{t+\xi} d\xi.$$

Letting $\mathcal{T}$ be the set of all stopping times with respect to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ taking values in the interval $[0,T]$, the value of the option at time 0 is

$$V(s_0, s_1) \quad = \quad \sup_{\tau \in \mathcal{T}} \mathbb{E}\left[\phi(Z(\tau))\right],$$

where $s_1(\xi)$ is the value of the stock at the past time $\xi \in [-\delta, 0]$. This problem is intrinsically infinite-dimensional; it falls into our setting as a special case by taking

$$\boldsymbol{\alpha}, \boldsymbol{\beta} \;=\; (1,0) \in H_w, \qquad \boldsymbol{\gamma} \;=\; \left(0,\; \frac{1}{\delta}\mathbf{1}_{[-\delta,0]}\right) \in H_w.$$

For more details about this problem we refer to [4].

Acknowledgements. The authors are grateful to two anonymous Referees whose suggestions allowed us to improve the first version of the paper. The authors also thank Mauro Rosestolato for very fruitful discussions and suggestions.

References

[1] O. Barndorff-Nielsen, F. Benth, and A. Veraart, *Modelling energy spot prices by Lévy semistationary processes.* CREATES Research Paper, 2010.

[2] H. Bauer and U. Rieder, *Stochastic control problems with delay*, Mathematical Methods of Operations Research, 62 (2005), pp. 411–427.

[3] A. Bensoussan, G. Da Prato, M. Delfour, and S. Mitter, *Representation and control of infinite dimensional systems. Second edition*, Systems & Control: Foundations & Applications, Birkhauser, 2007.

[4] M. Bernhart, P. Tankov, and X. Warin, *A finite dimensional approximation for pricing moving average options*, SIAM Journal of Financial Mathematics, 2 (2011), pp. 989–1013.

[5] V. Blaka Hallulli and T. Vargiolu, *Financial models with dependence on the past: a survey*, in Applied and Industrial Mathematics in Italy, M. Primicerio, R. Spigler, and V. Valente, eds., vol. 69 of Series on Advances in Mathematics for Applied Sciences, World Scientific, 2005, pp. 348–359.

[6] H. Brezis, *Functional analysis, Sobolev spaces and partial differential equations*, Springer, 2010.

[7] G. Da Prato and J. Zabczyck, *Stochastic equations in infinite dimension*, Encyclopedia of Mathematics and Its Applications. Cambridge University Press, 1992.

[8] E. B. Davies, *One-parameter semigroups*, Academic press, 1980.

[9] I. Elsanosi, B. Øksendal, and A. Sulem, *Some solvable stochastic control problems with delay*, Stochastics: An International Journal of Probability and Stochastic Processes, 71 (2000), pp. 69–89.

[10] K. Engel and R. Nagel, *One-paramater semigroups for linear evolution equations*, Springer, 2000.

[11] S. Federico, *A stochastic control problem with delay arising in a pension fund model*, Finance and Stochastics, 15 (2011), pp. 412–459.

[12] S. Federico, B. Goldys, and F. Gozzi, *HJB equations for the optimal control of differential equations with delays and state constraints, I: Regularity of viscosity solutions*, SIAM Journal on Control and Optimization, 48 (2010), pp. 4910–4937.

[13] ——, *HJB equations for the optimal control of differential equations with delays and state constraints, II: Verification and optimal feedbacks*, SIAM Journal on Control and Optimization, 49 (2011), pp. 2378–2414.

[14] S. Federico and B. Øksendal, *Optimal stopping of stochastic differential equations with delay driven by Lévy noise*, Potential analysis, 34 (2011), pp. 181–198.

[15] D. Filipović, *Invariant manifolds for weak solutions to stochastic equations*, Probability theory and related fields, 118 (2000), pp. 323–341.

[16] D. Filipović and J. Teichmann, *Existence of invariant manifolds for stochastic equations in infinite dimension*, Journal of Functional Analysis, 197 (2003), pp. 398–432.

[17] M. Fischer and G. Nappo, *Time discretisation and rate of convergence for the optimal control of continuous-time stochastic systems with delay*, Applied Mathematics and Optimization, 57 (2008), pp. 177–206.

[18] W. Fleming and H. Soner, *Controlled Markov Processes and Viscosity Solutions*, Springer, New York, 2006.

[19] P. Foschi and A. Pascucci, *Path dependent volatility*, Decisions in Economics and Finance, 31 (2008), pp. 13–32.

[20] M. Fuhrman, F. Masiero, and G. Tessitore, *Stochastic equations with delay: Optimal control via BSDEs and regular solutions of Hamilton-Jacobi-Bellman equations*, SIAM Journal on Control and Optimization, 48 (2010), pp. 4624–4651.

[21] P. V. Gapeev and M. Reiss, *An optimal stopping problem in a diffusion-type model with delay*, Statistics & probability letters, 76 (2006), pp. 601–608.

[22] L. Gawarecki and V. Mandrekar, *Stochastic Differential Equations in Infinite Dimensions*, vol. 99 of Research Notes in Mathematics, Pitman, 1984.

[23] F. Gozzi and C. Marinelli, *Stochastic optimal control of delay equations arising in advertising models*, in Stochastic Partial Differential Equations and Applications-VII. Levico Terme, Italy, G. Da Prato and L. Tubaro, eds., CRC Press, 2006, pp. 133–148.

[24] F. Gozzi, C. Marinelli, and S. Savin, *On controlled linear diffusions with delay in a model of optimal advertising under uncertainty with memory effects*, Journal of Optimization: Theory and Applications, 142 (2009), pp. 291–321.

[25] D. G. Hobson and L. C. Rogers, *Complete models with stochastic volatility*, Mathematical Finance, 8 (1998), pp. 27–48.

[26] V. B. Kolmanovskii and T. L. Maizenberg, *Optimal control of stochastic systems with aftereffect*, Avtomat. i Telemeh., 1 (1973), pp. 47–61.

[27] H. J. Kushner, *Numerical approximations for nonlinear stochastic systems with delays*, Stochastics: An International Journal of Probability and Stochastic Processes, 77 (2005), pp. 211–240.

[28] B. Larssen and N. H. Risebro, *When are HJB-equations in stochastic control of delay systems finite dimensional?*, Stochastic analysis and applications, 21 (2003), pp. 643–671.

[29] S.-E. A. Mohammed, *Stochastic differential systems with memory: theory, examples and applications*, in Stochastic analysis and related topics VI, Springer, 1998, pp. 1–77.

[30] B. Øksendal and A. Sulem, *A maximum principle for optimal control of stochastic systems with delay, with applications to finance*, in Proceedings of the conference on optimal control and partial differential equations, Paris, December 2000, J. L. Menaldi, E. Rofman, and A. Sulem, eds., Amsterdam, 2001, IOS Press, pp. 64–79.

[31] C. Prévôt and M. Röckner, *A Concise Course on Stochastic Partial Differential Equations*, Lecture Notes in Mathematics, Springer, 2007.

[32] M. Reiss and M. Fischer, *Discretisation of stochastic control problems for continuous time dynamics with delay*, Journal of Computational and Applied Mathematics, 205 (2007), pp. 969–981.

[33] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, vol. 293 of Grundlehren der Mathematischen Wissenschaften, Springer Verlag, 3rd ed., 1999.

[34] R. Vinter and R. Kwong, *The infinite time quadratic control problem for linear systems with state and control delays: an evolution equation approach*, SIAM Journal of Control and Optimization, 19 (1981), pp. 139–153.