Linear–Quadratic control for a class of stochastic Volterra equations: solvability and approximation

Eduardo Abi Jaber, Enzo Miller, Huyên Pham

Linear–Quadratic control for a class of stochastic Volterra equations: solvability and approximation

Eduardo ABI JABER [*] Enzo MILLER [†] Huyên PHAM [‡]

November 4, 2019

Abstract

We provide an exhaustive treatment of Linear–Quadratic control problems for a class of stochastic Volterra equations of convolution type, whose kernels are Laplace transforms of certain signed matrix measures which are not necessarily finite. These equations are in general neither Markovian nor semimartingales, and include the fractional Brownian motion with Hurst index smaller than 1/2 as a special case. We establish the correspondence of the initial problem with a possibly infinite dimensional Markovian one in a Banach space, which allows us to identify the Markovian controlled state variables. Using a refined martingale verification argument combined with a squares completion technique, we prove that the value function is of linear quadratic form in these state variables with a linear optimal feedback control, depending on non-standard Banach space valued Riccati equations. Furthermore, we show that the value function of the stochastic Volterra optimization problem can be approximated by that of conventional finite dimensional Markovian Linear–Quadratic problems, which is of crucial importance for numerical implementation.

MSC Classification: 93E20, 49N10, 60G22, 60H20.

Key words: Stochastic Volterra equations, linear-quadratic control, Riccati equations in Banach space.

[*]Université Paris 1 Panthéon-Sorbonne, Centre d'Economie de la Sorbonne, 106, Boulevard de l'Hôpital, 75013 Paris, eduardo.abi-jaber at univ-paris1.fr.

[†]LPSM, Université de Paris, Building Sophie Germain, Avenue de France, 75013 Paris, enzo.miller at polytechnique.org

[‡]LPSM, Université de Paris, Building Sophie Germain, Avenue de France, 75013 Paris, pham at lpsm.paris
The work of this author is supported by FiME (Finance for Energy Market Research Centre) and the "Finance et Développement Durable - Approches Quantitatives" EDF - CACIB Chair.

Contents

1 Introduction

Let us consider the basic problem of controlling the drift α of a real-valued Brownian motion W

$$X_t^\alpha = \int_0^t \alpha_s ds + W_t, \quad t \geq 0, \tag{1.1}$$

in order to steer the system to zero with minimal effort by minimizing over a finite horizon the cost functional

$$J(\alpha) = \mathbb{E}\Big[\int_0^T \big(|X_t^\alpha|^2 + \alpha_t^2\big) dt\Big].$$

This problem fits into the class of linear-quadratic (LQ) regulator problem, and can be explicitly solved by different methods including standard dynamic programming, maximum principle or spike variation methods relying on Itô stochastic calculus. It is well-known, see e.g. [30, Chapter 6], that the optimal control α^* is in linear feedback form with respect to the optimal state process $X^* = X^{\alpha^*}$:

$$\alpha_t^* = -\Gamma(t) X_t^*, \quad 0 \leq t \leq T,$$

where Γ is a deterministic nonnegative function solution to a Riccati equation, actually explicitly given by $\Gamma(t) = \tanh(T-t)$, and thus the associated optimal state process X^* is a mean-reverting Markov process.

Suppose now that the noise W is replaced by a Gaussian process with memory, typically a fractional Brownian motion, or more generally by stochastic Volterra processes. It is then natural to ask how the structure of the solution is modified, and how it can be derived, knowing that, in this case, usual methods for Markov processes and stochastic calculus for semimartingales can no longer be applied.

Stochastic Volterra processes appear in different applications for population dynamics, tumour growth, or energy finance, and provide suitable models for dynamics with memory and delay, see [8, 16, 25]. These processes have known a growing interest in finance with the recent empirical findings on rough volatility in [15]. Stochastic Volterra equations have been studied by numerous authors, see [4, 22, 23] and the references therein.

In this paper, we address the optimal control of d-dimensional stochastic Volterra equations of the form:

$$X_t^\alpha = g_0(t) + \int_0^t K(t-s)\Big(b(s, X_s^\alpha, \alpha_s)ds + \sigma(s, X_s^\alpha, \alpha_s)dW_s\Big), \tag{1.2}$$

where g_0 is a deterministic function and K is a (convolution) matrix-valued kernel of the form

$$K(t) = \int_{\mathbb{R}_+} e^{-\theta t}\mu(d\theta), \quad t > 0,$$

for some signed matrix measure μ. Our framework covers the case of the fractional kernel $K(t) = t^{H-1/2}/\Gamma(H+1/2)$ with $H \leq 1/2$, arising from the Mandelbrot-Van Ness representation of the fractional Brownian motion with Hurst index H. We shall mainly focus on the case where the coefficients b and σ are in linear form with respect to the state and control arguments, and the cost to be minimized is of linear-quadratic form.

Since the (controlled) stochastic Volterra process (1.2) is neither Markovian nor a semimartingale, it is natural to consider Markovian lifts for which suitable stochastic tools and control methods apply. Inspired by the Markovian representation of fractional Brownian motion introduced in [10], and more recently generalized to several un–controlled stochastic Volterra equations in [2, 11, 18], we establish the correspondence of the initial problem with a lifted Markovian controlled system $(Y_t^\alpha)_{t\in[0,T]}$ taking its values in the possibly infinite-dimensional Banach space $L^1(\mu)$. Next, in the LQ case, i.e., when b, σ are of linear form, and the cost function is linear-quadratic, we prove by means of a refined martingale verification argument combined with a squares completion technique, that the value function is of quadratic form while the optimal control is in linear feedback form with respect to these lifted state variables. The coefficients of the quadratic and linear form of the value function and optimal control are expressed in terms of a non-standard system of integral operator Riccati equations whose solvability (existence and uniqueness) is proved in [5]. A

related infinite-dimensional Riccati equation appeared in [7] for the minimization problem of an energy functional defined in terms of a non-singular (i.e. $K(0) < \infty$) completely monotone kernel. We stress that, although there exists several results for LQ control problems in infinite-dimension, and even for Volterra processes (see [9]), they cannot be applied in our Banach-space context as they only concern Hilbert spaces. As detailed above, the first contribution of our paper lies in the rigorous derivation of the optimal solution for the stochastic Volterra control problem. A second important feature of our approach is to provide a natural approximation of such solution by a suitable discretization of the measure μ, leading to conventional finite-dimensional LQ control problems, which involve standard matrix Riccati equations that can be numerically implemented.

The paper is organized as follows. In Section 2, we formulate the control problem, justify the correspondence with the lifted Markovian system in the Banach space $L^1(\mu)$, and formally derive the Riccati equation. Section 3 presents the main results:

(i) the analytic expression and solvability of the value function and optimal control in terms of a Banach-space valued Riccati equation. We illustrate our general result on the LQ regulator example mentioned in the beginning of the introduction with a fractional noise with Hurst parameter $H \leq 1/2$;

(ii) a general stability result for the solution of the stochastic Volterra control problem with respect to the kernel and its application for the approximation of the solution.

In Section 4, we prove a general existence result for SDEs with Lipschitz coefficients in Banach spaces, which is used in particular to get the existence of an optimal control for the LQ Volterra control problem. In Section 5, we provide a refined martingale verification theorem for LQ control problem in our context, which mainly relies on Itô's formula for quadratic functions in Banach spaces. The proof of the solvability result (i) is completed in Section 6, and that of the stability result (ii) is detailed in Section 7.

Related literature. The optimal control of stochastic Volterra equations has been considered in [29] by maximum principle method leading to a characterization of the solution in terms of a backward stochastic Volterra equation for the adjoint process. In [6], the authors also use the maximum principle together with Malliavin calculus to obtain a corresponding adjoint equation as a standard backward SDE. Although the kernel considered in these aforementioned papers is not restricted to be of convolution type, the required conditions do not allow singularity of K at zero, hence excluding the case of a fractional kernel with parameter $H < 1/2$. More recently, an extended Bellman equation has been derived in [17] for the associated controlled Volterra equation.

The solution to the LQ control problem as in (1.1) with controlled drift and additive noise has been obtained in [21] when the noise is a fractional Brownian motion with Hurst parameter $H > 1/2$, and in [12] when the noise is a general Gaussian process with an optimal control expressed as the sum of the well-known linear feedback control for the associated deterministic linear-quadratic control problem and the prediction of the response of the system to the future noise process. Recently, the paper [28] investigated LQ problem

of stochastic Volterra equations by providing characterizations of optimal control in terms of some forward-backward system, but leaving aside their solvability, and under some coefficients assumptions that preclude singular kernels such as the fractional kernel with parameter $H < 1/2$.

Notations. For a Banach space $\mathcal{B}$, $L^2([0,T],\mathcal{B})$ denotes the space of measurable and square integrable functions from $[0,T]$ to $\mathcal{B}$.

For any $d \times d_1$-matrix valued measure μ_1 on $\mathbb{R}_+$, we denote by $|\mu_1|$ its total variation, which is a scalar nonnegative measure, refer to [16, Section 3.5] for more details. The space $L^1(\mu_1)$ consists of μ_1-a.e. equivalence classes of $|\mu_1|$-integrable functions $\varphi : \mathbb{R}_+ \to \mathbb{R}^{d_1}$ endowed with the norm $\|\varphi\|_{L^1(\mu_1)} = \int_{\mathbb{R}_+} |\mu_1|(d\theta)|\varphi(\theta)|$, where we identify the function φ with its class of equivalence. For any such φ the integral

$$\int_{\mathbb{R}_+} \mu_1(d\theta)\varphi(\theta)$$

is well defined by virtue of the inequality

$$\left| \int_{\mathbb{R}_+} \mu_1(d\theta)\varphi(\theta) \right| \leq \int_{\mathbb{R}_+} |\mu_1|(d\theta)|\varphi(\theta)|,$$

see [16, Theorem 5.6]. If μ_2 is a $d \times d_2$-matrix valued measure, the space $L^1(\mu_1 \otimes \mu_2)$ consists of $\mu_1 \otimes \mu_2$-a.e. equivalence classes of $|\mu_1| \otimes |\mu_2|$-integrable functions $\Phi : \mathbb{R}_+^2 \to \mathbb{R}^{d \times d}$ endowed with the norm $\|\Phi\|_{L^1(\mu_1 \otimes \mu_2)} = \int_{\mathbb{R}_+^2} |\mu_1|(d\theta)|\Phi(\theta,\tau)||\mu_2|(d\theta) < \infty$. For any such Φ, the integral

$$\int_{\mathbb{R}_+^2} \mu_1(d\theta)^\top \Phi(\theta,\tau)\mu_2(d\tau)$$

is again well defined by virtue of [16, Theorem 5.6]. Both $(L^1(\mu_1), \|\cdot\|_{L^1(\mu)})$ and $(L^1(\mu_1 \otimes \mu_2), \|\cdot\|_{L^1(\mu_1 \otimes \mu_2)})$ are Banach spaces, see [24, Theorem 3.11]. We also denote by $L^\infty(\mu_1)$ the set of measurable functions $\psi : \mathbb{R}_+ \to \mathbb{R}^{d_1}$, which are bounded μ_1-a.e., and by $L^\infty(\mu_1 \otimes \mu_2)$ the set of measurable functions $\Phi : \mathbb{R}_+^2 \to \mathbb{R}^{d \times d}$, which are bounded $\mu_1 \otimes \mu_2$-a.e, that we endow with their usual norms $\|\psi\|_{L^\infty(\mu_1)}$ and $\|\Phi\|_{L^\infty(\mu_1 \otimes \mu_2)}$.

2 Formulation of the problem and preliminaries

Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a filtered probability space supporting a one dimensional Brownian motion W. Fix $T > 0$ and $d, d', m \in \mathbb{N}$. We consider a controlled d-dimensional stochastic Volterra equation

$$X_t^\alpha = g_0(t) + \int_0^t K(t-s)b(s, X_s^\alpha, \alpha_s)ds + \int_0^t K(t-s)\sigma(s, X_s^\alpha, \alpha_s)dW_s, \qquad (2.1)$$

where α is an element of the admissible set

$$\mathcal{A} = \left\{ \alpha : \Omega \times [0,T] \to \mathbb{R}^m \text{ progressively measurable such that } \sup_{0 \le t \le T} \mathbb{E}\left[|\alpha_t|^4\right] < \infty \right\},$$

$g_0 : [0,T] \to \mathbb{R}^d$ is a measurable function, $K : [0,T] \to \mathbb{R}^{d \times d'}$ is a measurable kernel, and $b, \sigma : [0,T] \times \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^{d'}$ are of affine form:

$$\begin{aligned} b(t,x,a) &= \beta(t) + Bx + Ca, \\ \sigma(t,x,a) &= \gamma(t) + Dx + Fa, \end{aligned}$$

where $B, D \in \mathbb{R}^{d' \times d}$, $C, F \in \mathbb{R}^{d' \times m}$, and $\beta, \gamma : [0,T] \to \mathbb{R}^{d'}$ are measurable functions. We are chiefly interested in the case where K is the Laplace transform

$$K(t) = \int_{\mathbb{R}_+} e^{-\theta t} \mu(d\theta), \quad t > 0, \tag{2.2}$$

of a signed $d \times d'$–measure μ satisfying

$$\int_{\mathbb{R}_+} \left(1 \wedge \theta^{-1/2}\right) |\mu|(d\theta) < \infty, \tag{2.3}$$

where $|\mu|$ denotes the total variation of μ. While condition (2.3) does not exclude $\mu_{ij}(\mathbb{R}_+) = \pm\infty$ for some $i \le d$, $j \le d'$, or equivalently a singularity of the kernel K at 0, it does ensure that $K \in L^2([0,T], \mathbb{R}^{d \times d'})$ and that $|\mu|$ is σ-finite, see Lemma A.1. The former implies that the stochastic convolution

$$t \mapsto \int_0^t K(t-s) \xi_s dW_s$$

is well defined as an Itô integral, for every $t \le T$, for any progressively measurable process ξ such that

$$\sup_{t \le T} \mathbb{E}\left[|\xi_t|^2\right] < \infty.$$

Indeed,

$$\mathbb{E}\left[\int_0^t |K(t-s)|^2 |\xi_s|^2 ds\right] \le \|K\|^2_{L^2(0,T)} \sup_{s \le T} \mathbb{E}\left[|\xi_s|^2\right] < \infty,$$

for every $t \le T$. The convolution

$$t \mapsto \int_0^t K(t-s) \xi_s ds,$$

is also well defined for every $t \le T$, by virtue of the Cauchy–Schwarz inequality.

We can now make precise the concept of solution to the controlled equation (2.1). By a solution to (2.1), we mean an $\mathbb{F}$-adapted process X^α with continuous sample paths such that (2.1) holds for all $t \leq T$, $\mathbb{P}$-almost surely. Under (2.2)-(2.3), assuming that β, γ are measurable and bounded, Theorem 4.4 shows that the controlled stochastic Volterra equation (2.1) admits a unique continuous solution X^α, for any continuous input curve g_0, and any admissible control $\alpha \in \mathcal{A}$. Furthermore, it holds that

$$\sup_{0\leq t\leq T} \mathbb{E}\left[|X_t^\alpha|^4\right] < \infty. \tag{2.4}$$

Remark 2.1. Notice that due to the possible singularity of the kernel K, and in contrast with standard stochastic differential equations, the solution X^α to the controlled stochastic Volterra equation does not satisfy in general the usual square integrability condition of the form: $\mathbb{E}[\sup_{0\leq t\leq T} |X_t^\alpha|^2] < \infty$. For this reason, we impose the stronger condition $\sup_{t\leq T} \mathbb{E}[|\alpha|_t^4] < \infty$ for the set of admissible controls $\mathcal{A}$, which will turn out to be crucial for the martingale verification result, see Section 5. □

We consider a cost functional given by

$$J(\alpha) = \mathbb{E}\Big[\int_0^T f(X_s^\alpha, \alpha_s)ds\Big], \tag{2.5}$$

where the running cost f has the following quadratic form

$$f(x, \alpha) = x^\top Q x + \alpha^\top N \alpha + 2x^\top L, \tag{2.6}$$

for some $Q \in \mathbb{S}_+^d, N \in \mathbb{S}_+^m$ and $L \in \mathbb{R}^d$. Here $\mathbb{S}_+^d$ denotes the set of d-dimensional nonnegative symmetric matrices. Note that by virtue of (2.4), $J(\alpha)$ is well defined for any $\alpha \in \mathcal{A}$. The aim is to solve

$$V_0 = \inf_{\alpha\in\mathcal{A}} J(\alpha). \tag{2.7}$$

Before going further, let us mention several kernels of interest that satisfy (2.2)-(2.3).

Example 2.2. (i) Smooth kernels: if $|\mu_{ij}(\mathbb{R}_+)| < \infty$, for every $i = 1, \dots, d, j = 1, \dots, d'$, then (2.3) is satisfied and K is infinitely differentiable on $[0, T]$. This is the case, for instance, when $\mu(d\theta) = \sum_{i=1}^n c_i^n \delta_{\theta_i^n}(d\theta)$, for some $c_i^n \in \mathbb{R}^{d\times d'}$ and $\theta_i^n \in \mathbb{R}_+$, $i = 1, \dots, n$, which corresponds to

$$K(t) = \sum_{i=1}^n c_i^n e^{-\theta_i^n t}.$$

(ii) The fractional kernel ($d = d' = 1$)

$$K_H(t) = \frac{t^{H-1/2}}{\Gamma(H + 1/2)}, \tag{2.8}$$

for some $H \in (0, 1/2)$, which is the Laplace transform of

$$\mu_H(d\theta) = \frac{\theta^{-H-1/2}}{\Gamma(H+1/2)\Gamma(1/2-H)} d\theta, \tag{2.9}$$

and more generally the Gamma kernel $K(t) = K_H(t)e^{-\zeta t}$ for $H \in (0, 1/2)$ and $\zeta \in \mathbb{R}$ for which

$$\mu(d\theta) = \frac{(\theta-\zeta)^{-H-1/2}\mathbb{1}_{(\zeta,\infty)}(\theta)}{\Gamma(H+1/2)\Gamma(1/2-H)} d\theta.$$

(iii) If K_1 and K_2 satisfy (2.2), then so does $K_1 + K_2$ and K_1K_2 with the respective measures $\mu_1 + \mu_2$ and $\mu_1 * \mu_2$. When μ_1, μ_2 satisfy (2.3), it is clear that $\mu_1 + \mu_2$ also satisfies (2.3). This condition is satisfied for the convolution $\mu_1 * \mu_2$ provided $\int_{[1,\infty)^2}(\theta+\tau)^{-1/2}\mu_1(d\theta)\mu_2(d\tau) < \infty$, which is the case for instance if either $\mu_1(\mathbb{R}_+)$ or $\mu_2(\mathbb{R}_+)$ are finite.

(iv) If K is a completely monotone kernel, i.e. K is infinitely differentiable on $(0, \infty)$ such that $(-1)^n K^{(n)}(t)$ is nonnegative for each $t > 0$, then, by Bernstein's theorem, there exists a nonnegative measure μ such that (2.2) holds, see [16, Theorem 5.2.5].

□

2.1 Markovian representation

The solution X^α of (2.1) is in general neither Markovian nor a semimartingale as illustrated by the Riemann–Liouville fractional Brownian motion

$$t \mapsto \frac{1}{\Gamma(H+1/2)} \int_0^t (t-s)^{H-1/2} dW_s, \quad H \in (0, 1/2],$$

which is Markovian and a martingale only for $H = 1/2$. Inspired by the Markovian representation of fractional Brownian motion introduced in [10], and more recently generalized to several un–controlled stochastic Volterra equations for kernels of the form (2.2), see [2, Section 4]; [11, Section 5.1]; [18]; we establish in the following theorem, by means of stochastic Fubini's theorem, the correspondence of (2.1) with a possibly infinite dimensional Markovian controlled system of the form

$$\begin{cases} dY_t^\alpha(\theta) &= \left(-\theta Y_t^\alpha(\theta) + \tilde{b}\left(t, \int_{\mathbb{R}_+} \mu(d\tau) Y_t^\alpha(\tau), \alpha_t\right)\right) dt \\ &\quad + \tilde{\sigma}\left(t, \int_{\mathbb{R}_+} \mu(d\tau) Y_t^\alpha(\tau), \alpha_t\right) dW_t \\ Y_0^\alpha(\theta) &= 0, \end{cases} \tag{2.10}$$

where the coefficients $\tilde{b} : [0,T] \times \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^{d'}$, $\tilde{\sigma} : [0,T] \times \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^{d'}$ are defined by

$$\tilde{b}(t,x,a) = \tilde{\beta}(t) + Bx + Ca,$$
$$\tilde{\sigma}(t,x,a) = \tilde{\gamma}(t) + Dx + Fa,$$

with

$$\tilde{\beta} = \beta + Bg_0 \text{ and } \tilde{\gamma} = \gamma + Dg_0.$$

Theorem 2.3. *Let g_0, β, γ be bounded functions on $[0,T]$ and K be given as in (2.2) such that (2.3) holds. Fix $\alpha \in \mathcal{A}$. Assume that there exists a progressively measurable process X^α that solves (2.1), $\mathbb{P}$-a.s., for each $t \leq T$, and that (2.4) holds. Then, for each $t \leq T$, X_t^α admits the representation*

$$X_t^\alpha = g_0(t) + \int_{\mathbb{R}_+} \mu(d\theta) Y_t^\alpha(\theta), \tag{2.11}$$

where, for each $\theta \in \mathbb{R}_+$,

$$Y_t^\alpha(\theta) = \int_0^t e^{-\theta(t-s)} b(s, X_s^\alpha, \alpha_s) ds + \int_0^t e^{-\theta(t-s)} \sigma(s, X_s^\alpha, \alpha_s) dW_s. \tag{2.12}$$

In particular, Y^α can be chosen to have continuous sample paths in $L^1(\mu)$, satisfies

$$\sup_{t \leq T} \mathbb{E}\left[\|Y_t^\alpha\|_{L^1(\mu)}^4 \right] < \infty, \tag{2.13}$$

$$\sup_{t \leq T} \sup_{\theta \in \mathbb{R}_+} |Y_t^\alpha(\theta)| < \infty, \tag{2.14}$$

and for each $\theta \in \mathbb{R}_+$, $t \mapsto Y_t^\alpha(\theta)$ solves (2.10). Conversely, assume that there exists a process Y^α continuous in $L^1(\mu)$ solution to (2.10), i.e., such that

$$\begin{aligned} Y_t^\alpha(\theta) = &\int_0^t e^{-\theta(t-s)} \tilde{b}\left(s, \int_{\mathbb{R}_+} \mu(d\tau) Y_s^\alpha(\tau), \alpha_s \right) ds \\ &+ \int_0^t e^{-\theta(t-s)} \tilde{\sigma}\left(s, \int_{\mathbb{R}_+} \mu(d\tau) Y_s^\alpha(\tau), \alpha_s \right) dW_s, \quad \mathbb{P} \otimes \mu - a.e. \end{aligned} \tag{2.15}$$

for each $t \leq T$, and that (2.13) holds. Then, the process X^α given by (2.11) is a continuous solution to (2.1) such that (2.4) holds.

Proof. Fix $t \leq T$ and set $Z_t^\alpha = \int_0^t b(s, X_s^\alpha, \alpha_s) ds + \int_0^t \sigma(s, X_s^\alpha, \alpha_s) dW_s$. We first plug (2.2) in (2.1) to get

$$X_t^\alpha - g_0(t) = \int_0^t K(t-s) dZ_s^\alpha = \int_0^t \left(\int_{\mathbb{R}_+} \mu(d\theta) e^{-\theta(t-s)} \right) dZ_s^\alpha.$$

An application of stochastic Fubini's, see [26, Theorem 2.2], yields

$$\int_0^t \left(\int_{\mathbb{R}_+} \mu(d\theta) e^{-\theta(t-s)} \right) dZ_s^\alpha = \int_{\mathbb{R}_+} \mu(d\theta) \left(\int_0^t e^{-\theta(t-s)} dZ_s^\alpha \right),$$

where the interchange is possible since by Jensen's inequality on the normalized measure $(1 \wedge \theta^{-1/2})\mu(d\theta)/\int_{\mathbb{R}_+}(1 \wedge \tau^{-1/2})\mu(d\tau)$ the term

$$\int_{\mathbb{R}_+} \left(\int_0^t e^{-2\theta(t-s)} \mathbb{E}\left[d\langle Z^\alpha \rangle_s \right] \right)^{1/2} |\mu|(d\theta)$$

is bounded from above, for some $c > 0$ by,

$$c \sqrt{\sup_{r \leq T} \left(|\gamma(r)|^2 + \mathbb{E}\left[|\alpha_r|^2 \right] + \mathbb{E}\left[|X_r^\alpha|^2 \right] \right)} \left(\int_{\mathbb{R}_+} \frac{1 - e^{-2\theta t}}{2\theta} \left(1 \wedge \theta^{1/2} \right) |\mu|(d\theta) \right)^{1/2}$$

which is finite due to the inequality

$$\frac{\left(1 - e^{-2\theta t}\right)}{2\theta} \leq \frac{1}{2} \left(1 \vee 2t\right) \left(1 \wedge \theta^{-1}\right), \tag{2.16}$$

condition (2.3), the boundedness of γ, the admissible set $\mathcal{A}$ and the estimate (2.4). The interchange is justified similarly for the drift part. If follows that

$$X_t^\alpha = g_0(t) + \int_{\mathbb{R}_+} \mu(d\theta) \left(\int_0^t e^{-\theta(t-s)} dZ_s^\alpha \right) = g_0(t) + \int_{\mathbb{R}_+} \mu(d\theta) Y_t^\alpha(\theta)$$

where $Y_t^\alpha(\theta)$ is given by (2.12) and corresponds to the variation of constants formula of (2.10). The claimed continuity statement together with (2.13)-(2.14) are proved in Lemma 4.3. The converse is proved along the exact same lines by working them backward. □

Remark 2.4. An alternative lift approach, in the spirit of [2, 13, 17, 20, 27], consists in introducing the double-indexed (controlled) processes

$$G_t^\alpha(u) = \mathbb{E}\left[X_u^\alpha - \int_t^u K(u-s) b(s, X_s^\alpha, \alpha_s) ds \;\middle|\; \mathcal{F}_t \right], \quad 0 \leq t \leq u \leq T.$$

The control problem can then be reformulated in terms of the infinite dimensional controlled Markov process $\{G_t^\alpha(.), t \in [0,T]\}$ with Itô dynamics

$$dG_t^\alpha(u) = K(u-t) \left(b(t, X_t^\alpha, \alpha_t) dt + \sigma(t, X_t^\alpha, \alpha_t) dW_t \right), \quad 0 \leq t < u \leq T.$$

□

2.2 Formal derivation of the solution

Thanks to Theorem 2.3, the possibly non-Markovian initial problem can be formally recast as a degenerate infinite dimensional Markovian problem in $L^1(\mu)$ on the state variables Y^α given by (2.10). To see this, we define the mean-reverting operator A^{mr} acting on measurable functions $\varphi \in L^1(\mu)$ by

$$(A^{mr}\varphi)(\theta) = -\theta\varphi(\theta), \quad \theta \in \mathbb{R}_+, \tag{2.17}$$

and consider the dual pairing

$$\langle \varphi, \psi \rangle_\mu \;\; = \;\; \int_{\mathbb{R}_+} \varphi(\theta)^\top \mu(d\theta)^\top \psi(\theta), \quad (\varphi, \psi) \in L^1(\mu) \times L^\infty(\mu^\top).$$

For any matrix–valued kernel G, we denote by $\boldsymbol{G}$ the integral operator induced by G, defined by:

$$(\boldsymbol{G}\phi)(\theta) = \int_{\mathbb{R}_+} G(\theta, \theta')\mu(d\theta')\phi(\theta').$$

Notice that when $G \in L^\infty(\mu \otimes \mu)$, the operator $\boldsymbol{G}$ is well-defined on $L^1(\mu)$, and we have $\boldsymbol{G}\phi \in L^\infty(\mu^\top)$ for $\phi \in L^1(\mu)$. In this case, $\langle \phi, \boldsymbol{G}\psi \rangle_\mu$ is well defined for all $\varphi, \psi \in L^1(\mu)$. When $G \in L^1(\mu \otimes \mu)$, the operator $\boldsymbol{G}$ is well-defined on $L^\infty(\mu)$, and we have $\boldsymbol{G}\phi \in L^1(\mu^\top)$, for $\phi \in L^\infty(\mu)$. In this case $\langle \boldsymbol{G}\phi, \psi \rangle_{\mu^\top}$ is well defined for all $\varphi, \psi \in L^\infty(\mu)$.

To fix ideas we set $g_0 = \beta = \gamma \equiv 0$ and $L = 0$. Noting that relation (2.15) is the mild form of the linear controlled dynamics in $L^1(\mu)$,

$$dY_t^\alpha = (A^{mr}Y_t^\alpha + \boldsymbol{B}Y_t^\alpha + C\alpha_t)\, dt + (\boldsymbol{D}Y_t^\alpha + F\alpha_t)\, dW_t, \quad Y_0^\alpha = 0,$$

we see that the optimization problem (2.7) can be reformulated as a Markovian problem in $L^1(\mu)$ with cost functional,

$$J(\alpha) = \mathbb{E}\left[\int_0^T \left(\langle Y_s^\alpha, \boldsymbol{Q}Y_s^\alpha \rangle_\mu + \alpha_s^\top N \alpha_s\right) ds\right], \tag{2.18}$$

where, by a slight abuse of notations, C and F denote the respective constant operators from $\mathbb{R}^m$ into $L^\infty(\mu)$ induced by the matrices C and F:

$$(Ca)(\theta) = Ca, \quad (Fa)(\theta) = Fa, \quad \theta \in \mathbb{R}_+, \quad a \in \mathbb{R}^m.$$

Their adjoint operators C^*, F^* from $L^1(\mu^\top)$ into $\mathbb{R}^m$ take the form

$$C^*g \;\; = \;\; C^\top \int_{\mathbb{R}_+} \mu(d\theta)^\top g(\theta), \quad F^*g \;\; = \;\; F^\top \int_{\mathbb{R}_+} \mu(d\theta)^\top g(\theta), \quad g \in L^1(\mu^\top).$$

Given the linear–quadratic structure of the problem, standard results in finite-dimensional stochastic control theory, see [30, Chapter 6], as well as in Hilbert spaces, see [14, 19], suggest that the optimal value process V^α associated to the functional (2.18) should be of linear–quadratic form

$$V_t^{\alpha^*} = \langle Y_t^{\alpha^*}, \mathbf{\Gamma}_t Y_t^{\alpha^*} \rangle_\mu,$$

with an optimal feedback control α^* satisfying

$$\alpha_t^* = -\left(N + F^* \mathbf{\Gamma}_t F\right)^{-1} \left(C^* \mathbf{\Gamma}_t + F^* \mathbf{\Gamma}_t \boldsymbol{D}\right) Y_t^{\alpha^*}, \quad 0 \le t \le T,$$

where $\mathbf{\Gamma}_t$ is a symmetric operator from $L^1(\mu)$ into $L^\infty(\mu^\top)$, and solves the operator Riccati equation:

$$\left\{ \begin{array}{rcl} \mathbf{\Gamma}_T & = & \mathbf{0} \\ \dot{\mathbf{\Gamma}}_t & = & -\mathbf{\Gamma}_t A^{mr} - (\mathbf{\Gamma}_t A^{mr})^* - \boldsymbol{Q} - \boldsymbol{D}^* \mathbf{\Gamma}_t \boldsymbol{D} - \boldsymbol{B}^* \mathbf{\Gamma}_t - (\boldsymbol{B}^* \mathbf{\Gamma}_t)^* \\ & & + \left(C^* \mathbf{\Gamma}_t + F^* \mathbf{\Gamma}_t \boldsymbol{D}\right)^* \left(N + F^* \mathbf{\Gamma}_t F\right)^{-1} \left(C^* \mathbf{\Gamma}_t + F^* \mathbf{\Gamma}_t \boldsymbol{D}\right), \quad t \in [0, T]. \end{array} \right.$$

In particular, when $\mathbf{\Gamma}$ is an integral operator, this formally induces the following Riccati equation for the associated (symmetric) kernel Γ valued in $L^1(\mu \otimes \mu)$:

$$\left\{ \begin{array}{rcl} \Gamma_T(\theta, \tau) & = & 0 \\ \dot{\Gamma}_t(\theta, \tau) & = & (\theta + \tau)\Gamma_t(\theta, \tau) - Q - D^\top \int_{\mathbb{R}_+^2} \mu(d\theta')^\top \Gamma_t(\theta', \tau') \mu(d\tau') D \\ & & - B^\top \int_{\mathbb{R}_+} \mu(d\theta')^\top \Gamma_t(\theta', \tau) - \int_{\mathbb{R}_+} \Gamma_t(\theta, \tau') \mu(d\tau') B + S_t(\theta)^\top \hat{N}_t^{-1} S_t(\tau), \end{array} \right.$$

where

$$S_t(\tau) = C^\top \int_{\mathbb{R}_+} \mu(d\theta)^\top \Gamma_t(\theta, \tau) + F^\top \int_{\mathbb{R}_+^2} \mu(d\theta')^\top \Gamma_t(\theta', \tau') \mu(d\tau') D$$

$$\hat{N}_t = N + F^\top \int_{\mathbb{R}_+^2} \mu(d\theta)^\top \Gamma_t(\theta, \tau) \mu(d\tau) F,$$

and provides an optimal control in the form

$$\alpha_t^* = -\hat{N}_t^{-1} \int_{\mathbb{R}_+} S_t(\theta) \mu(d\theta) Y_t^{\alpha^*}(\theta), \quad 0 \le t \le T.$$

Although the aforementioned infinite dimensional results provide formal expressions for the solution of the problem, they cannot be directly applied, since they concern Hilbert spaces. Here the infinite dimensional controlled process Y^α takes its values in the non reflexive Banach space $\left(L^1(\mu), \|\cdot\|_{L^1(\mu)}\right)$. The rigorous derivation of the solution is the first main objective of the present paper. Our second goal is to show how to obtain an analytic finite-dimensional approximation of the original control problem after a suitable discretization of the operator Riccati equation.

3 Main results

We collect in this section our main results.

3.1 Solvability: optimal control and value function

Let $\alpha \in \mathcal{A}$. Given the linear-quadratic structure of the problem and the formal analysis of Section 2.2, it is natural to consider a candidate optimal value process $(V_t^\alpha)_{t\leq T}$ of linear-quadratic form in the state variable Y^α given by (2.15), that is

$$V_t^\alpha = \int_{\mathbb{R}_+^2} Y_t^\alpha(\theta)^\top \mu(d\theta)^\top \Gamma_t(\theta,\tau)\mu(d\tau)Y_t^\alpha(\tau) + 2\int_{\mathbb{R}_+} \Lambda_t(\theta)^\top \mu(d\theta)Y_t^\alpha(\theta) + \chi_t, \quad (3.1)$$

where the functions $t \mapsto \Gamma_t, \Lambda_t, \chi_t$ are solutions, in a suitable sense, of the following system of Riccati equations:

$$\left\{\begin{array}{rcll} \dot\Gamma_t(\theta,\tau) &=& (\theta+\tau)\Gamma_t(\theta,\tau) - \mathcal{R}_1(\Gamma_t)(\theta,\tau), & \Gamma_T(\theta,\tau) = 0 \\ \dot\Lambda_t(\theta) &=& \theta\Lambda_t(\theta) - \mathcal{R}_2(t,\Gamma_t,\Lambda_t)(\theta), & \Lambda_T(\theta) = 0 \\ \dot\chi_t &=& -\mathcal{R}_3(t,\Gamma_t,\Lambda_t), & \chi_T = 0, \end{array}\right. \quad (3.2)$$

where we defined

$$\begin{aligned} \mathcal{R}_1(\Gamma)(\theta,\tau) =\ & Q + D^\top \int_{\mathbb{R}_+^2} \mu(d\theta')^\top \Gamma(\theta',\tau')\mu(d\tau')D + B^\top \int_{\mathbb{R}_+} \mu(d\theta')^\top \Gamma(\theta',\tau) \\ & + \int_{\mathbb{R}_+} \Gamma(\theta,\tau')\mu(d\tau')B - S(\Gamma)(\theta)^\top \hat N^{-1}(\Gamma)S(\Gamma)(\tau) \quad (3.3) \\ \mathcal{R}_2(t,\Gamma,\Lambda)(\theta) =\ & L + Qg_0(t) + B^\top \int_{\mathbb{R}_+} \mu(d\theta')^\top \Lambda(\theta') + \int_{\mathbb{R}_+} \Gamma(\theta,\tau')\mu(d\tau')\tilde\beta(t) \\ & + D^\top \int_{\mathbb{R}_+^2} \mu(d\theta')^\top \Gamma(\theta',\tau')\mu(d\tau')\tilde\gamma(t) - S(\Gamma)(\theta)^\top \hat N(\Gamma)^{-1}h(t,\Gamma,\Lambda) \quad (3.4) \\ \mathcal{R}_3(t,\Gamma,\Lambda) =\ & g_0(t)^\top Q g_0(t) + \tilde\gamma(t)^\top \int_{\mathbb{R}_+^2} \mu(d\theta')^\top \Gamma(\theta',\tau')\mu(d\tau')\tilde\gamma(t) \\ & + \tilde\beta(t)^\top \int_{\mathbb{R}_+} \mu(d\theta')^\top \Lambda(\theta') - h(t,\Gamma,\Lambda)\hat N(\Gamma)^{-1}h(t,\Gamma,\Lambda), \quad (3.5) \end{aligned}$$

with

$$\begin{aligned} S(\Gamma)(\tau) =\ & C^\top \int_{\mathbb{R}_+} \mu(d\theta)^\top \Gamma(\theta,\tau) + F^\top \int_{\mathbb{R}_+^2} \mu(d\theta')^\top \Gamma(\theta',\tau')\mu(d\tau')D \\ \hat N(\Gamma) =\ & N + F^\top \int_{\mathbb{R}_+^2} \mu(d\theta)^\top \Gamma(\theta,\tau)\mu(d\tau)F \quad (3.6) \\ h(t,\Gamma,\Lambda) =\ & C^\top \int_{\mathbb{R}_+} \mu(d\theta)^\top \Lambda(\theta) + F^\top \int_{\mathbb{R}_+^2} \mu(d\theta)^\top \Gamma(\theta,\tau)\mu(d\tau)\tilde\gamma(t). \end{aligned}$$

The two following definitions specify the concept of solution to the system (3.2).

Definition 3.1. *Let $\Gamma : \mathbb{R}^2_+ \to \mathbb{R}^{d\times d}$ such that $\Gamma \in L^\infty(\mu\otimes\mu)$. We say that Γ is symmetric if*

$$\Gamma(\theta,\tau) = \Gamma(\tau,\theta)^\top, \quad \mu\otimes\mu - a.e.$$

and nonnegative if

$$\int_{\mathbb{R}^2_+} \varphi(\theta)^\top \mu(d\theta)^\top \Gamma(\theta,\tau)\mu(d\tau)\varphi(\tau) \geq 0, \quad \text{for all } \varphi \in L^1(\mu).$$

We denote by $\mathbb{S}^d_+(\mu\otimes\mu)$ the set of all symmetric and nonnegative $\Gamma \in L^\infty(\mu\otimes\mu)$.

Remark 3.2. The integral operator $\mathbf{\Gamma}$ associated to a symmetric kernel $L^\infty(\mu\otimes\mu)$ is symmetric, in the sense that

$$\langle\varphi, \mathbf{\Gamma}\psi\rangle_\mu, = \langle\psi, \mathbf{\Gamma}\varphi\rangle_\mu, \quad \varphi,\psi \in L^1(\mu).$$

Moreover, the nonnegativity of Γ translates into

$$\langle\varphi, \mathbf{\Gamma}\varphi\rangle_\mu \geq 0, \quad \varphi \in L^1(\mu).$$

□

Definition 3.3. *By a solution to the system (3.2), we mean a triplet $(\Gamma,\Lambda,\chi) \in C([0,T], L^1(\mu\otimes\mu)) \times C([0,T], L^1(\mu^\top)) \times C([0,T],\mathbb{R})$ such that*

$$\Gamma_t(\theta,\tau) = \int_t^T e^{-(\theta+\tau)(s-t)}\mathcal{R}_1(\Gamma_s)(\theta,\tau)ds, \qquad 0\leq t\leq T, \quad \mu\otimes\mu - a.e. \tag{3.7}$$

$$\Lambda_t(\theta) = \int_t^T e^{-\theta(s-t)}\mathcal{R}_2(s,\Gamma_s,\Lambda_s)(\theta)ds, \qquad 0\leq t\leq T, \quad \mu - a.e. \tag{3.8}$$

$$\chi_t = \int_t^T \mathcal{R}_3(s,\Gamma_s,\Lambda_s)ds, \qquad 0\leq t\leq T, \tag{3.9}$$

where $\mathcal{R}_1$, $\mathcal{R}_2$ and $\mathcal{R}_3$ are defined respectively by (3.3), (3.4) and (3.5). In particular $\hat{N}(\Gamma_t)$ given by (3.6) is invertible for all $t\leq T$.

The existence and uniqueness of a solution to the Riccati system follows from [5], and is stated in the next theorem. Its proof is given in Section 6.

Theorem 3.4. *Let g_0,β,γ be bounded functions on $[0,T]$. Assume that μ satisfies (2.3) and that*

$$Q \in \mathbb{S}^d_+, \quad N - \lambda I_m \in \mathbb{S}^m_+, \tag{3.10}$$

for some $\lambda > 0$. *Then, there exists a unique triplet* $(\Gamma, \Lambda, \chi) \in C([0,T], L^1(\mu \otimes \mu)) \times C([0,T], L^1(\mu^\top)) \times C([0,T], \mathbb{R})$ *to the system of Riccati equation* (3.2) *such that* (3.7), (3.8), (3.9) *hold and* $\Gamma_t \in \mathbb{S}^d_+(\mu \otimes \mu)$, *for all* $t \leq T$. *Furthermore, there exists some positive constant* $M > 0$ *such that*

$$\int_{\mathbb{R}_+} |\mu|(d\tau)|\Gamma_t(\theta,\tau)| \leq M, \quad \mu - a.e., \quad 0 \leq t \leq T. \tag{3.11}$$

Remark 3.5. Since $\Gamma_t \in \mathbb{S}^d_+(\mu \otimes \mu)$, we have $\Gamma_t \in L^1(\mu \otimes \mu) \cap L^\infty(\mu \otimes \mu)$, for all $t \leq T$. Similarly, $\Lambda_t \in L^1(\mu^\top) \cap L^\infty(\mu^\top)$. To see this, it suffices to observe that since $\Lambda \in C([0,T], L^1(\mu^\top))$, it is bounded in $L^1(\mu^\top)$. Combined with the boundedness of Γ in $L^1(\mu \otimes \mu)$, the estimate (3.11) and the boundedness of the coefficients, we obtain

$$|\mathcal{R}_1(\Gamma_t)(\theta,\tau)| + |\mathcal{R}_2(t,\Gamma_t,\Lambda_t)(\theta)| \leq c, \quad \mu \otimes \mu - a.e., \quad t \leq T,$$

for some constant c. Finally, from (3.8), we get that $\Lambda_t \in L^\infty(\mu^\top)$, for all $t \leq T$. □

Remark 3.6. The process V^α given by (3.1) is well-defined and continuous, due to the continuity of (Γ, Λ, χ), that of Y^α from Theorem 2.3 together with the bounds (2.14) and Remark 3.5. □

Our first main result addresses the solvability of the problem (2.7). Theorem 3.7 establishes the existence of an optimal feedback control of linear form and provides an explicit expression for the value function in terms of the solution to the Riccati equation. The proof is collected in Section 6 and builds upon the results developed in Sections 4 and 5.

Theorem 3.7. *Let* β, γ *be bounded functions on* $[0,T]$ *and* g_0 *continuous. Fix* K, μ *as in* (2.2)-(2.3). *Under* (3.10), *let* (Γ, Λ, χ) *be the solution to the system of Riccati equation* (3.2) *produced by Theorem 3.4. Then, there exists an admissible control* $\alpha^* \in \mathcal{A}$ *with corresponding controlled trajectory* Y^{α^*} *as in* (2.15) *such that*

$$\alpha^*_t = -\hat{N}(\Gamma_t)^{-1}\Big(h(t,\Gamma_t,\Lambda_t) + \int_{\mathbb{R}_+} S(\Gamma_t)(\theta)\mu(d\theta)Y^{\alpha^*}_t(\theta)\Big) \tag{3.12}$$

for all $t \leq T$. *Furthermore,* α^* *is an admissible optimal control, in the sense that*

$$\inf_{\alpha \in \mathcal{A}} J(\alpha) = J(\alpha^*),$$

Y^{α^*} *is the optimally controlled trajectory of the state variable and* $V^{\alpha^*}_t$ *given by* (3.1) *is the optimal value process of the problem, that is*

$$V^{\alpha^*}_t = \inf_{\alpha \in \mathcal{A}_t(\alpha^*)} \mathbb{E}\left[\int_t^T f(X^\alpha_s, \alpha_s)ds \;\Big|\; \mathcal{F}_t\right], \quad 0 \leq t \leq T, \tag{3.13}$$

where $\mathcal{A}_t(\alpha) = \{\alpha' \in \mathcal{A} : \alpha'_s = \alpha_s,\ s \leq t\}$.

Remark 3.8. From (3.13), it follows that at initial time $t = 0$, the optimal value V_0 is equal to $V_0 = V_0^{\alpha^*} = \chi_0$, hence

$$V_0 = \int_0^T \mathcal{R}_3(t, \Gamma_t, \Lambda_t)dt.$$

In particular, for a constant initial condition $g_0(t) \equiv X_0$ for some $X_0 \in \mathbb{R}^d$, we have

$$V_0 = X_0^\top \Psi(T) X_0 + \Phi(T) X_0 + \xi(T),$$

for suitable functions Ψ, Φ, ξ, which corresponds to the usual linear–quadratic form in X_0. However, because of the possible non-Markovianity of the problem, for $t > 0$, the optimal value $V_t^{\alpha^*}$ is not necessarily linear–quadratic in $X_t^{\alpha^*}$ as in the standard case. □

The following example treats the LQ regulator problem (1.1) with a general Volterra noise.

Example 3.9. Let us consider a controlled equation with Volterra noise

$$X_t^\alpha = \int_0^t \alpha_s ds + \int_0^t \tilde{K}(t-s) dW_s,$$

$$J(\alpha) = \mathbb{E}\left[\int_0^T \left(QX_s^2 + N\alpha_s^2\right) ds\right],$$

where $\tilde{K}(t) = \int_{\mathbb{R}_+} e^{-\theta t} \tilde{\mu}(d\theta)$. Notice that X can be recast as

$$X_t^\alpha = \int_0^t K(t-s) \left(C\alpha_s ds + \gamma dW_s\right),$$

where K is the row vector $(1, \tilde{K})$, $C = (1, 0)^\top$ and $\gamma = (0, 1)^\top$. The kernel K is the Laplace transform of the 1×2-matrix measure $\mu = (\delta_0(d\theta), \tilde{\mu}(\mathrm{d}\theta))$. An application of Theorem 3.7 gives an optimal control of feedback form in Y:

$$\alpha_t^* = -\frac{1}{N}\Big[\Gamma_t(0,0) Y_t^1(0) + \int_{\mathbb{R}_+} \Gamma_t(\theta, 0) Y_t^2(\theta) \tilde{\mu}(d\theta)\Big], \tag{3.14}$$

where Γ is solution to the real-valued infinite dimensional Riccati equation

$$\Gamma_t(\theta, \tau) = \int_t^T e^{-(\theta+\tau)(s-t)} \left(Q - \Gamma_s(\theta, 0) N^{-1} \Gamma_s(0, \tau)\right) ds, \quad \tilde{\mu} \otimes \tilde{\mu} - a.e., \quad t \in [0, T],$$

and $Y_t(\theta) = (Y_t^1(\theta), Y_t^2(\theta))^\top = \int_0^t e^{-\theta(s-t)} \left(C\alpha_s^* ds + \gamma dW_s\right)$. In particular,

$$Y_t^2(\theta) = \int_0^t e^{-\theta(t-s)} dW_s$$

$$Y_t^1(0) = \int_0^t \alpha_s^* ds \;=\; X_t^* - \int_0^t \tilde{K}(t-s) dW_s \;=\; X_t^* - \int_{\mathbb{R}_+} Y_t^2(\theta) \tilde{\mu}(d\theta),$$

where $X^* = X^{\alpha^*}$, and the last equality holds by stochastic Fubini's theorem. Plugging the expressions of Y^1 and Y^2 into (3.14) yields

$$\alpha_t^* = -\frac{1}{N}\left(\Gamma_t(0,0)X_t^* + \int_{\mathbb{R}_+}\left(\Gamma_t(\theta,0) - \Gamma_t(0,0)\right)Y_t^2(\theta)\tilde{\mu}(d\theta)\right),$$

which is the sum of a feedback form in X^*, and a second term capturing the non Markovianity of X, as for example in the case of a fractional noise with Hurst parameter $H \leq 1/2$. Note that $\Gamma(0,0)$ satisfies the standard Riccati equation in LQ control problem. One can also note that when $\tilde{K} \equiv 1$ then $\tilde{\mu}(d\theta) = \delta_0(d\theta)$, which implies that the optimal control takes the standard feedback form $\alpha_t^* = -\frac{1}{N}\Gamma_t(0,0)X_t^*$. □

Remark 3.10. Conventional linear–quadratic models, see for instance [30, Chapter 7], are naturally nested in our framework. Indeed, they are recovered by setting $d = d'$ and $\mu = \delta_0 I_d$, which corresponds to $K(t) \equiv I_d$. In this case, the Riccati equations for $\Gamma(0,0), \Lambda(0), \chi$ reduce to the conventional matrix Riccati equations and $Y^\alpha = X^\alpha$ so that we recover the usual expression for the optimal control (3.12) and the value function

$$\begin{aligned}\alpha_t^* &= -\hat{N}(\Gamma_t(0))^{-1}\left(h(t,\Gamma_t(0,0),\Lambda_t(0)) + S(\Gamma_t)(0)X_t^{\alpha^*}\right),\\ V_t &= X_t^\top\Gamma_t(0,0)X_t + 2X_t^\top\Lambda_t(0) + \chi_t.\end{aligned}$$

□

Conventional linear–quadratic models can also be recovered by considering a kernel which is a weighted sum of exponentials as detailed in the following example. This will turn out to be of crucial importance in the next section.

Example 3.11. We set $d = d' = m = 1$ and

$$K^n(t) = \sum_{i=1}^n c_i^n e^{-\theta_i^n t}, \tag{3.15}$$

for some $n \in \mathbb{N}$, $c_i^n \in \mathbb{R}, \theta_i^n \geq 0$, $i = 1,\dots,n$. This corresponds to (2.2) with $\mu(d\theta) = \sum_{i=1}^n c_i^n \delta_{\theta_i^n}(d\theta)$ and Theorem 2.3 gives the representation

$$X_t^{n,\alpha} = g_0^n(t) + \sum_{i=1}^n c_i^n Y_t^{n,i,\alpha}, \tag{3.16}$$

where $Y^{n,i,\alpha} := Y^\alpha(\theta_i^n)$ are such that

$$\begin{aligned}dY_t^{n,i,\alpha} &= \Big[-\theta_i^n Y_t^{n,i,\alpha} + \tilde{b}\Big(t,\sum_{j=1}^n c_j^n Y_t^{n,j,\alpha},\alpha_t\Big)\Big]dt + \tilde{\sigma}\Big(t,\sum_{j=1}^n c_j^n Y_t^{n,j,\alpha},\alpha_t\Big)dW_t\\ Y_0^{n,i,\alpha} &= 0, \quad i = 1,\dots,n.\end{aligned} \tag{3.17}$$

Whence, the problem reduces to a conventional linear-quadratic control for the finite-dimensional controlled system $(Y^{n,i,\alpha})_{1\leq i\leq n}$. In particular, the system of Riccati (3.2) reduces to a a standard one in finite-dimension. For instance the equation for Γ reduces to the standard $n\times n$–matrix Riccati equation

$$\begin{cases} \dot{\Gamma}^n_t &= -Q^n - (B^n)^\top \Gamma^n_t - \Gamma^n_t B^n - (D^n)^\top \Gamma^n_t D^n \\ &\quad + \big((F^n)^\top \Gamma^n_t D^n + (C^n)^\top \Gamma^n_t\big)^\top \big(N^n + (F^n)^\top \Gamma^n_t F^n\big)^{-1} \big((F^n)^\top \Gamma^n_t D^n + (C^n)^\top \Gamma^n_t\big) \\ \Gamma^n_T &= 0, \end{cases} \tag{3.18}$$

where the coefficients $(B^n, C^n, D^n, F^n, N^n, Q^n) \in \mathbb{R}^{n\times n}\times\mathbb{R}^n\times\mathbb{R}^{n\times n}\times\mathbb{R}^n\times\mathbb{R}_+\times\mathbb{S}^n_+$ are defined by

$$\begin{aligned} B^n_{i,j} &= Bc^n_i - \theta^n_i \delta_{ij}, & D^n_{i,j} &= Dc^n_i, \\ C^n_i &= Cc^n_i, & F^n_i &= Fc^n_i, \\ Q^n_{i,j} &= Q, & N^n &= N, \end{aligned}$$

for all $1\leq i,j\leq n$. □

Remark 3.12. The proofs of Theorems 3.4 and 3.7 can be easily adapted to account for a multi-dimensional Brownian motion and time-dependent bounded coefficients. □

3.2 Stability and approximation by conventional LQ problems

The second main result of the paper concerns the approximation of the possibly non-Markovian control problem by sequences of finite dimensional Markovian ones, which is of crucial importance for numerical implementations. The main idea comes from the approximation of the measure μ, appearing in (2.2), by simpler measures μ^n, or equivalently approximating K by simpler kernels K^n given by

$$K^n(t) = \int_{\mathbb{R}_+} e^{-\theta t}\mu^n(d\theta), \quad t>0. \tag{3.19}$$

We also authorize the approximation of the input curve g_0. By substituting (K, g_0) with (K^n, g^n_0), the approximating problem reads

$$V^n_0 = \inf_{\alpha\in\mathcal{A}} J^n(\alpha) \tag{3.20}$$

where

$$\begin{aligned} J^n(\alpha) &= \inf_{\alpha\in\mathcal{A}} \mathbb{E}\left[\int_0^T \Big((X^{n,\alpha}_s)^\top Q X^{n,\alpha}_s + 2L_s^\top X^{n,\alpha}_s + \alpha_s^T N\alpha_s\Big)\, ds\right], \\ X^{n,\alpha}_t &= g^n_0(t) + \int_0^t K^n(t-s) b(s, X^{n,\alpha}_s, \alpha_s) ds + \int_0^t K^n(t-s)\sigma(s, X^{n,\alpha}_s, \alpha_s) dW_s. \end{aligned} \tag{3.21}$$

The following theorem establishes the stability of stochastic Volterra linear–quadratic control problems. Its proof is given in Section 7.2.

Theorem 3.13. *Let β, γ be bounded and measurable functions on $[0,T]$ and g_0 be continuous. Assume that μ satisfies (2.3) and let K be as in (2.2). Let $(g_0^n)_{n\geq 1}$ be a sequence of continuous functions and $(K^n)_{n\geq 1}$ be a sequence of kernels of the form (3.19) with respective measures μ^n satisfying (2.3), for each $n \in \mathbb{N}$. Assume (3.10) and that Q is invertible. Denote by V^* and V^{n*} the respective optimal value processes given by Theorem 3.7 for the respective inputs (g_0, K) and (g_0^n, K^n), for $n \geq 1$. If*

$$\|K^n - K\|_{L^2(0,T)} \to 0 \quad \text{and} \quad \|g_0^n - g_0\|_{L^2(0,T)} \to 0, \quad \text{as } n \to \infty, \tag{3.22}$$

then,

$$V_0^{n*} \to V_0^*, \quad \text{as } n \to \infty, \tag{3.23}$$

with a rate of convergence given by

$$|V_0^* - V_0^{n*}| \leq c\Big(\|g_0^n - g_0\|_{L^2(0,T)} + \|K^n - K\|_{L^2(0,T)}\Big), \tag{3.24}$$

for some positive constant c independent of n.

Combined with Example 3.11, Theorem 3.13 provides an approximation of linear–quadratic stochastic Volterra optimal control problems by conventional Markovian linear–quadratic models in finite dimension. To ease notations we restrict to the case $d = d' = m = 1$, for higher dimension matrices need to be replaced by tensors in what follows. The idea is to approximate μ by a discrete measure μ^n as follows. Fix $n \geq 1$ and $(\eta_i^n)_{0\leq i\leq n}$ a partition of $\mathbb{R}_+$. Let $\mu^n(d\theta) = \sum_{i=1}^n c_i^n \delta_{\theta_i^n}(d\theta)$ with

$$c_i^n = \int_{\eta_{i-1}^n}^{\eta_i^n} \mu(dx) \quad \text{and} \quad \theta_i^n = \frac{1}{c_i^n} \int_{\eta_{i-1}^n}^{\eta_i^n} \theta\mu(d\theta), \quad i = 1, \dots, n. \tag{3.25}$$

Then, for a suitable choice of the partition $(\eta_i^n)_{0\leq i\leq n}$, we obtain the convergence

$$\|K^n - K\|_{L^2(0,T)} \to 0, \quad \text{as } n \to \infty,$$

where K^n is given by (3.15), see for instance [3, Proposition 3.3 and Remark 3.4]. In particular, for the fractional kernel K_H given by (2.8), an even n, and the geometric partition $\eta_i^n = r_n^{i-n/2}$ for $i = 0, \dots, n$, for some $r_n > 1$, the coefficients (3.25) with μ_H as in (2.9) are explicitly given by

$$c_i^n = \frac{(r_n^{1-\alpha} - 1) r_n^{(\alpha-1)(1+n/2)}}{\Gamma(\alpha)\Gamma(2-\alpha)} r_n^{(1-\alpha)i} \quad \text{and} \quad x_i^n = \frac{1-\alpha}{2-\alpha} \frac{r_n^{2-\alpha} - 1}{r_n^{1-\alpha} - 1} r_n^{i-1-n/2}, \quad i = 1, \dots, n,$$

where $\alpha := H + 1/2$. If the sequence $(r_n)_{n\geq 1}$ satisfies

$$r_n \downarrow 1 \quad \text{and} \quad n \ln r_n \to \infty, \quad \text{as } n \to \infty,$$

then,

$$\|K^n - K_H\|_{L^2(0,T)} \to 0, \quad \text{as } n \to \infty,$$

see [1, Lemma A.3]. In practice, the free parameter r_n can be chosen by minimizing the L^2 norm between K^n and K_H, for instance if $n = 20$, setting $r_{20} = 2.5$ yields very good approximations for the un-controlled stochastic Volterra equation, see [1] for a more detailed practical study. For each n, the approximate control problem is a conventional linear quadratic one in finite dimension for the state variables (3.17) with the standard $n \times n$ matrix Riccati equation (3.18). This allows to numerically solve the Riccati equations and simulate the process $X^{n,\alpha}$ given by (3.16), leading to computation of the value function V_0^{n*} and the optimal control α^n as in (3.12) with μ replaced by μ^n.

4 An infinite dimensional SDE with Lipschitz coefficients

We aim to establish the existence of a solution to the stochastic Volterra equation and that of an admissible optimal control. For this, we shall study more generally the existence and uniqueness of a solution to an infinite dimensional stochastic differential equation (SDE) in $L^1(\mu)$. Throughout this section, we fix $t \in [0,T]$, $d, d', n \in \mathbb{N}$, $p \geq 2$, μ a $d \times d'$-measure satisfying (2.3), and W denotes an n–dimensional standard Brownian motion.

Let us consider the infinite dimensional SDE in $L^1(\mu)$:

$$d\widetilde{Y}_s = \big(A^{mr}\widetilde{Y}_s + \delta(s, \widetilde{Y}_s)\big)ds + \Sigma(s, \widetilde{Y}_s)dW_s, \qquad \widetilde{Y}_t = \xi, \tag{4.1}$$

on $[t,T]$, where A^{mr} is the mean-reverting operator as defined in (2.17), the inputs $\xi \in L^1(\mu)$, and $\delta : [0,T] \times \Omega \times L^1(\mu) \to L^\infty(\mu)$, $\Sigma : [0,T] \times \Omega \times L^1(\mu) \to L^\infty(\mu)^n$.

We look for $L^1(\mu)$–valued solutions to (4.1) in the strong probabilistic sense and in the mild analytical sense. More precisely, given a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_s)_{s\geq 0}, \mathbb{P})$ supporting a n dimensional Brownian motion W, we say that a progressively measurable process $\widetilde{Y}$ is a (mild) solution to (4.1) on $[t,T]$ if for each $s \in [t,T]$,

$$\begin{aligned}\widetilde{Y}_s(\theta) \;=\; e^{-\theta(s-t)}\xi(\theta) + \int_t^s e^{-\theta(s-u)}\delta(u, \widetilde{Y}_u)(\theta)du \\ + \int_t^s e^{-\theta(s-u)}\Sigma(u, \widetilde{Y}_u)(\theta)dW_u, \qquad \mu - a.e.,\end{aligned} \tag{4.2}$$

such that

$$\sup_{t\leq s\leq T} \mathbb{E}\left[\|\widetilde{Y}_s\|^p_{L^1(\mu)}\right] < \infty. \tag{4.3}$$

The following theorem establishes the strong existence and uniqueness of a solution to (4.2) under Lipschitz conditions.

Theorem 4.1. *Fix $p \geq 2$ and $t \leq T$. Assume that δ and Σ are progressively measurable and that there exists positive constants c_{LG}, c_{Lip}, and a progressively measurable process ϕ with*

$$\sup_{t\leq s\leq T} \mathbb{E}[|\phi_s|^p] < \infty,$$

such that for all $y, y' \in L^1(\mu)$, and $t \leq s \leq T$,

$$|\delta(s,y)(\theta)| + |\Sigma(s,y)(\theta)| \leq c_{\mathrm{LG}} \left(1 + |\phi_s| + \|y\|_{L^1(\mu)}\right), \tag{4.4}$$

$$|\delta(s,y)(\theta) - \delta(s,y')(\theta)| + |\Sigma(s,y)(\theta) - \Sigma(s,y')(\theta)| \leq c_{\mathrm{Lip}} \|y - y'\|_{L^1(\mu)}, \tag{4.5}$$

$\mathbb{P}\otimes\mu - a.e.$ Then, for any $\mathcal{F}_t$–measurable random variable $\mathbb{E}[\|\xi\|^p_{L^1(\mu)}] < \infty$, there exists a unique strong solution Y to (4.2) on $[t,T]$ such that (4.3) holds.

Proof. The proof is an application of the contraction mapping principle. We denote by $\mathcal{S}^p_{t,T}$ the space of progressively measurable processes $\widetilde{Y} : \Omega \times [t,T] \to L^1(\mu)$ such that

$$\|\widetilde{Y}\|_{\mathcal{S}^p_{t,T}} := \sup_{t\leq s\leq T} \mathbb{E}\left[\|\widetilde{Y}_s\|^p_{L^1(\mu)}\right]^{1/p} < \infty.$$

$(\mathcal{S}^p_{t,T}, \|\cdot\|_{\mathcal{S}^p_{t,T}})$ is a Banach space. We consider the following family of norms on $\mathcal{S}^p_{t,T}$:

$$\|\widetilde{Y}\|_\lambda := \sup_{t\leq s\leq T} e^{-\lambda(s-t)} \mathbb{E}\left[\|\widetilde{Y}_s\|^p_{L^1(\mu)}\right]^{1/p}, \quad \lambda > 0.$$

For every $\widetilde{Y} \in \mathcal{S}^p_{t,T}$, define a new process $\mathcal{T}\widetilde{Y}$ by

$$\begin{aligned}(\mathcal{T}\widetilde{Y})_s(\theta) =\ & e^{-\theta(s-t)}\xi(\theta) + \int_t^s e^{-\theta(s-u)}\delta(u,\widetilde{Y}_u)(\theta)du \\ & + \int_t^s e^{-\theta(s-u)}\Sigma(u,\widetilde{Y}_u)(\theta)dW_u \\ =\ & \mathbf{I}_s(\theta) + \mathbf{II}_s(\theta) + \mathbf{III}_s(\theta), \qquad \mu - a.e., \quad t \leq s \leq T.\end{aligned}$$

Since the norms $\|\cdot\|_{\mathcal{S}^p_{t,T}}$ and $\|\cdot\|_\lambda$ are equivalent, it is enough to find $\lambda > 0$ such that $\mathcal{T}$ defines a contraction on $(\mathcal{S}^p_{t,T}, \|\cdot\|_\lambda)$. That is, we look for $\lambda > 0$ and $M < 1$ such that

$$\|\mathcal{T}\widetilde{Y} - \mathcal{T}\widetilde{Z}\|_\lambda \leq M\|\widetilde{Y} - \widetilde{Z}\|_\lambda, \quad \widetilde{Y}, \widetilde{Z} \in \mathcal{S}^p_{t,T}. \tag{4.6}$$

Step 1: We first prove that $\mathcal{T}(\mathcal{S}^p_{t,T}) \subset \mathcal{S}^p_{t,T}$. Fix $\widetilde{Y} \in \mathcal{S}^p_{t,T}$ and $t \leq s \leq T$. $\mathcal{T}\widetilde{Y}$ is again progressively measurable. Jensen's inequality applied to the convex function $\|\cdot\|^p_{L^1(\mu)}$ leads to

$$\|(\mathcal{T}\widetilde{Y})_s\|^p_{L^1(\mu)} \leq 3^{p-1}\left(\|\mathbf{I}_s\|^p_{L^1(\mu)} + \|\mathbf{II}_s\|^p_{L^1(\mu)} + \|\mathbf{III}_s\|^p_{L^1(\mu)}\right).$$

Since $\mathbb{E}[\|\xi\|^p_{L^1(\mu)}] < \infty$, we have

$$\mathbb{E}\big[\|\mathbf{I}_s\|^p_{L^1(\mu)}\big] = \mathbb{E}\left[\|e^{-(\cdot)(s-t)}\xi\|^p_{L^1(\mu)}\right] \leq \mathbb{E}\left[\|\xi\|^p_{L^1(\mu)}\right] < \infty,$$

where we used the bound $e^{-\theta(s-t)} \leq 1$, since μ is supported on $\mathbb{R}_+$. Three successive application of Jensen's inequality on the normalized measures

$$\frac{(1 \wedge \theta^{-1/2})\mu(d\theta)}{\int_{\mathbb{R}_+}(1 \wedge \tau^{-1/2})\mu(d\tau)}, \quad \frac{du}{(s-t)}, \quad \frac{e^{-2\theta(s-u)}du}{\int_s^t e^{-2\theta(s-v)}dv},$$

yield for a constant c that may vary from line to line

$$\begin{aligned}
\|\mathbf{II}_s\|^p_{L^1(\mu)} &= \left(\int_{\mathbb{R}_+} |\mu|(d\theta)\left|\int_t^s e^{-\theta(s-u)}\delta(u,\widetilde{Y}_u)(\theta)du\right|\right)^p \\
&\leq c\int_{\mathbb{R}_+} |\mu|(d\theta)\left(1 \wedge \theta^{(p-1)/2}\right)\left|\int_t^s e^{-\theta(s-u)}\delta(u,\widetilde{Y}_u)(\theta)du\right|^p \\
&\leq c\int_{\mathbb{R}_+} |\mu|(d\theta)\left(1 \wedge \theta^{(p-1)/2}\right)\left|\int_t^s e^{-2\theta(s-u)}|\delta(u,\widetilde{Y}_u)(\theta)|^2du\right|^{p/2} \\
&\leq c\int_{\mathbb{R}_+} |\mu|(d\theta)\left(1 \wedge \theta^{(p-1)/2}\right) \\
&\qquad \times \int_t^s e^{-2\theta(s-u)}|\delta(u,\widetilde{Y}_u)(\theta)|^p du\left(\int_t^s e^{-2\theta(s-v)}dv\right)^{(p-2)/2}.
\end{aligned}$$

Taking expectation combined with the growth condition (4.4) and the fact that $\widetilde{Y} \in \mathcal{S}^p_{t,T}$ leads to

$$\begin{aligned}
\mathbb{E}\left[\|\mathbf{II}_s\|^p_{L^1(\mu)}\right] &\leq c\left(1 + \sup_{t\leq u\leq T}\mathbb{E}\left[|\phi_u|^p\right] + \|\widetilde{Y}\|^p_{\mathcal{S}^p_{t,T}}\right) \\
&\qquad \times \int_{\mathbb{R}_+} |\mu|(d\theta)\left(1 \wedge \theta^{(p-1)/2}\right)\left(\int_t^s e^{-2\theta(s-u)}du\right)^{p/2} \\
&= c\int_{\mathbb{R}_+} |\mu|(d\theta)\left(1 \wedge \theta^{(p-1)/2}\right)\left(\frac{1-e^{-2\theta(s-t)}}{2\theta}\right)^{p/2}.
\end{aligned}$$

Similarly, combining the same Jensen's inequalities with the Burkholder-Davis-Gundy

inequality, we get

$$\begin{aligned}
\mathbb{E}\left[\|\mathbf{III}_s\|^p_{L^1(\mu)}\right] &= \mathbb{E}\left[\left(\int_{\mathbb{R}_+}|\mu|(d\theta)\left|\int_t^s e^{-\theta(s-u)}\Sigma(u,\widetilde{Y}_u)(\theta)dW_u du\right|\right)^p\right] \\
&\leq c\int_{\mathbb{R}_+}|\mu|(d\theta)(1\wedge\theta^{(p-1)/2}) \\
&\qquad \times\int_t^s e^{-2\theta(s-u)}\mathbb{E}\left[|\Sigma(u,\widetilde{Y}_u)(\theta)|^p\right]du\left(\int_s^t e^{-2\theta(s-v)}dv\right)^{(p-2)/2} \\
&\leq c\int_{\mathbb{R}_+}|\mu|(d\theta)\left(1\wedge\theta^{(p-1)/2}\right)\left(\frac{1-e^{-2\theta(s-t)}}{2\theta}\right)^{p/2}
\end{aligned}$$

where the last inequality follows from the growth condition (4.4) and the fact that $\widetilde{Y}\in\mathcal{S}^p_{t,T}$. Recalling inequality (2.16), we get that

$$\int_{\mathbb{R}_+}|\mu|(d\theta)\left(1\wedge\theta^{(p-1)/2}\right)\left(\frac{1-e^{-2\theta(s-t)}}{2\theta}\right)^{p/2}\leq c\int_{\mathbb{R}_+}|\mu|(d\theta)\left(1\wedge\theta^{-1/2}\right) \tag{4.7}$$

which is finite due to condition (2.3). This shows that

$$\mathbb{E}\left[\|\mathbf{II}_s\|^p_{L^1(\mu)}\right]+\mathbb{E}\left[\|\mathbf{III}_s\|^p_{L^1(\mu)}\right]\ \leq\ c\ <\ \infty.$$

Combining the above proves that $\|\mathcal{T}\widetilde{Y}\|_{\mathcal{S}^p_{t,T}}<\infty$ and hence $\mathcal{T}:\mathcal{S}^p_{t,T}\to\mathcal{S}^p_{t,T}$.

Step 2: We prove that there exists $\lambda>0$ such that (4.6) holds. Let $\widetilde{Y},\widetilde{Z}\in\mathcal{S}^p_{t,T}$ such that $\|\widetilde{Y}\|_\lambda$ and $\|\widetilde{Z}\|_\lambda$ are finite. Similarly to Step 1, Jensen and Burkholder–Davis–Gundy inequalities combined with the Lipschitz condition (4.5) lead to

$$\|\mathcal{T}\widetilde{Y}-\mathcal{T}\widetilde{Z}\|^p_\lambda\ \leq\ M(\lambda)\|\widetilde{Y}-\widetilde{Z}\|^p_\lambda,$$

where

$$M(\lambda)=c\int_{\mathbb{R}_+}|\mu|(d\theta)\left(1\wedge\theta^{(p-1)/2}\right)\int_t^s e^{-2\theta(s-u)}e^{-\lambda p(s-u)}du\left(\int_t^s e^{-2\theta(s-v)}dv\right)^{(p-2)/2}.$$

By the dominated convegence theorem, recall (4.7), $M(\lambda)$ tends to 0 as λ goes to $+\infty$. We can therefore choose $\lambda_0>0$ so that (4.6) holds with $M(\lambda_0)<1$. An application of the contraction mapping theorem yields the claimed existence and uniqueness statement in $(\mathcal{S}^p_{t,T},\|\cdot\|_{\mathcal{S}^p_{t,T}})$ together with (4.3). □

Example 4.2. Fix $\alpha \in \mathcal{A}$, the conditions (4.4)-(4.5) are satisfied for the following specification of δ and Σ:

$$\delta(s,\omega,y)(\theta) = b_0(s,\omega,\theta) + \int_{\mathbb{R}_+} B_0(s,\omega,\theta,\tau)\mu(d\tau)y(\tau) + C_0(s,\omega,\theta)\alpha_s(\omega)$$
$$\Sigma(s,\omega,y)(\theta) = \gamma_0(s,\omega,\theta) + \int_{\mathbb{R}_+} D_0(s,\omega,\theta,\tau)\mu(d\tau)y(\tau) + F_0(s,\omega,\theta)\alpha_s(\omega),$$

where

$$|b_0(s,\theta)| + |\gamma_0(s,\theta)| + |C_0(s,\theta)| + |F_0(s,\theta)| \leq c, \quad \mathbb{P}\otimes\mu - a.e., \quad t \leq s \leq T, \tag{4.8}$$
$$|B_0(s,\theta,\tau)| + |D_0(s,\theta,\tau)| \leq c, \quad \mathbb{P}\otimes\mu\otimes\mu - a.e., \quad t \leq s \leq T, \tag{4.9}$$

for some constant c. □

The existence and uniqueness of a strong solution to the stochastic Volterra equation (2.1) readily follows from Theorem 4.1 when combined with Theorem 2.3. To prove continuity of the solution we need the following lemma.

Lemma 4.3. *Let $Z_t = \int_0^t b_s ds + \int_0^t \sigma_s dW_s$, $0 \leq t \leq T$, such that b and σ are progressively measurable and*

$$\sup_{t\leq T} \mathbb{E}\left[|b_t|^4\right] + \sup_{t\leq T} \mathbb{E}\left[|\sigma_t|^4\right] < \infty.$$

Then, the process

$$\widetilde{Y}_t(\theta) \;=\; \int_0^t e^{-\theta(t-s)} dZ_s, \quad \theta \in \mathbb{R}_+,$$

solution in the mild sense to

$$d\tilde{Y}_s = \; A^{mr}\tilde{Y}_s ds + dZ_s, \quad \tilde{Y}_0 \;=\; 0,$$

admits a continuous modification in $L^1(\mu)$ and satisfies (2.13)-(2.14).

Proof. The bound (2.13) follows along the lines of the estimates in step 1 of the proof of Theorem 4.1 with $p = 4$, for getting (4.3), by successive applications of Jensen inequalities. Let us now show (2.14) and the continuity statement. Fix $\theta \in \mathbb{R}_+$ and $t \leq T$. An integration by parts leads to

$$\widetilde{Y}_t(\theta) = e^{-\theta t} Z_t + \theta \int_0^t e^{-\theta(t-s)}(Z_t - Z_s) ds.$$

The Kolmogorov–Chentsov continuity criterion, yields that for each $\zeta \in (0, 1/4)$, the process Z admits a version with ζ-Hölder sample paths on $[0, T]$. We identify Z with this version so that

$$|Z_t(\omega) - Z_s(\omega)| \leq c_{T,\zeta}(\omega)|t - s|^\zeta, \quad s, t \leq T,$$

for some $c_{T,\zeta}(\omega) \geq 0$. Using this inequality and another integration by parts yields

$$\begin{aligned} |\widetilde{Y}_t(\theta,\omega)| \leq\ & c_{T,\zeta}(\omega)e^{-\theta t}t^\zeta + c_{T,\zeta}(\omega)\theta \int_0^t e^{-\theta u}u^\zeta du \\ =\ & c_{T,\zeta}(\omega)\zeta \int_0^t e^{-\theta u}u^{\zeta-1}du. \end{aligned}$$

This proves (2.14). Furthermore,

$$\sup_{t\leq T}|\widetilde{Y}_t(\theta,\omega)| \leq c_{T,\zeta}(\omega)\zeta \int_0^T e^{-\theta u}u^{\zeta-1}du,$$

where the right hand side is in $L^1(|\mu|)$ by virtue of the Cauchy–Schwarz inequality and Lemma A.1. Since $t \mapsto \widetilde{Y}_t(\theta,\omega)$ is continuous for each $\theta \in \mathbb{R}_+$, the dominated convergence theorem yields that the process $\widetilde{Y}$ is continuous in $L^1(\mu)$. □

Theorem 4.4. *Let g_0 be continuous, β, γ be bounded measurable functions on $[0,T]$ and K be a kernel as in (2.2) such that (2.3) holds. Fix an admissible control $\alpha \in \mathcal{A}$. The stochastic Volterra equation (2.1) admits a unique continuous and adapted strong solution X^α such that (2.4) holds.*

Proof. Existence, uniqueness and (2.4) are straightforward from Theorem 4.1 combined with Theorem 2.3 and Example 4.2 for the coefficients

$$\begin{aligned} b_0(s,\theta) =\ & \beta(s) + Bg_0(s), \quad B_0(s,\theta,\tau) = B, \quad C_0(s,\theta,\tau) = C, \\ \gamma_0(s,\theta) =\ & \gamma(s) + Dg_0(s), \quad D_0(s,\theta,\tau) = D, \quad F_0(s,\theta,\tau) = F. \end{aligned}$$

The statement concerning the continuity of X^α is a direct consequence of the continuity of Y^α established in Lemma 4.3 and the converse direction in Theorem 2.3. □

5 A martingale verification theorem

We first derive an Itô formula for quadratic functions in $L^1(\mu)$.

Lemma 5.1. *Fix a $d \times d'$-matrix measure μ on $\mathbb{R}_+$ such that $|\mu|$ is σ-finite, and $t \in [0,T]$. Let $\widetilde{Y}$ be a $L^1(\mu)$-valued progressively measurable processes solution in the mild sense to*

$$d\tilde{Y}_s = A^{mr}\tilde{Y}_s + b_s ds + \sigma_s dW_s, \quad t \leq s \leq T, \quad \tilde{Y}_t = \xi, \tag{5.1}$$

for $\xi \in L^1(\mu)$ and some progressively measurable b, σ valued in $L^\infty(\mu)$ and satisfying

$$\int_t^T \|b_s\|_{L^\infty(\mu)}ds + \int_t^T \|\sigma_s\|^2_{L^\infty(\mu)}ds < \infty, \quad \mathbb{P}-a.s. \tag{5.2}$$

Assume that $\widetilde{Y}$ is bounded in $s \in [t,T]$, $\mathbb{P}\otimes\mu$-a.e, and has continuous sample paths in $L^1(\mu)$. Let $\widetilde{\Gamma},\widetilde{\Lambda} \in C([t,T],L^1(\mu\otimes\mu)) \times C([t,T],L^1(\mu^\top))$ be solutions to

$$\begin{aligned}\dot{\widetilde{\Gamma}}_s(\theta,\tau) &= (\theta+\tau)\widetilde{\Gamma}_s(\theta,\tau) - R^1_s(\theta,\tau), && t\le s\le T, \quad \mu\otimes\mu - a.e.\\ \dot{\widetilde{\Lambda}}_s(\theta) &= \theta\widetilde{\Lambda}_s(\theta) - R^2_s(\theta), && t\le s\le T, \quad \mu - a.e. \end{aligned} \tag{5.3}$$

with $\widetilde{\Gamma}_T \in L^\infty(\mu\otimes\mu)$, $\widetilde{\Lambda}_T \in L^\infty(\mu^\top)$, and for some measurable functions $s\mapsto R^1_s$ and R^2_s valued respectively in $L^\infty(\mu\otimes\mu)$ and $L^\infty(\mu^\top)$, such that

$$\int_t^T \|R^1_s\|_{L^\infty(\mu\otimes\mu)}ds + \int_t^T \|R^2_s\|_{L^\infty(\mu^\top)}ds < \infty. \tag{5.4}$$

Then, for all $t\le s\le T$, $\widetilde{\Gamma}_s \in L^\infty(\mu\otimes\mu)$, $\widetilde{\Lambda}_s \in L^\infty(\mu^\top)$, so that the processes

$$U^1_s = \langle \widetilde{Y}_s, \widetilde{\boldsymbol{\Gamma}}_{\boldsymbol{s}}\widetilde{Y}_s\rangle_\mu = \langle \widetilde{\boldsymbol{\Gamma}}_{\boldsymbol{s}}\widetilde{Y}_s, \widetilde{Y}_s\rangle_{\mu^\top} \qquad U^2_s = \langle \tilde{Y}_s, \tilde{\Lambda}_s\rangle_\mu = \langle \widetilde{\Lambda}_s, \widetilde{Y}_s\rangle_{\mu^\top}, \quad t\le s\le T,$$

are well defined, where $\widetilde{\boldsymbol{\Gamma}}$ is the integral operator associated to the kernel $\widetilde{\Gamma}$. Furthermore, we have for $i=1,2$,

$$dU^i_s = \Delta^i_s ds + \Sigma^i_s dW_s, \qquad t\le s\le T, \tag{5.5}$$

where

$$\begin{aligned}\Delta^1_s &= -\langle \widetilde{Y}_s, \boldsymbol{R^1_s}\widetilde{Y}_s\rangle_\mu + \langle \widetilde{\boldsymbol{\Gamma}}_{\boldsymbol{s}}\sigma_s,\sigma_s\rangle_{\mu^\top} + \langle \widetilde{\boldsymbol{\Gamma}}_{\boldsymbol{s}}\widetilde{Y}_s, b_s\rangle_{\mu^\top} + \langle \widetilde{\boldsymbol{\Gamma}}_{\boldsymbol{s}}b_s, \widetilde{Y}_s\rangle_{\mu^\top},\\ \Sigma^1_s &= \langle \widetilde{\boldsymbol{\Gamma}}_{\boldsymbol{s}}\widetilde{Y}_s,\sigma_s\rangle_{\mu^\top} + \langle \widetilde{\boldsymbol{\Gamma}}_{\boldsymbol{s}}\sigma_s, \widetilde{Y}_s\rangle_{\mu^\top},\\ \Delta^2_s &= -\langle \widetilde{Y}_s, R^2_s\rangle_\mu + \langle \widetilde{\Lambda}_s, b_s\rangle_{\mu^\top}, \qquad \Sigma^2_s = \langle \widetilde{\Lambda}_s,\sigma_s\rangle_{\mu^\top},\end{aligned}$$

where $\boldsymbol{R^1}$ is the integral operator associated to the kernel R^1.

Proof. We illustrate the proof only for U^2, that of U^1 follows along the same lines. The idea is to apply Itô's formula θ by θ. By virtue of the inequality $|\mu(B)|\le|\mu|(B)$, for any Borel set B, and the σ-finiteness of $|\mu|$, an application of the Radon–Nikodym theorem yields the existence of a measurable function $h:\mathbb{R}_+\to\mathbb{R}^{d\times d'}$ such that

$$\mu(d\theta) = h(\theta)|\mu|(d\theta) \tag{5.6}$$

with $|h(\theta)|=1$, for all $\theta\in\mathbb{R}^d$, and $|h|\in L^1(|\mu|)$, see for instance [16, Lemma 3.5.9]. Recall that by definition of a mild solution to (5.1), we mean that

$$\begin{aligned}\widetilde{Y}_s(\theta) = e^{-\theta(s-t)}\xi(\theta) &+ \int_t^s e^{-\theta(s-u)}b_u(\theta)du\\ &+ \int_t^s e^{-\theta(s-u)}\sigma_u(\theta)dW_u, \quad t\le s\le T, \quad \mathbb{P}\otimes\mu - a.e.\end{aligned} \tag{5.7}$$

Fix $t \leq T$, and observe that the solution $\widetilde{\Lambda}$ to (5.3) is given by

$$\widetilde{\Lambda}_s(\theta) = e^{-\theta(T-s)}\widetilde{\Lambda}_T(\theta) + \int_s^T e^{-\theta(u-t)} R_u^2(\theta) du, \quad t \leq s \leq T, \quad \mu - a.e.,$$

which implies, with (5.4), that $\widetilde{\Lambda}_s \in L^\infty(\mu^\top)$, $t \leq s \leq T$, together with the boundedness of $s \mapsto \widetilde{\Lambda}_s$:

$$\sup_{t \leq s \leq T} \left|\widetilde{\Lambda}_s(\theta)\right| \leq \left\|\widetilde{\Lambda}_T\right\|_{L^\infty(\mu^\top)} + \int_t^T \|R_u^2\|_{L^\infty(\mu^\top)} du \; < \; \infty, \quad \mu - a.e. \tag{5.8}$$

Moreover, since $\widetilde{Y}$ is bounded in s, we have

$$\sup_{t \leq s \leq T} |\widetilde{Y}_s(\theta)| < \infty, \quad \mathbb{P} \otimes \mu - a.e. \tag{5.9}$$

Define the $\mathbb{P} \otimes \mu$-nullset

$$\mathcal{N} = \{(\omega, \theta) : \text{such that either (5.7) or (5.3) or (5.8) or (5.9) does not hold}\}.$$

Let $(\omega, \theta) \in \Omega \times \mathbb{R}_+ \setminus \mathcal{N}$ and observe that $s \mapsto \widetilde{Y}_s(\theta, \omega)$ and $s \mapsto \widetilde{\Lambda}_s(\theta)$ solve:

$$\begin{aligned} d\widetilde{Y}_s(\theta, \omega) &= \Big(-\theta \widetilde{Y}_s(\theta, \omega) + b_s(\theta, \omega)\Big) ds + \sigma_s(\theta, \omega) dW_s, \\ d\widetilde{\Lambda}_s(\theta) &= \Big(\theta \widetilde{\Lambda}_s(\theta) - R_s^2(\theta)\Big) ds. \end{aligned}$$

An application of Itô's formula to the process $u^2(\theta, \omega) : s \mapsto \widetilde{\Lambda}_s(\theta)^\top h(\theta) \widetilde{Y}_s(\theta)$ gives

$$u_s^2(\theta, \omega) = u_t^2(\theta, \omega) + \int_t^s \delta_u^2(\theta, \omega) du + \int_t^s \widetilde{\Lambda}_u(\theta)^\top h(\theta) \sigma_u(\theta, \omega) dW_u, \quad t \leq s, \tag{5.10}$$

with

$$\delta_u^2(\theta, \omega) \; = \; -R_u^2(\theta)^\top h(\theta) \widetilde{Y}_u(\theta, \omega) + \widetilde{\Lambda}_u(\theta)^\top h(\theta) b_u(\theta, \omega).$$

All the quantities appearing on the right hand side of (5.10) are well-defined thanks to the integrability assumptions (5.2)-(5.4) on the coefficients (b, σ, R^2) and the boundedness in s of $(\widetilde{\Lambda}_s, \widetilde{Y}_s)$ from (5.8)-(5.9). Whence, (5.10) holds $\mathbb{P} \otimes \mu$ almost everywhere. Next, by the boundedness (resulting from the continuity) of $s \mapsto \widetilde{Y}_s$ in $L^1(\mu)$, $s \mapsto \widetilde{\Lambda}_s$ in $L^1(\mu^\top)$, and again by the integrability conditions (5.2)-(5.4) on (b, σ, R^2), all the terms appearing in (5.10) are in $L^1(|\mu|)$ so that an integration with respect to the θ variable against $|\mu|$ combined with the identity (5.6) and the stochastic Fubini's theorem, see [26, Theorem 2.2], lead to (5.5). □

The next theorem establishes a martingale verification result for the possibly non-Markovian optimization problem (2.1)-(2.7).

Theorem 5.2. *Let β, γ be bounded functions on $[0,T]$, g_0 continuous and $N \in \mathbb{S}^m_+$. Fix K, μ as in (2.2)-(2.3). Assume that:*

(i) *there exists a solution $(\Gamma, \Lambda, \chi) \in C([0,T], L^1(\mu \otimes \mu)) \times C([0,T], L^1(\mu^\top)) \times C([0,T], \mathbb{R})$ such that (3.7), (3.8), (3.9) hold, and $\Gamma_t \in \mathbb{S}^d_+(\mu \otimes \mu)$, for all $t \leq T$, together with the estimate (3.11),*

(ii) *there exists an admissible control process $\alpha^* \in \mathcal{A}$ such that (3.12) holds for all $t \leq T$.*

Then, α^* is an optimal control, Y^{α^*} given by (2.10) is the optimally controlled trajectory of the state variable and V^{α^*} given by (3.1) is the optimal value process of the problem, in the sense that (3.13) holds, for all $t \leq T$.

Proof. Step 1. For any $\alpha \in \mathcal{A}$, we know from Theorem 4.4 that there exists a continuous solution X^α to (2.1) such that (2.4) holds. Let us then define the continuous process

$$M^\alpha_t = \int_0^t f(X^\alpha_s, \alpha_s) ds + V^\alpha_t - \int_0^t (\alpha_s - \alpha^*_s)^\top \hat{N}_s (\alpha_s - \alpha^*_s) ds,$$

where f is the quadratic function in (2.6), V^α is the process given by (3.1) from the solution (Γ, Λ, χ) to the Riccati equation (3.2), α^* is given by (3.12), and $\hat{N}$ by (3.6), recall Remark 3.6. The main point is to check that M^α is martingale for any $\alpha \in \mathcal{A}$. Indeed, if this the case, then, for each $t \leq T$, the equality $\mathbb{E}[M^\alpha_T | \mathcal{F}_t] = M^\alpha_t$ leads to

$$J_t(\alpha) - V^\alpha_t = \mathbb{E}\left[\int_t^T (\alpha_s - \alpha^*_s)^\top \hat{N}_s (\alpha_s - \alpha^*_s) ds \;\middle|\; \mathcal{F}_t \right], \tag{5.11}$$

where we have set $J_t(\alpha) = \mathbb{E}\left[\int_t^T f(X^\alpha_s, \alpha_s) ds \;\middle|\; \mathcal{F}_t \right]$, and used $V^\alpha_T = 0$, due to the vanishing terminal conditions of $(\Gamma, \Lambda, \Theta)$ and the continuity of V^α. Since $N \in \mathbb{S}^m_+$ and Γ is $\mathbb{S}^d_+(\mu \otimes \mu)$ valued, then $\hat{N} \in \mathbb{S}^m_+$ so that the right hand side of (5.11) is always nonnegative and vanishes for $\alpha = \alpha^*$. Fix now $t \leq T$, and observe that $V^{\alpha^*}_t = V^{\alpha'}_t$ for all $\alpha' \in \mathcal{A}_t(\alpha^*)$. We then deduce from (5.11) that

$$V^{\alpha^*}_t = J_t(\alpha^*) \;=\; \inf_{\alpha' \in \mathcal{A}_t(\alpha^*)} J_t(\alpha'),$$

which is the relation (3.13), and shows that α^* is an optimal control.

Step 2. We now prove that M^α is indeed a martingale by means of Itô's formula. To ease notations, we drop the superscript α from X^α and Y^α. The process V^α is written as

$$V^\alpha_t = U^1_t + 2U^2_t + \chi_t, \quad 0 \leq t \leq T,$$

where

$$U^1_t = \langle Y_t, \mathbf{\Gamma_t} Y_t \rangle_\mu, \qquad U^2_t \;=\; \langle \Lambda_t, Y_t \rangle_{\mu^\top},$$

and we recall that Y is bounded in (t,θ) from (2.14), and has continuous sample path in $L^1(\mu)$ by Theorem 2.3. From (2.14), and the admissibility condition on $\alpha \in \mathcal{A}$ combined with Cauchy-Schwarz inequality, it is clear that the drift $b_t := \tilde{\beta}(t) + \boldsymbol{B}Y_t + C\alpha_t$, and the diffusion coefficient $\sigma_t := \tilde{\gamma}(t) + \boldsymbol{D}Y_t + F\alpha_t$ of Y take values in $L^\infty(\mu)$, and satisfy the integrability condition (5.2). Moreover, from (3.11), and since $t \in [0,T] \mapsto \Gamma_t$, Λ_t are bounded (by continuity) in $L^1(\mu\otimes\mu)$ and $L^1(\mu^\top)$, we see that $t \mapsto R_t^1 := \mathcal{R}_1(\Gamma_t)$ is valued in $L^\infty(\mu\otimes\mu)$, $t \mapsto R_t^2 := \mathcal{R}_2(t,\Gamma_t,\Lambda_t)$ is valued in $L^\infty(\mu^\top)$, and satisfy (see Remark 3.5):

$$\sup_{t\le T}\big[\|R_t^1\|_{L^\infty(\mu\otimes\mu)} + \|R_t^2\|_{L^\infty(\mu^\top)}\big] < \infty,$$

which clearly implies the integrability condition (5.4). We can then apply Lemma 5.1 on U^1, U^2. Recalling that Γ is symmetric, this yields, after re-arranging the quadratic and linear terms in Y and α, using the equation for Γ in (3.2) and applying Fubini's theorem:

$$\begin{aligned}
dU_t^1 =& \int_{\mathbb{R}_+^2} Y_t(\theta)^\top \mu(d\theta)^\top \Big(S(\Gamma_t)(\theta)\hat{N}(\Gamma_t)^{-1}S(\Gamma_t)(\tau) - Q\Big)\mu(d\tau)Y_t(\tau)dt \\
&+ \int_{\mathbb{R}_+} 2Y_t(\theta)^\top \mu(d\theta)^\top \left(\int_{\mathbb{R}_+} \Gamma_t(\theta,\tau')\mu(d\tau')\tilde{\beta}(t) + D^\top \int_{\mathbb{R}_+^2} \mu(d\theta')^\top \Gamma_t(\theta',\tau')\mu(d\tau')\tilde{\gamma}(t)\right) dt \\
&+ \alpha_t^\top F^\top \int_{\mathbb{R}_+^2} \mu(d\theta)^\top \Gamma_t(\theta,\tau)\mu(d\tau)F\alpha_t dt \\
&+ 2\alpha_t^\top \left(\int_{\mathbb{R}_+} S(\Gamma_t)(\tau)\mu(d\tau)Y_t(\tau) + F^\top \int_{\mathbb{R}_+^2} \mu(d\tau)^\top \Gamma_t(\theta,\tau)\mu(d\tau)\tilde{\gamma}(t)\right) dt \\
&+ \left(\tilde{\gamma}(t)^\top \int_{\mathbb{R}_+^2} \mu(d\theta')^\top \Gamma_t(\theta',\tau')\mu(d\tau')\tilde{\gamma}(t)\right)dt \;+\; H_t^1 dW_t,
\end{aligned}$$

with

$$H_t^1 = 2\sigma(t,X_t,\alpha_t)^\top \int_{\mathbb{R}_+^2} \mu(d\theta)^\top \Gamma_t(\theta,\tau)\mu(d\tau)Y_t(\tau).$$

Similarly, using the equation for Λ in (3.2) we get

$$\begin{aligned}
dU_t^2 =& \int_{\mathbb{R}_+} Y_t(\theta)^\top \mu(d\theta)^\top \Big(S(\Gamma_t)(\theta)^\top \hat{N}(\Gamma_t)^{-1}h(t,\Gamma_t,\Lambda_t) - L\Big)dt \\
&+ \alpha_t^\top \left(C^\top \int_{\mathbb{R}_+} \mu(d\theta')^\top \Lambda_t(\theta')\right)dt + \tilde{\beta}(t)^\top \int_{\mathbb{R}_+} \mu(d\theta')^\top \Lambda_t(\theta')dt \;+\; H_t^2 dW_t,
\end{aligned}$$

where

$$H_t^2 = \sigma(t,X_t,\alpha_t)^\top \int_{\mathbb{R}_+} \mu(d\theta)^\top \Lambda_t(\theta).$$

Now we write that

$$dM_t^\alpha = \left(X_t^\top Q X_t + \alpha_t^\top N \alpha_t - (\alpha_t - \alpha_t^*)^\top \hat{N}_t (\alpha_t - \alpha_t^*) + L^\top X_t + \dot{\chi}_t \right) dt + dU_t^1 + 2dU_t^2.$$

Completing the squares for the terms in α, observing that

$$X_t^\top Q X_t = \int_{\mathbb{R}_+^2} Y_t(\theta)^\top \mu(d\theta)^\top Q \mu(d\tau) Y_t(\tau) + 2 g_0(t)^\top \int_{\mathbb{R}_+} \mu(d\theta) Y_t(\theta) + g_0(t)^\top Q g_0(t),$$
$$L^\top X_t = L^\top \int_{\mathbb{R}_+} \mu(d\theta) Y_t(\theta) + L^\top g_0(t),$$

using the equation for χ in (3.2), and adding all the above makes the drift in M^α vanish so that

$$dM_t^\alpha = \left(H_t^1 + 2H_t^2 \right) dW_t.$$

This shows that M^α is a local martingale. To argue true martingality, successive applications of Jensen and Cauchy–Schwarz inequalities combined with the bound (3.11) yield, for a constant c,

$$\begin{aligned} \int_0^T \mathbb{E}\left[|H_s^1|^2 \right] ds &\leq cM^2 \int_0^T \mathbb{E}\left[\left(1 + |X_s|^2 + |\alpha_s|^2 \right) \|Y_s\|_{L^1(\mu)}^2 \right] ds \\ &\leq cM^2 \int_0^T \mathbb{E}\left[\left(1 + |X_s|^4 + |\alpha_s|^4 \right) \right]^{1/2} \mathbb{E}\left[\|Y_s\|_{L^1(\mu)}^4 \right]^{1/2} ds \end{aligned}$$

which is finite due to (2.4), (2.13) and the admissibility of α. Since $\Lambda \in C([0,T], L^1(\mu^\top))$, we get similarly that

$$\int_0^T \mathbb{E}\left[|H_s^2|^2 \right] ds < \infty.$$

Whence, by the Burkholder-Davis-Gundy inequality, M^α is a true martingale, and the proof is complete. □

Remark 5.3. Theorem 5.2 is still valid if one adds a linear quadratic terminal cost to the cost functional (2.5), that is

$$J(\alpha) = \mathbb{E}\left[\int_0^T f(X_s^\alpha, \alpha_s) ds + (X_T^\alpha)^\top P X_T^\alpha + 2U^\top X_T^\alpha \right],$$

provided the terminal conditions of the system of Riccati equations (3.2) are updated to

$$\Gamma_T(\theta, \tau) = P, \quad \Lambda_T(\theta) = U + P g_0(T), \quad \chi_T = g_0(T)^\top P g_0(T) + 2U^\top g_0(T).$$

The main technical difficulty resides in Assumption (i). If K has no singularities at 0, then one can still construct continuous solutions to (3.2). However, in the presence of a singularity, the solution $t \to \Gamma_t$ inherits the singularity of the kernel and is no longer continuous but only lies in $L^1([0,T], L^1(\mu \otimes \mu))$. □

6 Proof of solvability result

Proof of Theorem 3.4. First, the existence and uniqueness of a solution $\Gamma \in C([0,T], L^1(\mu\otimes\mu))$ to the Riccati equation (3.7) satisfying the estimate (3.11) and such that $\Gamma_t \in \mathbb{S}^d_+(\mu\otimes\mu)$, for all $t \le T$, follow from [5, Theorem 2.3].

Second, we note that equation (3.8) for Λ is a Lyapunov equation of the form

$$\Psi_t(\theta,\tau) = \int_t^T e^{-(\theta+\tau)(s-t)} F(s,\Psi_s)(\theta,\tau)ds, \quad t \le T, \quad \mu\otimes\mu - a.e.$$

where

$$\begin{aligned} F(s,\Psi)(\theta,\tau) = \tilde{Q}_s(\theta,\tau) + \tilde{D}^1_s(\theta)^\top \int_{\mathbb{R}^2_+} \mu_1(d\theta')^\top \Psi(\theta',\tau')\mu_2(d\tau')\tilde{D}^2_s(\tau) \\ + \tilde{B}^1_s(\theta)^\top \int_{\mathbb{R}_+} \mu_1(d\theta')^\top \Psi(\theta',\tau) + \int_{\mathbb{R}_+} \Psi(\theta,\tau')\mu_2(d\tau')\tilde{B}^2_s(\tau), \end{aligned}$$

where μ_i, $i=1,2$, $d_{i1}\times d_{i2}$-matrix valued measures on $\mathbb{R}_+$, with

$$d_{11} = d, \quad d_{12} = d', \quad d_{21} = d_{22} = 1, \quad \mu_1 = \mu, \quad \mu_2 = 0,$$

and with coefficients

$$\begin{aligned} \tilde{Q}_t(\theta,\tau) &= L + Qg_0(t) + \int_{\mathbb{R}_+} \Gamma_t(\theta,\tau')\mu(d\tau')\widetilde{\beta}_t \\ &\quad - S(\Gamma_t)(\theta)^\top \hat{N}(\Gamma_t)^{-1}F^\top \int_{\mathbb{R}^2_+} \mu(d\theta')^\top \Gamma_t(\theta',\tau')\mu(d\tau')\widetilde{\gamma}_t, \\ \tilde{B}^1_t(\theta) &= B - C\hat{N}(\Gamma_t)^{-1}S(\Gamma_t)(\theta), \qquad \tilde{B}^2_t(\theta) = \tilde{D}^1_t(\theta) = \tilde{D}^2_t(\theta) = 0. \end{aligned}$$

From [5, Theorem 3.1], we then get the existence and uniqueness of a solution $\Lambda \in C([0,T], L^1(\mu^\top))$ to the equation (3.8). Finally, we notice that the right hand side of equation (3.9) for χ, does not involve χ. Therefore, the existence and the continuity of χ follow upon simple integration, which is justified by the boundedness of g_0, β, γ, that of Λ in $L^1(\mu^\top)$ and that of Γ in $L^1(\mu\otimes\mu)$ together with the bound (3.11). □

Proof of Theorem 3.7. The result is a direct consequence of Theorem 5.2 once we prove that conditions (i)-(ii) are satisfied. Condition (i) follows from Theorem 3.4. It remains to prove condition (ii), i. e. that there exists a progressively measurable process $\alpha^* \in \mathcal{A}$ associated to a controlled SDE $Y^{\alpha^*} \in L^1(\mu)$ such that (3.12) holds, and $\sup_{t\le T} \mathbb{E}\left[|\alpha^*_t|^4\right] < \infty$. To

this end, we consider the coefficients δ, Σ as in Example (4.2) with

$$\begin{aligned} b_0(t,\theta) &= \widetilde{\beta}(t) - C\hat{N}(\Gamma_t)^{-1}h(t,\Gamma_t,\Lambda_t) \\ \gamma_0(t,\theta) &= \widetilde{\gamma}(t) - F\hat{N}(\Gamma_t)^{-1}h(t,\Gamma_t,\Lambda_t) \\ B_0(t,\theta,\theta') &= B - C\hat{N}(\Gamma_t)^{-1}S(\Gamma_t)(\theta') \\ D_0(t,\theta,\theta') &= D - F\hat{N}(\Gamma_t)^{-1}S(\Gamma_t)(\theta') \\ C_0(t,\theta) &= F_0(t,\theta) = 0. \end{aligned}$$

By the boundedness of β, γ, g_0, on $[0,T]$, the continuity hence the boundedness on $[0,T]$ of Γ in $L^1(\mu \otimes \mu)$ and Λ in $L^1(\mu^\top)$ together with the bound in Theorem 3.4, the previous coefficients satisfy (4.8)-(4.9). Whence, for any $p \geq 2$, Example 4.2, combined with Theorem 4.1, yields the existence of a process Y^* with initial condition $Y_0^* \equiv 0$, for the coefficients δ, Σ as defined above and such that (4.3) holds. One can therefore define a process α^* by

$$\alpha_t^* = -\hat{N}(\Gamma_t)^{-1}\left(h(t,\Gamma_t,\Lambda_t) + \int_{\mathbb{R}_+} S(\Gamma_t)(\theta)\mu(d\theta)Y_t^*(\theta)\right),$$

and see, again from the boundedness on $[0,T]$ of (Γ,Λ) in $L^1(\mu\otimes\mu)\times L^1(\mu^\top)$ together with the bound (3.11), that

$$\mathbb{E}\big[|\alpha_t^*|^p\big] \leq c(1+M^4)\sup_{0\leq t\leq T}\mathbb{E}\big[\|Y_t^*\|_{L^1(\mu)}^p\big], \quad 0 \leq t \leq T,$$

which is finite due to (4.3). In particular, for $p=4$, we get that α^* lies in $\mathcal{A}$.
Finally, by construction, the coefficients of Y^* can be re-written in terms of α^* as

$$\begin{aligned} \delta(t,\omega,Y_t^*) &= \widetilde{\beta}(t) + B\int_{\mathbb{R}_+}\mu(d\theta)Y_t^*(\theta) + C\alpha_t^*, \\ \Sigma(t,\omega,Y_t^*) &= \widetilde{\gamma}(t) + D\int_{\mathbb{R}_+}\mu(d\theta)Y_t^*(\theta) + F\alpha_t^*, \end{aligned}$$

which means that $Y^* = Y^{\alpha^*}$, and ends the proof. □

7 Stability and approximation

In this section, we prove Theorem 3.13 starting with *a priori* L^2–estimates for the controlled stochastic Volterra equation before approximating the value function.

7.1 A-priori L^2–estimates for stochastic Volterra equations

Let X^α be the solution produced by Theorem 4.4. We provide an explicit bound for $\mathbb{E}\left[\|X^\alpha\|^2_{L^2(0,T)}\right]$, which is finite due to (2.4). For this, let us introduce the resolvent of the second kind R of a scalar kernel k, defined as the unique $L^1([0,T],\mathbb{R})$ solution to the linear convolution equation

$$R(t) = k(t) + \int_0^t k(t-s)R(s)ds = k(t) + \int_0^t R(t-s)k(s)ds, \quad t \leq T.$$

Recall that the resolvent R exists, for any kernel $k \in L^1([0,T],\mathbb{R})$, see [16, Theorems 2.3.1 and 2.3.5].

Lemma 7.1. *Fix $K \in L^2([0,T],\mathbb{R}^{d\times d'})$, $\alpha \in \mathcal{A}$, $g_0 \in L^2([0,T],\mathbb{R}^d)$ and $\beta,\gamma \in L^2([0,T],\mathbb{R}^{d'})$. If X^α is a progressively measurable process satisfying (2.1) with*

$$\mathbb{E}\left[\|X^\alpha\|^2_{L^2(0,T)}\right] < \infty, \tag{7.1}$$

then, it holds that

$$\mathbb{E}\left[\|X^\alpha\|^2_{L^2(0,T)}\right] \leq c m_T(g_0,K,\alpha)\left(1 + \|R\|_{L^1(0,T)}\right),$$

where c is a constant only depending on (T,B,C,D,F),

$$m_T(g_0,K,\alpha) = \|g_0\|^2_{L^2(0,T)} + \|K\|^2_{L^2(0,T)}\left(\|\beta\|^2_{L^2(0,T)} + \|\gamma\|^2_{L^2(0,T)} + \mathbb{E}\big[\|\alpha\|^2_{L^2(0,T)}\big]\right)$$

and R is the resolvent of $c|K|^2$.

Proof. Throughout the proof, we make use of the notations $(f*g)(t) = \int_0^t f(t-s)g(s)ds$ and $(f*dZ)_t = \int_0^t f(t-s)dZ_s$, and c will denote a constant depending exclusively on (T,B,C,D,F) that may vary from line to line. We first observe that by Jensen's inequality

$$\begin{aligned}\|X^\alpha_t\|^2_{L^2(0,T)} &\leq 5\|g_0\|^2_{L^2(0,T)} + 5\|K*(\beta + C\alpha)\|^2_{L^2(0,T)} + 5\|K*BX^\alpha\|^2_{L^2(0,T)} \\ &\quad + 5\|K*DX^\alpha dW\|^2_{L^2(0,T)} + 5\|K*(\gamma + F\alpha)\,dW\|^2_{L^2(0,T)} \\ &= 5(\mathbf{I}+\mathbf{II}+\mathbf{III}+\mathbf{IV}+\mathbf{V}).\end{aligned}$$

An application of Young and Cauchy–Schwarz inequalities yields

$$\mathbf{II} \leq c\|K\|^2_{L^2(0,T)}\left(\|\beta\|^2_{L^2(0,T)} + \|\alpha\|^2_{L^2(0,T)}\right).$$

Successive applications of Cauchy–Schwarz, Tonelli's theorem and changes of variables lead to

$$\begin{aligned}
\mathbf{III} &\leq c \int_0^T \int_0^t |K(t-s)|^2 |X_s^\alpha|^2 ds dt \\
&= c \int_0^T \int_0^t |K(s)|^2 |X_{t-s}^\alpha|^2 ds dt \\
&= c \int_0^T |K(s)|^2 \int_s^T |X_{t-s}^\alpha|^2 dt ds \\
&= c \int_0^T |K(T-s)|^2 \int_0^s |X_u^\alpha|^2 du ds \\
&= c \int_0^T |K(T-s)|^2 \|X^\alpha\|_{L^2(0,s)}^2 ds.
\end{aligned}$$

Taking the expectation, we get

$$\begin{aligned}
\mathbb{E}[\mathbf{II} + \mathbf{III}] \leq c\|K\|_{L^2(0,T)}^2 \left(\|\beta\|_{L^2(0,T)}^2 + \mathbb{E}\left[\|\alpha\|_{L^2(0,T)}^2 \right] \right) \\
+ c \int_0^T |K(T-s)|^2 \mathbb{E}\left[\|X^\alpha\|_{L^2(0,s)}^2 \right] ds.
\end{aligned}$$

Similarly, Itô's isometry combined with Tonelli's theorem and multiple changes of variables give

$$\begin{aligned}
\mathbb{E}[\mathbf{IV}] &= |D|^2 \mathbb{E}\left[\int_0^T \int_0^t |K(t-s)|^2 |X_s^\alpha|^2 ds dt \right] \\
&= |D|^2 \int_0^T |K(T-s)|^2 \mathbb{E}\left[\|X^\alpha\|_{L^2(0,s)}^2 \right] ds.
\end{aligned}$$

Another application of Itô's isometry and Young's inequality shows that

$$\begin{aligned}
\mathbb{E}[\mathbf{V}] &\leq c\| |K|^2 * \left(|\gamma|^2 + |F|^2 |\alpha|^2 \right) \|_{L^1(0,T)}^2 \\
&\leq c\|K\|_{L^2(0,T)}^2 \left(\|\gamma\|_{L^2(0,T)}^2 + \mathbb{E}\left[\|\alpha\|_{L^2(0,T)}^2 \right] \right).
\end{aligned}$$

Combining the above yields

$$\begin{aligned}
\mathbb{E}\left[\|X^\alpha\|_{L^2(0,T)}^2 \right] &\leq c\|g_0\|_{L^2(0,T)}^2 + c\|K\|_{L^2(0,T)}^2 \left(\|\beta\|_{L^2(0,T)}^2 + \|\gamma\|_{L^2(0,T)}^2 + \mathbb{E}\left[\|\alpha\|_{L^2(0,T)}^2 \right] \right) \\
&\quad + c \int_0^T |K(T-s)|^2 \mathbb{E}\left[\|X^\alpha\|_{L^2(0,s)}^2 \right] ds, \\
&\leq c m_T(g_0, K, \alpha) \left(1 + \|R\|_{L^1(0,T)} \right)
\end{aligned}$$

where the last line follows from the generalized Gronwall inequality for convolution equations with R the resolvent of $c|K|^2$, see [16, Theorem 9.8.2]. □

Lemma 7.2. *Fix $n \in \mathbb{N}$. Let $K, K^n \in L^2([0,T], \mathbb{R}^{d\times d'})$, $\alpha \in \mathcal{A}$, $g_0, g_0^n \in L^2([0,T], \mathbb{R}^d)$ and $\beta, \gamma \in L^2([0,T], \mathbb{R}^{d'})$. Assume that there exist two progressively measurable processes X and X^n satisfying (2.1) and (7.1) for the respective inputs (g_0, K, α) and (g_0^n, K^n, α). Then,*

$$\mathbb{E}\left[\|X^n - X\|^2_{L^2(0,T)}\right] \leq c m_n \left(1 + \|R^n\|_{L^1(0,T)}\right), \tag{7.2}$$

where

$$m_n = \|g_0^n - g_0\|^2_{L^2(0,T)} + \|K^n - K\|^2_{L^2(0,T)} \left(\mathbb{E}\left[\|X\|^2_{L^2(0,T)}\right] + \mathbb{E}\left[\|\alpha\|^2_{L^2(0,T)}\right]\right)$$

and R^n is the resolvent of $c|K^n|^2$. If in addition

$$\|g_0^n - g_0\|_{L^2(0,T)} \to 0, \quad \|K^n - K\|_{L^2(0,T)} \to 0, \tag{7.3}$$

as $n \to \infty$, then,

$$\mathbb{E}\left[\|X^n - X\|^2_{L^2(0,T)}\right] \to 0, \quad \text{as } n \to \infty. \tag{7.4}$$

Proof. Fix $t \leq T$. We start by writing

$$\begin{aligned}
X_t - X_t^n &= (g_0(t) - g_0^n(t)) + \int_0^t (K(t-s) - K^n(t-s))\,(BX_s + C\alpha_s)\,ds \\
&\quad - \int_0^t K^n(t-s)B(X_s^n - X_s)ds \\
&\quad + \int_0^t (K(t-s) - K^n(t-s))\,(DX_s + F\alpha_s)\,dW_s \\
&\quad - \int_0^t K^n(t-s)D(X_s^n - X_s)dW_s \\
&= \mathbf{I}_t + \mathbf{II}_t + \mathbf{III}_t + \mathbf{IV}_t + \mathbf{V}_t.
\end{aligned}$$

In the sequel, c denotes a constant independent of n that may vary from line to line. Repeating the same argument as in the proof of Lemma 7.1, we get

$$\mathbb{E}\left[\|\mathbf{II}\|^2_{L^2(0,T)} + \|\mathbf{IV}\|^2_{L^2(0,T)}\right] \leq c\|K^n - K\|^2_{L^2(0,T)} \left(\mathbb{E}\left[\|X\|^2_{L^2(0,T)}\right] + \mathbb{E}\left[\|\alpha\|^2_{L^2(0,T)}\right]\right),$$

which is finite due to Lemma 7.1. Similarly,

$$\mathbb{E}\left[\|\mathbf{III}\|^2_{L^2(0,T)} + \|\mathbf{V}\|^2_{L^2(0,T)}\right] \leq c \int_0^T |K^n(T-s)|^2 \mathbb{E}\left[\|X^n - X\|^2_{L^2(0,s)}\right] ds.$$

Combining the above and invoking [16, Theorem 9.8.2] for the generalized Gronwall inequality for convolution equations yields the estimate (7.2). We now prove that its right

hand side goes to 0, as n goes to infinity. We first note that $R^n \to R$ in L^1, by virtue of the continuous dependence of the resolvent on the kernel combined with the L^2–convergence of $(K^n)_{n\geq 1}$ in (7.3), see [16, Lemma 9.3.11]. Consequently, the sequences $(\|R^n\|_{L^1(0,T)})_{n\geq 1}$ and $(\|K^n\|_{L^2(0,T)})_{n\geq 1}$ are uniformly bounded in n, that is

$$\sup_{n\geq 1} \|R^n\|_{L^1(0,T)} + \sup_{n\geq 1} \|K^n\|_{L^2(0,T)} < \infty. \tag{7.5}$$

Thus, it follows from (7.2) that it is enough to prove that $m_n \to 0$ to get the claimed convergence (7.4). This is straightforward from (7.3) and the proof is complete. □

7.2 Approximation of the value function

The proof of Theorem 3.13 now follows from the two following lemmas. In the sequel, we work under the assumptions of Theorem 3.13 and we recall the expressions of $J^n, X^{n,\alpha}$ in (3.21). To ease notations, we drop the α superscripts.

Lemma 7.3. *Let $\alpha \in \mathcal{A}$. Under (3.22) we have*

$$|J(\alpha) - J^n(\alpha)|^2 \leq c\left(2 + \left(\mathbb{E}\left[\|\alpha\|^2_{L^2(0,T)}\right]\right)^2\right)\left(\|g_0^n - g_0\|^2_{L^2(0,T)} + \|K^n - K\|^2_{L^2(0,T)}\right),$$

where c is a constant independent of n.

Proof. Fix $\alpha \in \mathcal{A}$, we start by writing

$$\begin{aligned} J(\alpha) - J^n(\alpha) &= \mathbb{E}\left[\int_0^T \left(X_s^\top Q X_s - (X_s^n)^\top Q X_s^n\right) ds\right] + \mathbb{E}\left[\int_0^T (X_s - X_s^n)^\top L ds\right] \\ &= \mathbb{E}\left[\int_0^T (X_s - X_s^n)^\top Q (X_s + X_s^n) ds\right] + \mathbb{E}\left[\int_0^T (X_s - X_s^n)^\top L ds\right] \\ &= \mathbf{I} + \mathbf{II} \end{aligned}$$

so that

$$|J(\alpha) - J^n(\alpha)|^2 \leq 2\left(\mathbf{I}^2 + \mathbf{II}^2\right).$$

We let c denote a constant independent of n that may vary from line to line. Successive applications of Cauchy-Schwarz inequality and Lemma 7.2 yield

$$\begin{aligned} \mathbf{II}^2 &\leq c\mathbb{E}\left[\|X - X^n\|^2_{L^2(0,T)}\right] \\ &\leq c\left(1 + \|R^n\|_{L^1(0,T)}\right)\left(1 + \mathbb{E}\left[\|\alpha\|^2_{L^2(0,T)}\right]\right)\left(\|g_0^n - g_0\|^2_{L^2(0,T)} + \|K^n - K\|^2_{L^2(0,T)}\right) \end{aligned}$$

where R^n is the resolvent of $c|K^n|^2$. By virtue of the L^2 convergence of the kernels $(K^n)_{n\geq 1}$ in (3.22), $\|R^n\|_{L^1(0,T)}$ is uniformly bounded in n, see (7.5). Whence,

$$\mathbf{II}^2 \leq c\left(1+\mathbb{E}\left[\|\alpha\|^2_{L^2(0,T)}\right]\right)\left(\|g_0^n-g_0\|^2_{L^2(0,T)}+\|K^n-K\|^2_{L^2(0,T)}\right).$$

Similarly, we get from Lemmas 7.1 and 7.2

$$\begin{aligned}\mathbf{I}^2 &\leq c\left(\mathbb{E}\left[\|X\|^2_{L^2(0,T)}\right]+\mathbb{E}\left[\|X^n\|^2_{L^2(0,T)}\right]\right)\mathbb{E}\left[\|X-X^n\|^2_{L^2(0,T)}\right]\\ &\leq c\left(1+\left(\mathbb{E}\left[\|\alpha\|^2_{L^2(0,T)}\right]\right)^2\right)\left(\|g_0^n-g_0\|^2_{L^2(0,T)}+\|K^n-K\|^2_{L^2(0,T)}\right),\end{aligned}$$

where the last inequality follows from the fact that $\sup_{n\geq 1}\mathbb{E}\left[\|X^n\|^2_{L^2(0,T)}\right]<\infty$, since $\mathbb{E}\left[\|X^n-X\|^2_{L^2(0,T)}\right]\to 0$ from Lemma 7.2. Combining the above yields the desired estimate. □

Lemma 7.4. *Assume (3.10), (3.22) and that Q is invertible. Let α^* and α^{n*} be the optimal controls produced by Theorem 3.7 respectively for the problem (2.5) and its approximation (3.20). There exists a constant $\kappa>0$ such that*

$$\mathbb{E}\left[\|\alpha^*\|^2_{L^2(0,T)}\right]+\sup_{n\geq 1}\mathbb{E}\left[\|\alpha^{n*}\|^2_{L^2(0,T)}\right]\leq\kappa. \tag{7.6}$$

Proof. Under (3.10), there exists $c>0$ such that

$$|a|^2\leq ca^\top Na,\quad a\in\mathbb{R}^m.$$

Denoting by $X^n=X^{n,\alpha^{n*}}$, it follows that

$$\begin{aligned}\mathbb{E}\left[\|\alpha^{n*}\|^2_{L^2(0,T)}\right] &\leq (1\vee c)\mathbb{E}\left[\int_0^T\left(\left(\alpha_s^{n*}\right)^\top N\alpha_s^{n*}+\left(X_s^n+Q^{-1}L\right)^\top Q\left(X_s^n+Q^{-1}L\right)\right)ds\right]\\ &= (1\vee c)\left(J^n(\alpha^{n*})+L^\top Q^{-1}L\right)\\ &\leq (1\vee c)\left(J^n(\mathbf{0})+L^\top Q^{-1}L\right),\end{aligned}$$

for all $n\in\mathbb{N}$, where the last inequality follows from the optimality of α^{n*} and $\mathbf{0}$ corresponds to the admissible control $\alpha_s=0$, for all $s\leq T$. Applying Lemma 7.3, with $\alpha=\alpha^n=\mathbf{0}$, we obtain the convergence of the un-controlled functional cost: $\lim_{n\to\infty}J^n(\mathbf{0})=J(\mathbf{0})$, which ensures that $J^n(\mathbf{0})$ is uniformly bounded in n. We then deduce the existence of a constant κ such that (7.6) holds. □

The proof of Theorem 3.13 is now straightforward.

Proof of Theorem 3.13. Fix an arbitrary $\varepsilon > 0$, and let $c_\varepsilon > 0$ to be determined later. First note that Lemma 7.4 ensures the existence a constant $\kappa > 0$ such that for all $n \in \mathbb{N}$:

$$\begin{aligned} V_0^n &= \inf_{\alpha \in \mathcal{A}} J^n(\alpha) = \inf_{\alpha \in \mathcal{A}_\kappa} J^n(\alpha), \\ V_0 &= \inf_{\alpha \in \mathcal{A}} J(\alpha) = \inf_{\alpha \in \mathcal{A}_\kappa} J(\alpha), \end{aligned} \tag{7.7}$$

where $\mathcal{A}_\kappa = \{\alpha \in \mathcal{A} : \mathbb{E}[\|\alpha\|^2_{L^2(0,T)}] \le \kappa\}$. Under condition (3.22), there exists $n_\varepsilon \in \mathbb{N}$ such that for every $n \ge n_\varepsilon$ we have $\|g_0^n - g_0\|^2_{L^2(0,T)} + \|K^n - K\|^2_{L^2(0,T)} \le c_\varepsilon$. By Lemma 7.3, it follows that for any $\alpha \in \mathcal{A}_\kappa$, and $n \ge n_\varepsilon$,

$$|J(\alpha) - J^n(\alpha)|^2 \le c(2 + \kappa^2)c_\varepsilon \;=\; \varepsilon,$$

by choosing $c_\varepsilon = \frac{\epsilon}{c(2+\kappa^2)}$. Combined with (7.7), this gives (3.23) and also (3.24). □

A An elementary lemma

Lemma A.1. *Let K be given as in (2.2), and $\overline{K}$ defined by*

$$\overline{K}(t) = \int_{\mathbb{R}_+} e^{-\theta t} |\mu|(d\theta), \quad t > 0.$$

Assume that (2.3) holds, then

$$\int_0^T |K(s)|^2 ds \le \int_0^T \overline{K}(s)^2 ds < \infty.$$

Furthermore, $|\mu|$ is σ-finite.

Proof. Since, $|K(t)| \le \overline{K}(t)$, for all $t > 0$, it is clear that $\int_0^T |K(s)|^2 ds \le \int_0^T \overline{K}(s)^2 ds$. Furthermore,

$$\begin{aligned} \|\overline{K}\|_{L^2(0,T)} &= \left\| \int_{\mathbb{R}_+} e^{-\theta(\cdot)} |\mu|(d\theta) \right\|_{L^2(0,T)} \\ &\le \int_{\mathbb{R}_+} \left\| e^{-\theta(\cdot)} \right\|_{L^2(0,T)} |\mu|(d\theta) \;=\; \int_{\mathbb{R}_+} \sqrt{\frac{1 - e^{-2\theta T}}{2\theta}} |\mu|(d\theta) \end{aligned}$$

which is finite due to inequality (2.16) and condition (2.3). To prove that $|\mu|$ is σ-finite, we observe that $\mathbb{R}_+ = \cup_{n \in \mathbb{N}} [0, n]$ such that for each $n \ge 1$,

$$\begin{aligned} |\mu|([0,n]) \;=\; \int_0^1 |\mu|(d\theta) + \int_1^n |\mu|(d\theta) &\le \int_0^1 |\mu|(d\theta) + \sqrt{n} \int_1^n \theta^{-1/2} |\mu|(d\theta) \\ &\le \sqrt{n} \int_{\mathbb{R}_+} (1 \wedge \theta^{-1/2}) |\mu|(d\theta) < \infty. \end{aligned}$$

□

References

[1] E. Abi Jaber. Lifting the Heston model. *Quantitative Finance*, pages 1–19, 2019.

[2] E. Abi Jaber and O. El Euch. Markovian structure of the Volterra Heston model. *Statistics & Probability Letters*, 149:63–72, 2019.

[3] E. Abi Jaber and O. El Euch. Multifactor approximation of rough volatility models. *SIAM Journal on Financial Mathematics*, 10(2):309–349, 2019.

[4] E. Abi Jaber, M. Larsson, and S. Pulido. Affine Volterra processes. *The Annals of Applied Probability*, 29(5):3155–3200, 2019.

[5] E. Abi Jaber, E. Miller, and H. Pham. Integral operator Riccati equations arising in stochastic Volterra control problems. 2019.

[6] N. Agram and B. Øksendal. Malliavin calculus and optimal control of stochastic Volterra equations. *Journal of Optimization Theory and Applications*, 167(3):1070–1094, 2015.

[7] A. Alfonsi and A. Schied. Capacitary measures for completely monotone kernels via singular control. *SIAM Journal on Control and Optimization*, 51(2):1758–1780, 2013.

[8] O. Barndorff-Nielsen, F. Benth, and A. Veraart. Ambit processes and stochastic partial differential equations. In *Advances mathematical methods for finance, eds G. Di Nunno and B. Oksendal*, pages 35–74, 2011.

[9] S. Bonaccorsi, F. Confortola, and E. Mastrogiacomo. Optimal control for stochastic Volterra equations with completely monotone kernels. *SIAM Journal on Control and Optimization*, 50(2):748–789, 2012.

[10] P. Carmona and L. Coutin. Fractional Brownian motion and the Markov property. *Electronic Communications in Probability*, 3:95–107, 1998.

[11] C. Cuchiero and J. Teichmann. Generalized Feller processes and Markovian lifts of stochastic Volterra processes: the affine case. *arXiv preprint arXiv:1804.10450*, 2018.

[12] T. E. Duncan and B. Pasik-Duncan. Linear-quadratic fractional Gaussian control. *SIAM Journal on Control and Optimization*, 51(6):4504–4519, 2013.

[13] O. El Euch and M. Rosenbaum. Perfect hedging in rough Heston models. *The Annals of Applied Probability*, 28(6):3813–3856, 2018.

[14] F. Flandoli. Direct solution of a Riccati equation arising in a stochastic control problem with control and observation on the boundary. *Applied Mathematics and Optimization*, 14(1):107–129, 1986.

[15] J. Gatheral, T. Jaisson, and M. Rosenbaum. Volatility is rough. *Quantitative Finance*, 18(6):933–949, 2018.

[16] G. Gripenberg, S.-O. Londen, and O. Staffans. *Volterra integral and functional equations*, volume 34 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1990.

[17] B. Han and H. Y. Wong. Time-consistent feedback strategies with Volterra processes. *arXiv preprint arXiv:1907.11378*, 2019.

[18] P. Harms and D. Stefanovits. Affine representations of fractional processes with applications in mathematical finance. *Stochastic Processes and their Applications*, 129(4):1185–1228, 2019.

[19] Y. Hu and S. Tang. Stochastic LQ and associated Riccati equation of PDEs driven by state-and control-dependent White noise. *arXiv preprint arXiv:1809.05308*, 2018.

[20] A. Jacquier and M. Oumgari. Deep PPDEs for rough local stochastic volatility. 2019.

[21] M. Kleptsyna, A. L. Breton, and M. Viot. About the linear quadratic regulator problem under a fractional Brownian perturbation,. *ESAIM Probab. Stat*, 9:161–170, 2003.

[22] L. Mytnik and T. S. Salisbury. Uniqueness for Volterra-type stochastic integral equations. *arXiv preprint arXiv:1502.05513*, 2015.

[23] E. Pardoux and P. Protter. Stochastic Volterra equations with anticipating coefficients. *Ann. Probab.*, 18(4):1635–1655, 1990.

[24] W. Rudin. *Real and complex analysis*. Tata McGraw-hill education, 2006.

[25] J. Schmiegel. Self-scaling tumor growth. *Physica A: Statistical Mechanics and its Applications,*, 367(C):509–524, 2006.

[26] M. Veraar. The stochastic Fubini theorem revisited. *Stochastics*, 84(4):543–551, 2012.

[27] F. Viens and J. Zhang. A martingale approach for fractional Brownian motions and related path dependent PDEs. *Annals of Applied Probability, to appear*, 2018.

[28] T. Wang. Linear quadratic control problems of stochastic Volterra integral equations. *ESAIM: Control, Optimisation and Calculus of Variations*, 24:1849–1879, 2018.

[29] J. Yong. Backward stochastic Volterra integral equations and some related problems. *Stochastic Processes and their Applications*, 116(5):779–795, 2006.

[30] J. Yong and X. Y. Zhou. *Stochastic controls: Hamiltonian systems and HJB equations*, volume 43. Springer Verlag, 1999.