Linear-Quadratic Stochastic Delayed Control and Deep Learning Resolution

William Lefebvre\textsuperscript{1,2} · Enzo Miller\textsuperscript{2}

Received: 12 March 2021 / Accepted: 11 August 2021 / Published online: 23 September 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract
We consider a simple class of stochastic control problems with a delayed control, in both the drift and the diffusion part of the state stochastic differential equation. We provide a new characterization of the solution in terms of a set of Riccati partial differential equations. Existence and uniqueness of a solution are obtained under a sufficient condition expressed directly as a relation between the time horizon, the drift, the volatility and the delay. Furthermore, a deep learning scheme (The code is available in a IPython notebook.) is designed and used to illustrate the effect of the delay feature on the Markowitz portfolio allocation problem with execution delay.

Keywords Linear-quadratic stochastic control · Delay · Riccati PDEs · Markowitz portfolio allocation · Deep learning

Mathematics Subject Classification 93E20 · 60H10 · 34K50

1 Introduction

The control of systems whose dynamic contains delays on the state and/or control has attracted the attention of the optimization and engineering communities in the last decades due to its wide variety of applications, allowing to tackle problems where the past of a system influences its present or where an agent controls a system with
a latency. As a non-exhaustive list of applications, we may cite the following papers, classified by their applications domain: engineering [15,26]; advertising [11,12,20,23]; Learning by doing with memory effect [7]; growth model with lags between investment decision and project completion [2–4,13,17]; investment [18,27]. More recently, the introduction of delayed control together with mean-field effects was studied [6,10] and new machine learning methods have been designed to numerically solve stochastic control problems with delay [14]. We also refer to the monograph [24] to find literature on the various effects of delays on traffic flow modeling, chemical processes, population dynamics, supply chain, etc.

In the optimal control community, two main approaches have emerged: the structural state method and the extended state method, and we refer to [5, Part II, Chapter 3] for the study of the latter in the deterministic case and [9] for the structural state approach in the stochastic case. Let us also mention the paper [9] for an overview and exhaustive list of references.

In this paper, we aim at studying the challenging case where there is a delayed control both in the drift and volatility. Except in [9], this situation is not treated theoretically nor numerically in the references above. The main difficulty comes from the fact that the natural formulation of a control problem with delayed control involves a boundary control problem. Indeed, assume for instance that \( X \) denotes a state variable following the simple dynamic \( \dot{X}_t = \alpha_{t-d} \), where \( \alpha \) denotes the control. For any time \( t \) and index \( s \in [-d, 0] \), set \( u_t(s) = \alpha_{t+s} \), the memory of the control \( \alpha \). Then, note that \( \partial_t u_t(s) = \partial_s u_t(s) \) and \( u_t(0) = \alpha_t \). Thus, the natural infinite-dimensional formulation of the controlled system is

\[
\begin{align*}
\text{State eq. on } (X, u) & \quad \left\{ \begin{array}{l}
\dot{X}_t = \mathcal{M}u_t \\
(\partial_t - \partial_s)u_t(s) = 0,
\end{array} \right.
\text{Boundary constraint} & \quad \{ B u_t = \alpha_t, \\
\text{Initial conditions} & \quad \{ u_0(s) = \gamma_s, \quad X_0 = x,
\end{align*}
\]

where \( t, s \in [0, T] \times [-d, 0] \), \( \mathcal{M}u := u(-d) \), \( Bu := u(0) \) and \( \gamma \) is the initial value of the control over \([-d, 0]\). Consequently, any delayed controlled problem where the delay appears in the control variable can be recast as a boundary control problem whose geometry is parametrized by the delay \( d \). More details concerning this geometry will be given in Sect. 2.

**Main contributions.** Our goal is to shed some lights on the difficulty related to delayed control on the volatility and to provide a practical and simple tool for designing a numerical scheme practitioners can use. In this paper, we study the most simple linear-quadratic control problem with delayed control both in drift and volatility. The optimal feedback control and the value function are given in terms of Riccati partial differential equations and the extended state \( (x, u) \in H = \mathbb{R} \times L^2([0, T], \mathbb{R}) \), where \( x \) denotes the position and \( s \in [-d, 0] \mapsto u(s) \) the memory of the control. The existence and uniqueness of these are proven under a condition, emerging from the delay feature, involving the drift \( b \), the volatility \( \sigma \), the delay \( d \) and the horizon \( T \). Finally, we adopt a deep learning approach in the spirit of the papers [22] (Physics Informed
Neural Network) and [25] (Deep Galerkin) to propose a numerical scheme. Our results are illustrated on the celebrated Markowitz portfolio allocation problem where we take into account execution delay. We believe the semi-explicit resolution of infinite-dimensional control problems by means of deep learning method will open the door to several interesting applications such as quick simulations of richer models and precise benchmarking of reinforcement learning algorithms.

**Outline of the paper.** The rest of the paper is organized as follows: In Sect. 2, we formulate the stochastic delayed control problem and derive an heuristic approach through a lifting in an infinite-dimensional space, namely the extended state space in the spirit of [16], but without the use of semi-group theory. We state in Sect. 3 a verification theorem and prove existence and uniqueness results for the Riccati PDEs. A deep learning-based numerical scheme with two applications on Markowitz portfolio allocation is given in Sect. 4, with a detailed analysis of the effect of the delay feature on the allocation strategy.

**Notations.**
Given a probability space \((\Omega, \mathcal{F}, \mathbb{P})\), a filtration \(\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}\) satisfying the usual conditions and \(a < b\) two real numbers, we denote by

\[
L^2([a, b], \mathbb{R}) = \left\{ Y : [a, b] \mapsto \mathbb{R}, \text{ s.t. } \int_a^b |Y_t|^2 \, dt < \infty \right\},
\]

\[
L^2_{\mathbb{F}}([a, b], \mathbb{R}) = \left\{ Y : \Omega \times [a, b] \mapsto \mathbb{R}, \mathbb{F} - \text{ prog. measurable s.t. } \times \mathbb{E}\left[ \int_a^b |Y_t|^2 \, dt \right] < \infty \right\}.
\]

Here \(| \cdot |\) denotes the Euclidean norm on \(\mathbb{R}\) or \(\mathbb{R}^d\), and \(H = \mathbb{R} \times L^2([0, T], \mathbb{R})\) denotes the extended state space endowed with the scalar product \(\langle x, y \rangle_H = x_0 y_0 + \int_0^T x_1(s) y_1(s) \, ds\). For any \(z = (x, u) \in H\), we use the notation \(z_0 = x\) and \(z_1 = u\).

### 2 Formulation of the Problem and Heuristic Approach

Let \((\Omega, \mathcal{F}, \mathbb{P} := (\mathcal{F}_t)_{t \leq 0}, \mathbb{P})\) be a complete filtered probability space on which a real-valued Brownian motion \((W_t)_{t \leq 0}\) is defined and consider the simple system defined on \([0, T]\) by the following dynamics

\[
\begin{dcases}
    dX_t^\alpha = \alpha_{t-d} (b dt + \sigma dW_t), & 0 \leq t \leq T, \\
    X_0 = x, & \alpha_s = \gamma(s), \quad s \in [-d, 0].
\end{dcases}
\] (2.1)

endowed with the cost functional

\[
J(\alpha) = \mathbb{E}\left[ (X_T^\alpha)^2 \right],
\] (2.2)
where \( \gamma \in L^2([-d, 0], \mathbb{R}) \) and \( \alpha \) models the control chosen in the set of admissible strategies \( \mathcal{A} \):

\[
\mathcal{A} = \left\{ \alpha \in L^2_{\mathbb{F}}([0, T], \mathbb{R}) \text{ such that (2.1) has a solution satisfying} \right. \\
\left. \times \mathbb{E} \left[ \sup_{t \leq T} |X_t^\alpha|^2 \right] < \infty \right\}.
\]

For any \( 0 \leq a < b \leq T \), we also define the set \( \mathcal{A}_{a,b} \) as the restriction of \( \mathcal{A} \) to \( L^2_{\mathbb{F}}([a, b], \mathbb{R}) \).

**Remark 2.1** At this point, we may expect a priori that the optimization problem (2.1)–(2.2) admits an optimizer provided \( \sigma \neq 0 \), even if the control is not directly penalized. The intuition behind this a priori belief is that the more \( \alpha \) is aggressive in bringing \( X \) to 0, the more the variance of \( X \) increases due to the diffusion term. It is the case in the classical LQ stochastic optimization problem with controlled volatility such as

\[
dX_t^\alpha = \alpha_t (b dt + \sigma dW_t), \quad t \leq T,
\]

\[X_0 = x,
\]

\[J(\alpha) = \mathbb{E}[(X_T^\alpha)^2],
\]

where the optimal control reads \( \alpha_t^* = -\frac{b}{\sigma^2} X_t^\alpha \) and the value function \( V_t = e^{(t-T)\frac{\sigma^2}{\delta}} \).

A surprising finding in our paper is the necessity, due to the delay feature, for a more restricting condition on the diffusion coefficients to show the existence of an optimizer. See Proposition 3.3.

**Remark 2.2** In the rest of the paper, we focus on the one-dimensional case with delayed control both in drift and volatility which features the main difficulties related to the presence of the delay. Although Proposition 3.3 concerning the existence and uniqueness of a Riccati-PDE system does not directly extend to the multidimensional case, the verification theorem can easily be adapted to the multidimensional case with delayed state and control.

The first step consists in lifting the dynamics in the infinite-dimensional Hilbert space \( H = \mathbb{R} \times L^2([0, T], \mathbb{R}) \), where the system is naturally Markovian. To do so, denote \( u_t(s) = \alpha_{t+s} \) for any \( t \leq T \) and \( s \in [-d, 0] \), a transport of the control. The dynamics (2.1) then reads

\[
\begin{align*}
\begin{cases}
\text{d}Z_t^\alpha = AZ_t^\alpha dt + BZ_t^\alpha dW_t + C \text{d}\alpha_t, & 0 \leq t \leq T, \\
Z_0 = (x, \gamma),
\end{cases}
\end{align*}
\]

where \( Z_t^\alpha \) is defined as the \( H = \mathbb{R} \times L^2([0, T], \mathbb{R}) \)-valued random process \( Z_t^\alpha = (X_t^\alpha, u_t(\cdot)) \) and

\[
A = \begin{pmatrix}
0 & b \delta_{-d} \\
0 & \partial_s
\end{pmatrix}, \quad B = \begin{pmatrix}
0 & \sigma \delta_{-d} \\
0 & 0
\end{pmatrix}, \quad C = \begin{pmatrix}
0 \\
1 \partial(\cdot)
\end{pmatrix}.
\]
where \( l_0(s) = 0 \) if \(-d \leq s < 0\) and \( l_0(0) = 1\). Thus, \( l_0(\cdot) = 0_{L^2}\). Let \( V \) be the value function

\[
V(t, z) = V(t, (x, u)) = \inf_{\alpha \in \mathcal{A}_t,T} \mathbb{E}[(Z^\alpha_T)_{0}] = \inf_{\alpha \in \mathcal{A}_t,T} \mathbb{E}[(X^\alpha_T)^2], \quad z \in H,
\]

where \( Z^\alpha \) denotes the solution to (2.3) starting from \( z = (x, u) \) at time \( t \). Then, assuming \( V \in C^{1,2} \left([0, T] \times L^2([-d, 0]), \mathbb{R}\right)\), the dynamic programming principle reads

\[
V(t, z) = \inf_{\alpha \in \mathcal{A}_t,t+h} \mathbb{E}[V(t+h, Z^\alpha_{t+h})]
\]

\[
= \inf_{\alpha \in \mathcal{A}_t,t+h} \mathbb{E} \left[ V(t, z) + \int_t^{t+h} \partial_t V(s, Z^\alpha_s)ds + \int_t^{t+h} \partial_z V(s, Z^\alpha_s)dZ^\alpha_s \right]
\]

\[
\quad + \frac{1}{2} \int_t^{t+h} \partial^2_z V(s, Z^\alpha_s)d(Z^\alpha_s)_{s},
\]

\[
= \inf_{\alpha \in \mathcal{A}_t,t+h} \mathbb{E} \left[ V(t, z) + \int_t^{t+h} \partial_t V(s, Z^\alpha_s)ds + \int_t^{t+h} \partial_z V(s, Z^\alpha_s)
\quad \times (AZ^\alpha_s ds + Cdx_s) \right] + \frac{1}{2} \int_t^{t+h} \partial^2_z V(s, Z^\alpha_s)d(Z^\alpha_s)_{s}.
\]

As a result, simplifying by \( V(t, z) \), dividing by \( h \) and letting \( h \to 0 \) yields (informally) the Hamilton–Jacobi equation

\[
\partial_t V + \inf_{\alpha \in \mathbb{R}} \{ (\partial_z V, Az)_H + \partial^2_z V(Bz \otimes Bz) \} = 0, \quad t \leq T, \quad z \in \mathbb{R} \times L^2([0, T], \mathbb{R}),
\]

\[
V(T, z) = z_0^2. \tag{2.4}
\]

Recall that in Eq. (2.4), we have \( z_1(0) = u(0) = \alpha \). Let us now assume that the value function \( V \) is of the following form

\[
V(t, z) = \langle P_t z, z \rangle_H,
\]

where \( P \in C([0, T], \mathcal{L}(H, H)) \) is a self-adjoint bounded positive operator valued function of the form

\[
P_t : (x, y(\cdot)) \mapsto \begin{pmatrix}
P_{11}(t)x + \int_{-d}^{0} P_{12}(t, s)y(s)ds \\
P_{12}(t, \cdot) + P_{22}(t, \cdot)y(\cdot) + \int_{-d}^{0} P_{22}(t, \cdot, s)y(s)ds
\end{pmatrix}.
\]

Thus, for any \( z = (x, u) \in H \) such that \( u(0) = \alpha \) and \( t \leq T \), Eq. (2.4) reduces to

\[
\langle \dot{P}_t z, z \rangle_H + \inf_{\alpha \in \mathbb{R}} \{ \langle P_t Az, z \rangle_H + \langle P_t z, Az \rangle_H + \langle P_t Bz, Bz \rangle_H \} = 0. \tag{2.5}
\]
Furthermore, using the boundary condition \( u(0) = \alpha \) together with integration by part, we have

\[
\langle P_I z, Az \rangle_H = (P_I z)_0(Az)_0 + \int_{-d}^0 (P_I z)_1(s)(Az)_1(s)ds
\]

\[
= bu(-d) \left( P_{11}(t)x + \int_{-d}^0 P_{12}(t, s)u(s)ds \right) \\
+ \alpha x P_{12}(t, 0) - u(-d)x P_{12}(t, -d) \\
- x \int_{-d}^0 \partial_s P_{12}(t, s)u(s)ds + \alpha \int_{-d}^0 P_{22}(t, 0, s)u(s)ds \\
- u(-d) \int_{-d}^0 P_{22}(t, -d, s)u(s)ds - \int_{-d}^0 \int_{-d}^0 \partial_s P_{22}(t, s, r)u(s)u(r)dsdr \\
+ \frac{1}{2} \left( \alpha^2 P_{22}^2(t, 0) - u(-d)^2 P_{22}^2(t, -d) \right) - \int_{-d}^0 \partial_r P_{22}^2(t, s)ds,
\]

(2.6)

and

\[
\langle P_I Bz, Bz \rangle_H = \sigma^2 P_{11}(t)u(-d)^2.
\]

(2.7)

**Remark 2.3** In (2.6), along with the integration by part, formulas such as \( 2u\partial_s u = \partial_s u^2 \) were (formally) used. However, as it appears in the verification Theorem 3.1, the feedback optimal control obtained is as regular as the controlled process \( X^\alpha \) and thus as regular as the Brownian motion \( W \). This is why our approach is only heuristic and justifies the need for the verification Theorem 3.1.

As a consequence, the minimizer of the Hamiltonian in (2.5) reads

\[
\alpha^*(t, z) = -\frac{1}{P_{22}(t, 0)} \left( x P_{12}(t, 0) + \int_{-d}^0 P_{22}(t, 0, s)u(s)ds \right).
\]

(2.8)

**Remark 2.4** Note that when \( d \to 0 \), then \( \alpha^*(t, z) \to -\frac{b}{\sigma^2}x \) which agrees with the optimal strategy in the undelayed case.

Combining (2.5), (2.6) and (2.7) yields the set of Riccati partial differential equations

\[
\dot{P}_{11}(t) = \frac{P_{12}(t, 0)^2}{P_{22}^2(t, 0)}, \quad (\partial_t - \partial_s)(P_{12})(t, s) = \frac{P_{12}(t, 0)P_{22}(t, s, 0)}{P_{22}^2(t, 0)},
\]

\[
(\partial_t - \partial_s)(P_{22})(t, s) = 0, \quad (\partial_t - \partial_s - \partial_r)(P_{22})(t, s, r) = \frac{P_{22}(t, s, 0)P_{22}(t, 0, r)}{P_{22}^2(t, 0)},
\]

(2.9)

accompanied by the boundary conditions

\[
\partial_s P_{22}(0, 0) = 0, \quad \partial_r P_{22}(0, 0, 0) = 0.
\]
\[ P_{12}(t, -d) = bP_{11}(t), \quad P_{22}(t, -d) = \sigma^2 P_{11}(t), \]
\[ P_{22}(t, s, -d) = bP_{12}(t, s), \quad P_{22}(t, -d, r) = bP_{12}(t, r), \] (2.10)
and the final conditions
\[ P_{11}(T) = 1, \quad P_{12}(T, s) = P_{22}(T, s) = P_{22}(T, s, r) = 0, \] (2.11)
for almost every \( s, r \in [-d, 0] \).

**Remark 2.5** Looking at the expression (2.8), we can already guess some effects of the existence of a delay on the optimal strategy. Indeed, from (2.10) one notes that \( P_{12} \approx b, P_{22} \approx \sigma^2, P_{22} \approx b^2 \), and we may write
\[ \alpha \approx \frac{-1}{\sigma^2}(bx + db^2\alpha) \approx \frac{-bx}{\sigma^2(1 + d(b/\sigma)^2)}. \]

In Sect. 4, we illustrate numerically the various effects of the delayed control through two examples of Markowitz portfolio allocation with execution delay.

Note that due to the existence of the delay, the value function is independent of the control chosen after \( T - d \), so that \( P_{12}(t, s) = P_{22}(t, s) = P_{22}(t, s, r) = 0 \) whenever \( t + s \geq T - d \) or \( t + r \geq T - d \). Similarly, the optimal control defined in (2.8) is ill defined on \([T - d, T]\) so we decide to set to zero the control after time \( T - d \) and rewrite
\[ \alpha^*(t, z) = -\frac{1}{P_{22}(t, 0)} \left( xP_{12}(t, 0) + \int_{-d}^0 P_{22}(t, 0, s)u(s)ds \right). \] (2.12)

Thus, to make sense of the set of Ricatti partial differential equations (2.9)–(2.10)–(2.11) and the optimal control (2.12), we adopt the convention \( 0^2/0 = 0 \) and define the concept of solution as follows

**Definition 2.6** A 4-uplets \( P = (P_{11}, P_{12}, P_{22}, P_{22}) \) is said to be a solution to (2.9)–(2.10)–(2.11) if \( P_{11} : [0, T] \mapsto \mathbb{R}, P_{22}, P_{22} : [0, T] \times [-d, 0] \mapsto \mathbb{R} \) and \( P_{22} : [0, T] \times [-d, 0]^2 \mapsto \mathbb{R} \) are piecewise absolutely continuous functions satisfying (2.9)–(2.10)–(2.11) with \( P_{22}(t) > 0 \) for any \( t < T - d \).

The reason we chose *piecewise* absolutely continuous functions as our set of functions is because we expect the kernel \( P \) to be discontinuous. To illustrate this consideration, cut the domain \( D \) into three pieces \( D = [0, T] \times [-d, 0]^2 = D_a \cup D_b \cup D_c \) as represented in Fig. 1, with
\[ D_a = [0, T - d] \times [-d, 0]^2, \]
\[ D_b = \{(t, s, r) \in D \text{ s.t. } t > T - d, \ t + s \lor r < T - d\}, \]
\[ D_c = \{(t, s, r) \in D \text{ s.t. } t > T - d, \ t + s \lor r \geq T - d\}. \]
and note that, necessarily, $P_{12}$, $P_{22}$, and $P_{22}$ are null on $D_c$ but not on the remaining domain, see also the numerical simulations in Fig. 7.

In the next section, we provide a proof of the existence and uniqueness of system (2.9)–(2.10)–(2.11), and a verification theorem yielding rigorously the optimal control and value of (2.1)–(2.2).

### 3 Verification and Existence Results

In this section, we establish a verification result for the optimization problem (2.1)–(2.2).

**Theorem 3.1** (Verification Theorem) Assume that

(i) There exists a solution $P$ to (2.9)–(2.10)–(2.11) in the sense of Definition 2.6,

(ii) The control strategy defined as

$$
\alpha^*_t = \frac{-1_{t \leq T-d}}{P_{22}(t, 0)} \left\{ X_t^{\alpha^*} P_{11}(t, 0) + \int_{t-d}^{t} P_{22}(t, 0, s - t) \alpha^*_s ds \right\},
$$

where $X^{\alpha^*}$ denotes the controlled state is an admissible control.

Then the optimization problem (2.1)–(2.2) admits (3.1) as an optimal feedback control. Furthermore, for $z = (x, \gamma) \in H$, the value is given by
\[ V(z) = P_{11}(0)x^2 + 2x \int_{-d}^0 P_{12}(0, s)\gamma_z ds + \int_{-d}^0 P_{22}(0, s)\gamma_z^2 ds \]
\[ + \int_{[-d,0]^2} \gamma_z\gamma_u P_{22}(0, s, r) d\sigma dr \]
\[ = (P_0z, z)_H. \]

Proof The proof is a basic application of the martingale optimality principle, see [8]. Let \( \alpha \in \mathcal{A} \) and define

\[ V_t^\alpha = P_{11}(t)(X_t^\alpha)^2 + 2X_t^\alpha \int_{-d}^t P_{12}(t, s - t)\alpha_s ds + \int_{-d}^t P_{22}(t, s - t)\alpha_s^2 ds \]
\[ + \int_{[t-d,t]^2} P_{22}(t, s - t, r - t)\alpha_s\alpha_r d\sigma dr \]
\[ = (P_tZ_t^\alpha, Z_t^\alpha)_H. \]

An application of Itô’s formula to (3.3) combined with differentiation under the integral symbol, authorized by the assumed boundedness of \( P \) and its derivatives, yield

\[ dV_t^\alpha = \left\{ 1_t(X_t^\alpha)^2 + 2 \left( 2_tX_t^\alpha + 3_tX_t^\alpha + 4_t\int_{-d}^t \alpha_s ds \right) \right\} dt + Z_t^\alpha dW_t, \]

where we have set

\[ 1_t = \dot{P}_{11}(t) \]
\[ 2_t = P_{12}(t, 0) \]
\[ 3_t = 2bP_{11}(t) - P_{12}(t, -d) \]
\[ 4_t(s) = (\partial_t - \partial_s)(E_2)(t, s - t) \]
\[ 5_t = \sigma^2 P_{11}(t) - P_{22}(t, -d) \]
\[ 6_t(s) = \left( bP_{12}(t, \cdot) - P_{22}(t, \cdot, -d) \right)(s - t) \]
\[ 7_t(s) = P_{22}(t, 0, s - t) \]
\[ 8_t(s) = (\partial_t - \partial_s)(P_{22})(t, s - t) \]
\[ 9_t(s, r) = (\partial_t - \partial_s - \partial_r)(P_{22})(t, s - t, r - t), \]

and

\[ Z_t^\alpha = 2\sigma_\alpha_{t-d} \left( X_t^\alpha P_{11}(t) + \int_{-d}^t \alpha_s P_{12}(t, s - t) ds \right). \]

Then, using the set of constraints (2.9)–(2.10)–(2.11) together with (3.1) and a completion of the square in \( \alpha \) yield

\[ dV_t^\alpha = \left( P_{22}(t, 0) (\alpha_t - T(\alpha)_t)^2 \right) dt + Z_t^\alpha dW_t, \]

\( \Box \) Springer
where $T(\alpha)$ is defined as

$$T(\alpha)_t = -\frac{1_{t \leq T-d}}{P_{22}(t,0)} \left\{ X_\alpha^{t_2} P_{12}(t,0) + \int_{t-d}^t \alpha_s P_{22}(t,0, s-t) ds \right\}, \quad t \leq T.$$ 

Note that since the kernels $P_i$’s are bounded and the control $\alpha^*$ is assumed to be admissible, $\alpha^* \in L^2_\mathbb{F}([-d, T], \mathbb{R})$, $X^\alpha$ is continuous and the stochastic integral $\int_0^T Z_s^\alpha dW_s$ is well posed.

Furthermore, it is a local martingale. Thus, there exists a localizing increasing sequence of stopping times $\{\tau_k\}_{k \geq 1}$ converging to $T$ such that $\int_0^{T \wedge \tau_k} Z_s^\alpha dW_s$ is a martingale for every $k \geq 1$. Then, for any $k \geq 1$

$$\mathbb{E}\left[ V_{T \wedge \tau_k}^\alpha \right] = V_0^\alpha + \mathbb{E}\left[ \int_0^{T \wedge \tau_k} P_{22}(s,0)(\alpha_s - T(\alpha)_s) ds \right].$$

Note that $t \mapsto V_t^\alpha$ is continuous since $P$ is bounded and $\alpha \in L^2_\mathbb{F}([-d, T], \mathbb{R})$. Thus, an application of the dominated convergence theorem on the left term (recall that $\mathbb{E}\left[ \sup_{t \leq T} |X_t^\alpha|^2 \right] < \infty$, $\alpha \in L^2_\mathbb{F}([-d, T], \mathbb{R})$ as $\alpha \in \mathcal{A}$) combined with the monotone convergence theorem on the right term yields, as $k \to \infty$

$$\mathbb{E}\left[ V_T^\alpha \right] = \mathbb{E}[(X_T^\alpha)^2] = V_0^\alpha + \mathbb{E}\left[ \int_0^T P_{22}(s,0)(\alpha_s - T(\alpha)_s)^2 ds \right].$$

Note that here we used the assumption $P_{22}(t,0) \geq 0$ on $[0,T]$. Since $P_{22}$ is nonnegative, we obtain that the optimal strategy is given by $\alpha^*$ and that the optimal value equals (3.2). \qed

**Proposition 3.2** Assume that there exists a bounded 4-uplets $P$ solution to (2.9)–(2.10)–(2.11) in the sense of Definition 2.6. Then (3.1) defines an admissible control.

**Proof** Let $\gamma \in L^2([-d, 0], \mathbb{R})$. To prove the claim, note that it suffices to show that the equation

$$\begin{cases} 
\alpha_t = \frac{-1_{t \leq T-d}}{P_{22}(t,0)} \left\{ P_{12}(t,0) \left( x + \int_0^t \alpha_{s-d}(bds + \sigma dW_s) \right) \right. \\
\left. + \int_{t-d}^t \alpha_s P_{22}(t,0, s-t) ds \right\}, \\
\alpha_s = \gamma_s, \quad s \in [-d, 0],
\end{cases} \quad (3.4)$$

admits a solution in $L^2_\mathbb{F}([0, T], \mathbb{R})$ and that the process $X$, then defined as

$$X_t = x + \int_0^t \alpha_{s-d}(bds + \sigma dW_s), \quad t \leq T, \quad (3.5)$$
satisfies $\mathbb{E}\left[\sup_{t \leq T} |X_t|^2\right] < \infty$. To prove the first point, consider the linear operator $\phi$ on $L^2_{\mathcal{F}}([0, T], \mathbb{R})$ defined as, for any $a \in L^2_{\mathcal{F}}([0, T], \mathbb{R})$

$$
\phi(a)_t = \frac{-1}{P_{22}(t, 0)} \left\{ \begin{array}{l}
P_{12}(t, 0) \left( x + \int_0^t \hat{a}_{s-d} (b ds + \sigma dW_s) \right) \\
+ \int_0^t \hat{a}_{s} P_{22}(t, 0, s-t) ds \end{array} \right\},
$$

where $\hat{a}_t = 1_{t \leq 0} \gamma_t + 1_{t > 0} a_t$. For $\lambda \leq 0$, we endow $L^2_{\mathcal{F}}([0, T], \mathbb{R})$ with the norm

$$
\|a\|_{2, \lambda} = \sqrt{\int_0^T e^{-\lambda s} |a_s|^2 ds}. \text{ Then, for any } a, a' \in L^2_{\mathcal{F}}([0, T], \mathbb{R}), \text{ we have}
$$

$$
\|\phi(a) - \phi(a')\|^2_{2, \lambda} = \mathbb{E}\left[ \int_0^T e^{-\lambda s} |\phi(a)_s - \phi(a')_s|^2 ds \right] = 2(I + II).
$$

(3.6)

An application of Jensen’s inequality on the normalized measure $\frac{dr}{s}$ on $[0, s]$, combined with $\frac{1-e^{-\lambda(T-t)}}{\lambda} \leq (1 \vee T)(1 \wedge \lambda^{-1})$, and the Burkholder–Davis–Gundy inequality lead to

$$
I = \mathbb{E}\left[ \int_0^T e^{-\lambda s} \left| \frac{P_{12}(s, 0)}{P_{22}(s, 0)} \int_0^s (\hat{a}_{r-d} - \hat{a}_{r-d}') (b dr + \sigma dW_r) \right|^2 ds \right]
\leq \sup_{s \leq T} \left| \frac{P_{12}(s, 0)}{P_{22}(s, 0)} \right|^2 \left\{ \mathbb{E}\left[ \int_0^T e^{-\lambda s} \left( \int_0^s (\hat{a}_{r-d} - \hat{a}_{r-d}') dr \right)^2 ds \right] + \mathbb{E}\left[ \int_0^T e^{-\lambda s} \left( \int_0^s (\hat{a}_{r-d} - \hat{a}_{r-d}') \sigma dW_r \right)^2 ds \right] \right\}
\leq c(1 \wedge \lambda^{-1}) \mathbb{E}\left[ \int_0^T e^{-\lambda r} (\hat{a}_{r-d} - \hat{a}_{r-d}')^2 dr \right],
$$

where $c > 0$ depends only on $b, \sigma, T$ and $\sup_{s \leq T} \left| \frac{P_{12}(s, 0)}{P_{22}(s, 0)} \right|$. Furthermore, we have

$$
II = \mathbb{E}\left[ \int_0^T e^{-\lambda s} \left| \frac{1}{P_{22}(s, 0)} \int_0^s ((\hat{a}_r - \hat{a}_r') P_{22}(r, 0, s-r) dr \right|^2 ds \right]
\leq \sup_{s \leq T} \frac{P_{22}(s, 0, r)}{P_{22}(s, 0)} \mathbb{E}\left[ \int_0^T e^{-\lambda s} \left( \int_0^s (\hat{a}_r - \hat{a}_r') dr \right)^2 ds \right]
\leq \hat{c}(1 \wedge \lambda^{-1}) \mathbb{E}\left[ \int_0^T e^{-\lambda r} (\hat{a}_r - \hat{a}_r')^2 dr \right],
$$

\( \odot \) Springer
where \( \hat{c} > 0 \) depends only on \( T \) and \( \sup_{r \in [-d, 0]} \left| \frac{P_{22}(s, 0, r)}{P_{22}(s, 0)} \right| \). Consequently, (3.6) reduces to
\[
\| \phi(a) - \phi(a') \|_{2, \lambda}^2 \leq (c + \hat{c})(1 \wedge \lambda^{-1}) \| a - a' \|_{2, \lambda}^2.
\]

As a result, for \( \lambda \) large enough, \( \phi \) is a contraction on the Banach space \( \left( L^2_{\mathbb{F}}([0, T], \mathbb{R}), \| \cdot \|_{2, \lambda} \right) \), thus proving the existence of \( \alpha \in L^2_{\mathbb{F}}([0, T], \mathbb{R}) \) solution to (3.4). Finally, an application of Burkholder–Davis–Gundy’s inequality to (3.5) yields \( E \left[ \sup_{t \leq T} |X_t|^2 \right] \).

The proof is thus complete. \( \square \)

Next, we give a sufficient condition for the existence of \( P = (P_{11}, P_{12}, P_{22}, P_{22}) \) in terms of \( b, \sigma, d \) and \( T \). Let \( a = (a_n)_{n \geq 1} \) be the sequence defined as
\[
\begin{align*}
a_0 & = 1, \\
da_{n+1} & = a_n - \frac{d}{a_n} \left( \frac{b}{\sigma} \right)^2, \quad n \geq 0.
\end{align*}
\]

Let us denote \( \mathcal{N} : (d, b, \sigma) \mapsto \inf \{ n \geq 1 : a_n > 0 \text{ and } a_{n+1} \leq 0 \} \). Clearly, \( \mathcal{N} \) is a well defined finite valued function on \( \mathbb{R}^3 \) whose image is not restricted to \( \{0\} \).

**Proposition 3.3** Assume \( \mathcal{N}(d, b, \sigma) \geq 2 \) and \( T < \mathcal{N}(d, b, \sigma)d \). Then (2.9)–(2.10)–(2.11) has a unique solution in the sense of Definition 2.6 on \([0, T]\) with \( 0 < a_{\mathcal{N}(d, b, \sigma)} \leq P_{11}(0) < 1 \).

**Proof** See Appendix A. \( \square \)

**Remark 3.4** Note that when \( d = 0 \), the sufficient condition above reduces to \( \sigma \neq 0 \).

Let us give some intuition as to why the delay feature induces the condition on the coefficients described above to ensure existence. We focus on the first slice \([T - 2d, T - d] \times [-d, 0]\) of the domain, where the solution \( P \) is not trivial.

First, note that since \( P_{22} \) is a transport of \( P_{11} \) which takes the form \( P_{22}(t, s) = \sigma^2 P_{11}(t + s + d)1_{t + s + d \leq T} = \sigma^2 1_{t + s + d \leq T} \). On this slice, the kernel \( P_{11} \) can be expressed in the following integral form
\[
P_{11}(t) = 1 - \left( \frac{b}{\sigma} \right)^2 \int_t^{T - d} \left( \frac{P_{12}(s, 0)}{b} \right)^2 ds, \quad (3.8)
\]
for \( t \in [T - 2d, T - d] \). Looking at \( P_{12}(\cdot, 0) \), we have
\[ P_{12}(t, 0) = b - \sigma^2 \int_{t}^{T-d} P_{12}(x, 0) \times \frac{P_{22}(x, t-x, 0)}{b} \, dx, \]

\[ \lesssim b^2 \]

see also (A.4). Below, we represent the value of the normalized kernel \( P_{12}/b \) in the different areas of the domain \([T - 2d, T] \times [-d, 0] \). If we visualize the evolution of the normalized kernel \( P_{12}/b \) in a backward way on the slice \([T - 2d, T - d]\), we see that this term is equal to a transport of its value on the boundary \( \frac{P_{12}(T-d,s)}{b} = 1 \), represented by the blue arrows, minus the integral of a positive source term which is independent of \( t \in [T - 2d, T - d] \mapsto P_{11}(t) \).

Consequently, the delay \( d > 0 \) makes the integral term in (3.8) independent of \( t \in [T - 2d, T - d] \mapsto P_{11}(t) \) and of the order of \( d \left( \frac{b}{\sigma} \right)^2 \). If this quantity is too large, the kernel \( P_{11} \) can then reach negative values, thus making \( P_{22} \) negative on the next slice \([T - 3d, T - 2d]\) and therefore preventing the system (2.9)–(2.10)–(2.11) from
Several sequences \((a_n)_{n \geq 0}\) with \(d = 0.1\) for some values of \(b/\sigma\). Note that for \(b/\sigma = 0.76\), \(T = d \times 9 = 0.9\) is the maximum horizon allowed according to condition (3.7) as \(a_{10} < 0 < a_9\) in such case.

having a solution. Repeating this argument from slice to slice of size \(d\) in a backward manner induces the aforementioned sufficient condition. Note that these arguments break down when \(d = 0\). These arguments are precisely developed in Appendix A. See also Fig. 2 for some instances of \((a_n)_{n \geq 0}\).

4 Deep Learning Scheme

4.1 A Quick Reminder of PINNs and Deep Galerkin Method for PDEs

In order to solve (2.9)–(2.10)–(2.11), we will make use of neural networks in the spirit of the emerging Physics Informed Neural Networks (PINNs) and Deep Galerkin literature, see [25] and [22] to name just a few. We first recall some of the main ideas. Assume we have a nonlinear partial differential equation of the form

\[
\partial_t u + \mathcal{G}(u) = 0, \quad \text{on } \Omega,
\]

\[
u = g, \quad \text{on } \partial \Omega,
\]

where \(\mathcal{G}\) is a nonlinear operator, \(\Omega\) a bounded open subset and \(g\) a function on the boundary of the domain. The main idea is to approximate the solution \(u\) to (4.1) by a deep neural network. Let us call \(t \mapsto u(t, \Theta)\) this network, where \(\Theta\) and \(t\) denote respectively its parameters and a generic element of \(\Omega \cup \partial \Omega\). The goal is to find a \(\Theta\) so that \(t \mapsto u(t, \Theta)\) is the closest to satisfying (4.1) in the sense of the neural network loss. To do so, the idea is to proceed by minimizing the mean square error loss.
Fig. 3 Structure of the model used to solve (2.9)–(2.10)–(2.11)

\[
L(\Theta, T) = L^\Omega(\Theta, T) + L^{\partial \Omega}(\Theta, T),
\]

\[
L^\Omega(\Theta, T) = \frac{1}{|T|} \sum_{t \in T} |(\partial_t + \mathcal{G})u(t, \Theta)|^2 1_{t \in \Omega},
\]

\[
L^{\partial \Omega}(\Theta, T) = \frac{1}{|T|} \sum_{t \in T} |u(t, \Theta) - g(t)|^2 1_{t \in \partial \Omega},
\]

where \( L^{\partial \Omega} \) is the loss associated with the final and boundary constraints on \( \partial \Omega \), \( L^\Omega \) the loss associated to the PDE constraint \( \partial_t u + \mathcal{G}(u) = 0 \) on \( \Omega \) and \( T \) a random subset of \( \partial \Omega \cup \Omega \).

### 4.2 A Tailor-Made Algorithm for Riccati Partial Differential Equations

Although (2.9)–(2.10)–(2.11) naturally fits the framework of PINNs, we make use of the structure exhibited on the operator \( P \) to build a tailor-made algorithm to approximate the system of Riccati transport PDEs (2.9)–(2.10)–(2.11).

**Step 1:** We define one neural network for each kernel \( P_{11}, P_{12}, P_{22} \), as described in Fig. 3. Note that usually a unique neural network is used as a surrogate to the function that is to be approximated.

**Step 2:** We build specific loss functionals for each of our neural networks. To illustrate this, consider for instance the constraint imposed on the derivative of \( t \mapsto P_{11}(t) \):

\[
\dot{P}_{11}(t) = \frac{P_{12}(t, 0)^2}{P_{22}(t, 0)}, \quad t \in [0, T].
\]
Note here that, as in the previous sections, we use the convention $0^2/0 = 0$. As a result (4.2) can be rewritten as

$$
\dot{P}_{11}(t) = \frac{P_{12}(t, 0)^2}{P_{22}(t, 0)}, \quad t \leq T - d,
$$

$$
\dot{P}_{11}(t) = 0, \quad t > T - d.
$$

Thus, a natural contribution to the total loss function to enforce (4.2) would be

$$
\mathcal{L}_{11}(\Theta_{11}, \Theta_{12}, \Theta_{22}, T) = \frac{1}{|T|} \sum_{t \in T} \left( \frac{\partial_t P_{11}(\Theta_{11}, t) - P_{12}(\Theta_{12}, t, 0)^2}{P_{22}(\Theta_{22}, t, 0)} \right)^2, \quad T \subset [0, T],
$$

and the natural gradient descent step associated to the constraint (4.2) would be

$$
\Theta_i \leftarrow \Theta_i - \eta \nabla_{\Theta_i} \mathcal{L}_{11}(\Theta_{11}, \Theta_{12}, \Theta_{22}, T), \quad i \in \{11, 12, 22\},
$$

Consequently, the constraint (4.2) a priori entails the updating of $P_{11}$, $P_{12}$ and $P_{22}$. In particular, it requires to compute the gradient of $(\Theta_{12}, \Theta_{22}) \mapsto \sum_{t \in T} \frac{P_{12}(\Theta_{12}, t, 0)^2}{P_{22}(\Theta_{22}, t, 0)}$ which is expected to be highly unstable as $t, s \mapsto P_{22}(t, s)$ vanishes for $t + s \geq T - d$.

To mitigate this issue, the term $t \mapsto \frac{P_{12}(\Theta_{12}, t, 0)^2}{P_{22}(\Theta_{22}, t, 0)}$ is considered as an exogenous source term for $P_{11}$ which is fixed when we train $P_{11}$.

$$
\dot{P}_{11}(t) = \frac{P_{12}(t, 0)^2}{\underbrace{P_{22}(t, 0)}}, \quad t \in [0, T].
$$

Seen as a fixed exogenous source term when $P_{11}$ is trained

To implement this idea, a second set of neural networks $(\tilde{\Theta}_k(\hat{\Theta}_k))_{k \in \{11, 12, 22, 22\}}$ is initialized with $\hat{\Theta}_k = \Theta_k$ for $k \in \{11, 12, 22, 22\}$ at initialization. These additional networks are then used as surrogates to the right-hand side source terms and will not be used for the computation of the gradients of the losses. They will only be updated at the end of each batch training. Consequently, the gradient descent scheme implemented for each batch $T$ is the following:

$$
\Theta_{11} \leftarrow \Theta_{11} - \eta \nabla_{\Theta_{11}} \left( \mathcal{L}_{11}^\Omega(\Theta_{11}, \hat{\Theta}_{12}, \hat{\Theta}_{22}, T) + \mathcal{L}_1^\Omega(\Theta_{11}, T) \right),
$$

$$
\Theta_{12} \leftarrow \Theta_{12} - \eta \nabla_{\Theta_{12}} \left( \mathcal{L}_{12}^\Omega(\Theta_{12}, \hat{\Theta}_{12}, \hat{\Theta}_{22}, T) + \mathcal{L}_1^\Omega(\Theta_{11}, \Theta_{12}, T) + \mathcal{L}_2^\Omega(\Theta_{12}, T) \right),
$$

$$
\Theta_{22} \leftarrow \Theta_{22} - \eta \nabla_{\Theta_{22}} \left( \mathcal{L}_{22}^\Omega(\Theta_{22}, \hat{\Theta}_{22}, T) + \mathcal{L}_1^\Omega(\Theta_{12}, \Theta_{22}, T) + \mathcal{L}_2^\Omega(\Theta_{22}, T) \right),
$$

$$
\Theta_{22} \leftarrow \Theta_{22} - \eta \nabla_{\Theta_{22}} \left( \mathcal{L}_4^\Omega(\Theta_{22}, \hat{\Theta}_{22}, T) + \mathcal{L}_2^\Omega(\Theta_{12}, \Theta_{22}, \hat{\Theta}_{22}, T) + \mathcal{L}_2^\Omega(\Theta_{22}, T) \right),
$$

$$
\tilde{\Theta}_k \leftarrow \Theta_k, \quad k \in \{11, 12, 22, 22\},
$$

$$
(4.3)
$$
where $\mathcal{L}^b_{11}$'s stand for the boundary loss and the $\mathcal{L}^b_{11}$'s stand for the final loss. The precise definitions of the losses are given below.

For each neural network $P_k(\Theta_k), k \in \{11, 12, 22, 22\}$, a follower network is initialized $\tilde{F}_k(\tilde{\Theta}_k), k \in \{11, 12, 22, 22\}$. These follower networks serve as surrogate for the source terms in (2.9)–(2.10)–(2.11). They condition the loss functionals that are used to train the $P_k$'s and are updated at the end of each batch as described in Algorithm 1.

**Losses of $P_{11}$:**

\[
\mathcal{L}^b_{11}(\Theta_{11}, \tilde{\Theta}_{12}, \tilde{\Theta}_{22}) = \frac{1}{|T|} \sum_{t \in T} \left( \partial_t P_{11}(t, \Theta_{11}) - \tilde{P}_{12}(t, 0, \tilde{\Theta}_{12})^2 \right)^2,
\]

\[
\mathcal{L}^f_{11}(\Theta_{11}, T) = \frac{1}{|T|} \sum_{t \in T} \left( (P_{11}(t, \Theta_{11}) - 1)^2 1_{t=T} \right).
\]

**Losses of $P_{12}$:**

\[
\mathcal{L}^b_{12}(\Theta_{12}, \tilde{\Theta}_{22}, \tilde{\Theta}_{22}, T) = \frac{1}{|T|} \sum_{t, s \in T} \left( \partial_t - \partial_s P_{12}(t, s, \Theta_{12}) - \frac{\tilde{P}_{12}(t, 0, \tilde{\Theta}_{12}) \tilde{P}_{22}(t, s, 0, \tilde{\Theta}_{22})}{\tilde{P}_{22}(t, 0, \tilde{\Theta}_{22})} \right)^2,
\]

\[
\mathcal{L}^b_{12}(\tilde{\Theta}_{11}, \Theta_{12}, T) = \frac{1}{|T|} \sum_{t, s \in T} \left( P_{12}(t, s, \Theta_{12}) - b \tilde{P}_{11}(t, \tilde{\Theta}_{11}) \right)^2 1_{s=0, t \neq 0},
\]

\[
\mathcal{L}^f_{12}(\Theta_{12}, T) = \frac{1}{|T|} \sum_{t, s \in T} \left( P_{12}(t, s, \Theta_{12})^2 1_{t=T} \right).
\]

**Losses of $P_{22}$:**

\[
\mathcal{L}^b_{22}(\Theta_{22}, T) = \frac{1}{|T|} \sum_{t, s \in T} \left( (\partial_t - \partial_s) P_{22}(t, s, \Theta_{22}) \right)^2,
\]

\[
\mathcal{L}^b_{22}(\tilde{\Theta}_{11}, \Theta_{22}, T) = \frac{1}{|T|} \sum_{t, s \in T} \left( P_{22}(t, s, \Theta_{22}) - \sigma^2 \tilde{P}_{11}(t, \tilde{\Theta}_{11}) \right)^2 1_{s=0, t \neq T},
\]

\[
\mathcal{L}^f_{22}(\Theta_{22}, T) = \frac{1}{|T|} \sum_{t, s \in T} \left( P_{22}(t, s, \Theta_{22})^2 1_{t=T} \right).
\]
Losses of $P_{22}$:

$$
\mathcal{L}_{22}^\Omega(\Theta_{22}, \tilde{\Theta}_{22}, T) = \frac{1}{|T|} \sum_{t, s, r \in T} \left( (\partial_t \alpha_t - \partial_s \alpha_s - \partial_r \alpha_r) P_4(t, s, r, \Theta_{22}) - \frac{\tilde{P}_{22}(t, 0, r, \tilde{\Theta}_{22}) \tilde{P}_{22}(t, s, 0, \tilde{\Theta}_{22})}{\tilde{P}_{22}(t, 0, \tilde{\Theta}_{22})} \right)^2,
$$

$$
\mathcal{L}_{22}^\Omega(\tilde{\Theta}_{12}, \Theta_{22}, T) = \frac{1}{|T|} \sum_{t, s, r \in T} \left( P_{22}(t, s, r, \Theta_{22}) - b \tilde{P}_{22}(t, s, \tilde{\Theta}_{12}) \right)^2 1_{r=0, t \neq T},
$$

$$
\mathcal{L}_{22}^\Omega(\Theta_{22}, T) = \frac{1}{|T|} \sum_{t, s, r \in T} \left( P_{22}(t, s, r, \Theta_{22}) \right)^2 1_{t=0, s \neq T}.
$$

\textbf{Algorithm 1:} Deep learning scheme to solve (2.9)–(2.10)–(2.11)

\textbf{Result:} A set of optimized parameters $\Theta^* = (\Theta^*_k)_{k \in \{11, 12, \tilde{2}2, 22\}}$;

Initialize the learning rate $\eta$, the neural networks $P(\Theta) = (P_k(\Theta_k))_k$ and $\tilde{P}(\tilde{\Theta}) = (\tilde{P}_k(\tilde{\Theta}_k))_k$;

Copy the weights $\tilde{\Theta}_k \leftarrow \Theta_k, k \in \{11, 12, \tilde{2}2, 22\}$;

\textbf{for each batch do}

\hspace{1cm} Randomly sample $T \subset [0, T] \times [-d, 0]^2$;

\hspace{1cm} Compute the gradient $\nabla_{\Theta} \mathcal{L}(\Theta, \tilde{\Theta}, T)$ as in (4.3);

\hspace{1cm} Update $\Theta \leftarrow \Theta - \eta \nabla_{\Theta} \mathcal{L}(\Theta, \tilde{\Theta}, T)$;

\hspace{1cm} Update $\tilde{\Theta} \leftarrow \tilde{\Theta}$;

\textbf{end}

\textbf{Return:} The set of optimized parameters $\Theta^*$.

As an example, we show in Fig. 4 the losses $\mathcal{L}_{ij}$ obtained during the training of the neural network in the case of a mean–variance portfolio selection with one delayed asset, presented in the next section. During our numerical experiments, the training of the neural networks was stopped after ten consecutive epochs without improvement of the sum of the four losses.

\section{5 Applications to Mean–Variance Portfolio Selection with Execution Delay}

\subsection{5.1 One Asset with Delay}

We now aim at solving the celebrated example of mean–variance portfolio selection, see [19], with execution delay in the spirit of the problem of hedging of European options with execution delay presented in [9]. We present here the settings. Let us consider a standard Black–Scholes financial market, composed of a risk-less asset with zero interest rate.
Fig. 4 Losses $L_{11}$, $L_{12}$, $L_{22}$ and $L_{22}$ during a 300 epochs training with a learning rate of $10^{-4}$ and parameters $\sigma = 1, \lambda = 0.5, d = 1$ and $T = 5$. Here the losses $L_{11}$, $L_{12}$ and $L_{22}$ have been multiplied by a factor 100 and the loss $L_{22}$ by a factor 1000 to prevent gradients from reaching small values too quickly and to take into account the different learning speeds of the $P_{ij}$.

and a risky asset with dynamics

$$dS_t = S_t \{(\sigma \lambda) \, dt + \sigma \, dW_t\}, \quad t \in [0, T],$$

where $\lambda$ and $\sigma$ are constants representing respectively the risk premium and the volatility of the risky asset. At every time $t \in [0, T]$ the investor chooses, based on the information $\mathcal{F}_t$, to allocate the amount of money $\alpha_t \in \mathbb{R}$ into the risky asset. However, due to execution delays this order will be executed at time $t + d$. Set $N^\alpha_t$ (respectively $N^0_t$) the number of risky (respectively risk-less) shares held at time $t$. Then, given an investment strategy $\alpha \in \mathcal{A}$, the value $(X^\alpha_t)_{t \in [0, T]}$ of the portfolio, that we suppose self-financing, follows the dynamics

$$dX^\alpha_t = N^\alpha_t dS_t + (dN^\alpha_t) S_t + (dN^0_t) S^0_t$$

$$= N^\alpha_t S_t \{(\sigma \lambda) \, dt + \sigma \, dW_t\}.$$

Consequently, the controlled state equation of the portfolio’s value is of the form

$$dX^\alpha_t = \alpha_{t-d} ((\sigma \lambda) \, dt + \sigma \, dW_t), \quad t \in [0, T],$$

$$X_0 = x_0, \quad \alpha_s = \gamma_s, \quad \forall s \in [-d, 0].$$
with $x_0 > 0$ and $y \in L^2([-d, 0], \mathbb{R})$. The Mean–Variance portfolio selection problem in continuous-time consists in solving the following constrained problem

\[
\begin{align*}
\min_{\alpha \in \mathcal{A}} & \quad \text{Var}(X_T^\alpha) \\
\text{s.t.} & \quad \mathbb{E}[X_T^\alpha] = c.
\end{align*}
\] (5.1)

It is well-known that problem (5.1) is equivalent to the following max–min problem, see [21, Sect. 6.6.2]

\[
\begin{align*}
\max_{\eta \in \mathbb{R}} \min_{\alpha \in \mathcal{A}} & \quad \mathbb{E}\left[(X_T^\alpha - (c - \eta))^2\right] - \eta^2.
\end{align*}
\] (5.2)

Thus, solving problem (5.1) involves two steps. First, the internal minimization problem in terms of the Lagrange multiplier $\eta$ has to be solved. Second, the optimal value of $\eta$ for the external maximization problem has to be determined. Thus, with $\xi = c - \eta$, we first define the **Inner optimization problem**:

\[
\begin{align*}
\min_{\alpha \in \mathcal{A}} & \quad \mathbb{E}\left[(X_T^\alpha - \xi)^2\right].
\end{align*}
\] (5.3)

Note that, by setting $\tilde{X}^\alpha = X^\alpha - \xi$, the inner problem (5.3) fits into the delayed LQ control problem analyzed in Sect. 3. We first solve the inner optimization problem (5.3) in the following lemma.

**Lemma 5.1** Fix $\eta \in \mathbb{R}$ and $\xi = c - \eta$. Assume $T < dN(d, (\sigma\lambda), \sigma)$ and define $\alpha^*(\xi)$ as the investment strategy

\[
\alpha^*_t(\xi) = \frac{-1_{t \leq T-d}P_{22}(t, 0)}{P_{22}(t, 0)} \left\{ (X^\alpha_t - \xi)P_{12}(t, 0) + \int_{t-d}^t \alpha^*_s(\xi)P_{22}(t, 0, s - t)ds \right\},
\] (5.4)

where $P$ denotes the solution to (2.9)–(2.10)–(2.11) in the sense of Definition 2.6. Then, the inner minimization problem (5.3) admits $\alpha^*(\xi)$ as an admissible optimal feedback strategy and the optimal value is

\[
V_0(\xi) = P_{11}(0)(x_0 - \xi)^2 + R(x_0 - \xi, \gamma),
\] (5.5)

where $R(\gamma)$ denotes the cost associated to the initial investment strategy $\gamma$ on $[-d, 0]$

\[
R(x, \gamma) = 2x \int_{-d}^0 \gamma_sP_{12}(0, s)ds + \int_{-d}^0 \gamma_s^2P_{22}(0, s)ds + \int_{[0,d]} \gamma_s\gamma_uP_{22}(0, s, r)dsdr.
\]

**Proof** First, note that Proposition 3.3 yields the existence and uniqueness of a solution $P$ to (2.9)–(2.10)–(2.11). Furthermore the admissibility of $\alpha^*(\xi)$ results from Proposition 3.2. For any $\alpha \in \mathcal{A}$, define $\tilde{X}^\alpha_t = X^\alpha_t - \xi$. Then, by Itô’s formula we have

\[
d\tilde{X}^\alpha_t = \alpha_{t-d}((\sigma\lambda)dt + \sigma dW_t), \quad t \in [0, T],
\]
\[ \tilde{X}_0 = x_0 - \xi, \quad \alpha_s = \gamma_s, \quad \forall s \in [-d, 0]. \]

As a result, \( \tilde{X}_0 \) and \( X_0 \) have the same dynamics and \( \tilde{X}_0^\alpha = X_0^\alpha - \xi \) so that problem (5.3) can be alternatively written as

\[
\min_{\alpha \in A} \mathbb{E} \left[ \left( \tilde{X}_0^\alpha \right)^2 \right].
\]

Thus, the optimality of \( \alpha^*(\xi) \) and the value (5.5) are immediately given by the verification Theorem 3.1. \(\Box\)

**Theorem 5.2** Assume \( T < dN(d, (\sigma\lambda), \sigma) \). Then, the optimal investment strategy for the maximization problem (5.1) is given by \( a^*(\xi^*) \) defined in (5.4) with \( \xi^* = c - \eta^* \) and

\[
\eta^* = \frac{K(\gamma) + P_{11}(0)(x_0 - c)}{1 - P_{11}(0)}, \quad K(\gamma) = \int_{-d}^{0} \gamma_s P_{12}(0, s) ds. \quad (5.6)
\]

Furthermore, the value of (5.1) is

\[
\text{Var}(X_T^{\alpha^*}) = \frac{P_{11}(0)}{1 - P_{11}(0)} (x_0 - c + K(\gamma))^2 + \int_{-d}^{0} \gamma_s^2 P_{22}(0, s) ds + \int_{[-d,0]^2} \gamma_s \gamma_u P_{22}(0, s, r) ds dr. \quad (5.7)
\]

**Proof** As \( T < dN(d, (\sigma\lambda), \sigma) \), Proposition 3.3 ensures the existence and uniqueness of a solution \( P \) to (2.9)–(2.10)–(2.11). From Lemma 5.1 and (5.2), we have that the max-min problem (5.2), which is equivalent to (5.1), reduces to

\[
\max_{\eta \in \mathbb{R}} \left\{ V_0(c - \eta) - \eta^2 \right\} = \max_{\eta \in \mathbb{R}} \left\{ P_{11}(0) (x_0 - (c - \eta))^2 + R(x_0 - (c - \eta), \gamma) - \eta^2 \right\}.
\]

Furthermore, since \( T < dN(d, (\sigma\lambda), \sigma) \), Proposition 3.3 ensures \( 0 < P_{11}(0) < 1 \) so that the maximization problem is strictly concave. Consequently, \( \eta^* \) given by (5.6) is the optimal parameter. Setting \( \xi^* = c - \eta^* \) in (5.4) and (5.5) results in the optimality of \( \alpha^*(\xi^*) \), and the optimal value (5.7) for the mean–variance problem (5.1). \(\Box\)

**Remark 5.3** In the absence of pre-investment strategy, \( \gamma = 0 \), we recover the usual form of the efficient frontier formula

\[
\text{Var}(X_T^{\alpha^*}) = \frac{P_{11}(0)}{1 - P_{11}(0)} (x_0 - c)^2.
\]
Fig. 5 Efficient frontier with $\sigma = 1, \lambda = 0.5, T = 5$ and $\gamma \equiv 0$

Our observations from the simulations are the following.

**Efficient frontier:** In Fig. 5, we plot the efficient frontier for different delays $d$. Note that the greater the delay, the greater is the variance. This could have been foreseen by observing that, when the initial control is set to 0, i.e. $\gamma = 0_{L^2}$, the value function takes the form $V(x, 0_{L^2}) = P_{11}(0) x^2$, see (3.2). As the value function is clearly an increasing function of the delay, the terminal variance of the portfolio $\text{Var}(X^*_{\alpha}) = \frac{P_{11}(0)}{1-P_{11}(0)} (x_0 - c)^2$ is also an increasing function of the delay.

**Destabilization effect:** In Fig. 6, we plot different scenarios of portfolio allocation. We observe a destabilization effect and a supplement of volatility induced by the delay feature. We also note the tendency to invest more aggressively for greater values of the delay, as the investor has less time to ensure that the promised yield is achieved. We propose the following interpretation: In the classical setting, where $d = 0$, if $Y^*$ denotes the optimal portfolio value process, the optimal investment strategy is of the form $\alpha^*_t = -\frac{b}{\sigma^2} (Y^*_{\alpha} - \mu^*)$ for a certain constant $\mu^* > c$. It can then easily be shown that $Y^* \leq \mu^*$. Thus, the optimal strategy consists in aiming from below at a fixed target $\mu^*$. When $d > 0$, the optimal control is composed of an additional *inertial term*

$$
\alpha^*_t(\xi^*) = -\frac{1}{P_{22}(t, 0)} \left\{ \frac{(X^*_{\alpha} - \xi^*) P_{12}(t, 0)}{1-P_{11}(0)} + \int_{t-d}^{t} \alpha^*_s(\xi) P_{22}(t, 0, s-t) ds \right\}.
$$

so that, contrary to the case where $d = 0$, the optimal control does not cancel when the target $\xi^*$ is attained. Also, note that at every time $t$, the agent does not have any control on the near future from $t$ to $t + d$.

**Kernel $P$:** In Fig. 7, we plot the kernels $P_{11}, P_{12}, P_{22}$, and $P_{22}$. Note the discontinuity between $D_B$ and $D_C$ also described in Fig. 1.
Fig. 6 Different scenarios of the optimal portfolio with $c = 1.6$, $\sigma = 1$, $\lambda = 0.5$, and $T = 5$. Left: $t \mapsto \alpha^*$, right: $t \mapsto X_t^*$. Note the destabilization effect and the supplement of volatility induced by the delay feature. Note also the tendency to invest more aggressively the delayed investor has, as she has less time to ensure the promised yield, with $\xi^* = 2.57$ for a delay $d = 0.5$, $\xi^* = 2.68$ for $d = 1$, $\xi^* = 2.80$ for $d = 1.5$ and $\xi^* = 2.97$ for $d = 2$.

5.2 One Asset with Delay and One Without

To further explore the effect of the delay on the control, we now study a toy example where the investor has two investment opportunities, one with a delayed execution and one without. More precisely, consider the following portfolio dynamic
Fig. 7 Numerical results of Algorithm 1 with $\sigma = 1, \lambda = 0.5, d = 1.5,$ and $T = 5$

\[
\begin{align*}
\mathbf{dX}^\text{(a,b)}_t &= \alpha_t \left\{ (\sigma_1 \lambda_1) dt + \sigma_1 dW^1_t \right\} + \beta_t \left\{ (\sigma_2 \lambda_2) dt + \sigma_2 dW^2_t \right\}, \quad t \in [0, T], \\
X_0 &= x_0, \quad \beta_s = \gamma_s, \quad s \in [-d, 0], \\
\langle W^1, W^2 \rangle_t &= \rho t,
\end{align*}
\]

where $x_0 > 0$ and $\gamma \in L^2([-d, 0], \mathbb{R})$, together with the same optimization objective (5.1) as before. Here, $\alpha_t$ and $\beta_t$ correspond respectively to the amounts of money the investor decides to invest at time $t$ in the undelayed and the delayed risky assets. The constants $\lambda_i$ and $\sigma_i$ represent respectively the risk premium and the volatility of the risky asset $i$. Following the heuristic approach of Sect. 2, we define the following set of Riccati-PDEs on $[0, T] \times [-d, 0]^2$

\[
\begin{align*}
\dot{P}_{11}(t) &= \lambda^2_P P_{11}(t) + \frac{P_{12}(t, 0)^2}{P_{22}(t, 0)}, \\
(\partial_t - \partial_s)(P_{12})(t, s) &= \lambda^2_P P_{12}(t, s) + \frac{P_{12}(t, 0) P_{22}(t, s, 0)}{P_{22}(t, 0)}, \\
(\partial_t - \partial_s)(P_{22})(t, s) &= 0, \\
(\partial_t - \partial_s - \partial_r)(P_{22})(t, s, r) &= \lambda^2_P \frac{P_{12}(t, s) P_{12}(t, r)}{P_{11}(t)} + \frac{P_{22}(t, s, 0) P_{22}(t, 0, r)}{P_{22}(t, 0)},
\end{align*}
\]
accompanied by the boundary conditions, for almost any \( t, s \in [0, T] \times [-d, 0] \)

\[
P_{12}(t, -d) = \lambda_2 \sigma_2 \left( 1 - \rho \frac{\lambda_1}{\lambda_2} \right) P_{11}(t), \quad P_{22}(t, -d) = \sigma_2^2 \left( 1 - \rho^2 \right) P_{11}(t),
\]

\[
P_{22}(t, s, -d) = \lambda_2 \sigma_2 \left( 1 - \rho \frac{\lambda_1}{\lambda_2} \right) P_{12}(t, s), \quad P_{22}(t, -d, s) = \lambda_2 \sigma_2 \left( 1 - \rho \frac{\lambda_1}{\lambda_2} \right) P_{12}(t, s),
\]

and the terminal constraints

\[
P_{11}(T) = 1, \quad P_{12}(T, s) = P_{22}(T, s) = P_{22}(T, s, r) = 0,
\]

for almost every \( s, r \in [-d, 0] \).

As in the previous section, we first solve the inner optimization problem \((5.3)\).

**Lemma 5.4** Fix \( \eta \in \mathbb{R} \) and \( \xi = c - \eta \). Assume \((5.8)-(5.9)-(5.10)\) admits a piecewise absolutely continuous solution with \( P_{22}(t) > 0 \) for any \( t \leq T - d \), and define the couple \((\alpha^*, \beta^*)(\xi)\) as the investment strategies

\[
\alpha^*_t(\xi) = -\left\{ \frac{\lambda_1}{\sigma_1} (X^*_t - \xi) + \frac{\sigma_2}{\sigma_1} \beta^*_{t-d} + \frac{\lambda_1}{\sigma_1} P_{11}(t) \int_{t-d}^t \beta^*_s(\xi) P_{12}(t, s-t) ds \right\},
\]

\[
\beta^*_t(\xi) = -\frac{1}{P_{22}(t, 0)} \left\{ P_{12}(t, 0) (X^*_t - \xi) + \int_{t-d}^t \beta^*_s(\xi) P_{22}(t, 0, r-t) dr \right\},
\]

where \( X^* \) denotes the state process \( X(\alpha^*, \beta^*) \). Then, the inner minimization problem \((5.3)\) admits \((\alpha^*(\xi), \beta^*(\xi))\) as an optimal feedback strategy and the optimal value is

\[
V_0(\xi) = P_{11}(0)(x_0 - \xi)^2 + R(x_0 - \xi, \gamma),
\]

where \( R(\gamma) \) denotes the cost associated to the initial investment strategy \( \gamma \) on \([-d, 0]\)

\[
R(x, \gamma) = 2x \int_{-d}^0 \gamma_s P_{12}(0, s) ds + \int_{-d}^0 \gamma_s^2 P_{22}(0, s) ds + \int_{[-d,0]^2} \gamma_s \gamma_u P_{22}(0, s, r) ds dr.
\]

**Proof** The proof is similar to the one of Lemma 5.1. \(\square\)

Finally, the parameter \( \eta^* \) and efficient frontier \( \text{Var}(X^*_T) = f(c) \) are given by the same formulas \((5.6)\) and \((5.7)\) as in in the mono-asset case, \( \gamma \) being the pre-investment strategy of the delayed asset.

**Remark 5.5** One surprise that emerges is that the “buy the good stock sell the bad one” criterion is unchanged for the delayed asset. Indeed, the sign of the control for this asset is still given by the sign of \( 1 - \rho \frac{\lambda_1}{\lambda_2} \), that fixes the sign of \( P_{12} \) and \( P_{22} \), as it would be in the case without delay\(^1\), see the boundary conditions \((5.9)\). But this threshold

\(^1\) When \( d = 0 \), recall that \( \alpha^*_t = \frac{\lambda_1 P_t}{\sigma_1 (1 - \rho^2)} (\xi^* - X^*_t) \) and \( \beta^*_t = \frac{\lambda_2 P_t}{\sigma_2 (1 - \rho^2)} (\xi^* - X^*_t) \) with \( P \) being a positive function and \( \xi^* \geq X^* \). Thus, in the classical setting, the buy or sell thresholds are \((1 - \rho \frac{\lambda_2}{\lambda_1}) \) and \((1 - \rho \frac{\lambda_1}{\lambda_2}) \).
Fig. 8 $t \mapsto (\alpha^*_t, \beta^*_t)$, with $\sigma_1 = \sigma_2 = 1$, $\lambda_1 = \lambda_2 = 0.5$ and $T = 5$. Blue (solid) : $\alpha^*$, orange (dashed) : $\beta^*$. The same realizations of $W$ and $B$ were used for all experiments. Note that the more positively correlated the assets are, the more favored the undelayed asset is (Color figure online)

Fig. 9 $t \mapsto X^*_t$, with $\sigma_1 = \sigma_2 = 1$, $\lambda_1 = \lambda_2 = 0.5$, $T = 5$ and $d = 1.5$ for $\rho = -0.7$, 0 and 0.7. The same realizations of the Brownian motions $W^1$ and $B$ was used for all experiments.
Fig. 10  \( t \mapsto X^*_t \), with \( \sigma_1 = \sigma_2 = 1, \lambda_1 = \lambda_2 = 0.5, T = 5 \) and \( \rho = -0.7 \) for \( d = 1.5, 1 \) and 0.5. The same realizations of the Brownian motions \( W^1 \) and \( B \) was used for all experiments.

disappears in the undelayed asset’s control as now only the term \( \frac{\lambda_1}{\sigma_1} \) remains in the mean-reverting term.

**Numerical simulations:** To exhibit the effect of the correlation \( \rho \), we generate two independent Brownian motions \( (W^1_t)_{t \in [0,T]} \) and \( (B_t)_{t \in [0,T]} \) and define the Brownian motion \( (W^2_t)_{t \in [0,T]} \) as

\[
W^2_t = \rho W^1_t + \sqrt{1 - \rho^2} B_t, \quad t \in [0,T].
\]

We then compare different scenarios with different values of correlation \( \rho \) and delay \( d \) while fixing \( W^1 \) and \( B \). The numerical simulations can be found in Figs. 8, 9 and 10. As it could have been expected, we see from (5.9) and Fig. 8, that the greater \( \rho \) is, the more favored the undelayed asset is.

6 Conclusions

In this article, we provide a solution to a simple linear-quadratic control problem with delayed control both in drift and volatility. After deriving the solution heuristically, we derive rigorously the solution by a martingale optimality approach. A verification theorem is proved for this stochastic control problem. The solution is obtained in terms of a system of Riccati partial differential equations for which we prove existence and uniqueness under a condition on the coefficient. We then solve numerically this system using a tailor-made deep-learning algorithm based on the Deep Galerkin and PINNs literature. Finally, this method is applied to the problem of the mean–variance portfolio construction for one delayed asset and for two assets where one is delayed and one is not.
A Proof of Proposition 3.3

Our proof extends [1, see Theorem 5] to the case where the volatility is controlled. It consists in slicing the domain \( D \) in slices of size \( d \) and proceeding by a backward recursion. More precisely, we show existence and uniqueness over a sequence of slices \( (\left[T - (n + 1)d, T - nd\right] \times [-d, 0]^2)^n \). We then concatenate the sequence of absolutely continuous solutions obtained, which yields a piecewise absolutely continuous solution. In each slice, the proof consists of the following steps

1. Show that there exists a unique solution on a small interval;
2. Prove that the local solution is Lipschitz;
3. As a result extend the solution to the whole slice.

We finally concatenate the sequence of solutions obtained above.

A.1 Slice \( t \in [T - d, T] \), Initialization

On \( D_b \cup D_c \), the constraints (2.9)–(2.10)–(2.11) on \( P_{12}, P_{22}^\sigma \) and \( P_{22} \) reduce to linear homogeneous transport equations admitting closed form solutions given, for every \((t, s, r) \in D_b \cup D_c \), by

\[
P_{12}(t, s) = bP_{11}(t + s + d)1_{t+s+d \leq T}, \quad P_{22}^\sigma(t, s) = \sigma^2P_{11}(t + s + d)1_{t+s+d \leq T}, \quad P_{22}(t, s, r) = b^2P_{11}(t + s \vee r + d)1_{t+s \vee r + d \leq T}.
\]

Or, as \( P_{11}(t) = 1 \) for any \( t \geq T - d \), we then have for every \((t, s, r) \in D_b \cup D_c \)

\[
P_{11}(t) = 1, \quad P_{12}(t, s) = b1_{t+s+d \leq T}, \quad P_{22}^\sigma(t, s) = \sigma^21_{t+s+d \leq T}, \quad P_{22}(t, s, r) = b^21_{t+s \vee r + d \leq T}.
\]

The existence and uniqueness in the sense of Definition 2.6 are thus trivially proved on \([T - d, T]\).

A.2 Slice \([T - 2d, T - d]\)

On \([T - 2d, T - d] \times [-d, 0]^2\), we have \( P_{22}^\sigma(t, s) = \sigma^2P_{11}(t + s + d) \) so that \( P_{22}(t, 0) = \sigma^2P_{11}(t + d) = \sigma^2 \). Consequently, the system (2.9)–(2.10)–(2.11) reduces to

\[
\dot{P}_{11}(t) = \frac{P_{12}(t, 0)^2}{\sigma^2}.
\]
\begin{align*}
(\partial_t - \partial_s)(P_{12})(t, s) &= \frac{P_{12}(t, 0)P_{22}(t, s, 0)}{\sigma^2}, \\
(\partial_t - \partial_s - \partial_r)(P_{22})(t, s, r) &= \frac{P_{22}(t, s, 0)P_{22}(t, 0, r)}{\sigma^2}, \tag{A.1}
\end{align*}
with terminal conditions
\begin{align*}
P_{11}(T - d) &= 1, \quad P_{12}(T - d, s) = b, \quad P_{22}(T - d, s, r) = b^2, \tag{A.2}
\end{align*}
and boundary constraints
\begin{align*}
P_{12}(t, -d) &= b P_{11}(t), \quad P_{22}(t, s, -d) = b P_{12}(t, s). \tag{A.3}
\end{align*}
Thus, for every \((t, s, r) \in [T - 2d, T - d] \times [-d, 0]^2\), the set of equations (A.1) and constraints (A.2)–(A.3) can be rewritten in the following integral form
\begin{align*}
P_{11}(t) &= 1 - \sigma^{-2} \int_t^{T-d} P_{12}(x, 0)^2 dx, \\
P_{12}(t, s) &= b P_{11}((T - d) \land (t + s + d)) \\
&\quad - \sigma^{-2} \int_t^{(T-d)\land(t+s+d)} P_{12}(x, 0) P_{22}(x, t + s - x, 0) dx, \\
P_{22}(t, s, r) &= b P_{12}((T - d) \land (t + s \land r + d), (s - r) \lor (r - s) - d) \\
&\quad - \sigma^{-2} \int_t^{(T-d)\land(t+s\land r+d)} P_{22}(x, t + s - x, 0) P_{22}(x, 0, t + r - x) dx. \tag{A.4}
\end{align*}
We then make use of the following lemma to prove local existence of a solution.

\textbf{Lemma A.1} \textit{There exists } \tau \in (0, d] \textit{ such that system (A.4) has a unique absolutely continuous solution on } [T - \tau - d, T - d] \times [-d, 0]^2. \textit{ }

\textbf{Proof} \textit{Let } \tau \in (0, d] \textit{ and } S_\tau \textit{ denote the Banach space of absolutely continuous functions } \xi = (\xi_1(\cdot), \xi_2(\cdot, \cdot), \xi_3(\cdot, \cdot, \cdot)) \textit{ defined on }
\begin{align*}
D_\tau = \{(t, s, r) | T - d - \tau \leq t \leq T - d, -d \leq s, r \leq 0\},
\end{align*}
edowed with the sup-norm
\begin{align*}
\|\xi\|_\infty = \|\xi_1\|_\infty + \|\xi_2\|_\infty + \|\xi_3\|_\infty,
\end{align*}
where \(\|\xi_1\|_\infty, \|\xi_2\|_\infty \text{ and } \|\xi_3\|_\infty\) denote, with a slight abuse of notation, the respective sup-norm on \([T - d - \tau, T - d], [T - d - \tau, T - d] \times [-d, 0] \text{ and } [T - d - \tau, T - d] \times [-d, 0]^2\). \textit{Let } B_\tau \textit{ denote the ball in } S_\tau
\begin{align*}
B_\tau = \{(\xi_1, \xi_2, \xi_3) \in S_\tau : \|\xi_1 - 1\|_\infty \leq 1/2, \|\xi_2 - b\|_\infty \leq |b|/2, \|\xi_3 - b^2\|_\infty \leq b^2/2\}.
\end{align*}
On $B_{\tau}$, we denote by $\phi = (\phi_1, \phi_2, \phi_3)$ the operator defined as follows

$$(\phi_1 \xi)(t) = 1 - \sigma^{-2} \int_{t}^{T-d} \xi_2(x, 0)^2 dx$$

$$(\phi_2 \xi)(t, s) = b \phi_1(\xi)((T - d) \wedge (t + s + d))$$

$$- \sigma^{-2} \int_{t}^{(T-d)\wedge(t+s+d)} \xi_2(r, 0)\xi_3(r, t + s - x, 0) dx$$

$$(\phi_3 \xi)(t, s, r) = b \phi_2(\xi)((T - d) \wedge (t + s \wedge r + d), (s - r) \vee (r - s) - d)$$

$$- \sigma^{-2} \int_{t}^{(T-d)\wedge(t+s\wedge r+d)} \xi_3(x, t + s - x, 0)\xi_3(x, 0, t + r - x) dx.$$ 

Clearly, there exists $\tilde{\tau} > 0$ such that for any $\tau \leq \tilde{\tau}$, $\phi(B_{\tau}) \rightarrow B_{\tau}$. We show a contraction property on $\phi$. For any $\xi, \xi' \in B_{\tau}$, we have the following inequalities

$$\|\phi_1(\xi) - \phi_1(\xi')\|_{\infty} \leq 4\tau\sigma^{-2}|b|\|\xi_2 - \xi_2'|\|_{\infty},$$

$$\|\phi_2(\xi) - \phi_2(\xi')\|_{\infty} \leq 4\tau\sigma^{-2}\left(|b|\|\xi_2 - \xi_2'|\|_{\infty} + |b|^2\|\xi_3 - \xi_3'|\|_{\infty}\right)$$

$$+ |b|\|\phi_1(\xi) - \phi_1(\xi')\|_{\infty},$$

$$\|\phi_3(\xi) - \phi_3(\xi')\|_{\infty} \leq |b|\|\phi_2(\xi) - \phi_2(\xi')\|_{\infty} + 4\tau\sigma^{-2}|b|^2\|\xi_3 - \xi_3'|\|_{\infty}.$$ 

Consequently, the operator $\phi$ satisfies

$$\|\phi(\xi) - \phi(\xi')\|_{\infty} \leq \tau m\|\xi - \xi'\|_{\infty},$$

where $m > 0$ depends on $b$ and $\sigma$. Therefore, for $\tau < \tilde{\tau} \wedge m^{-1}$, the operator $\phi$ is a contraction of $B_{\tau}$ into itself. Thus, $\phi$ admits a unique fixed point in $B_{\tau}$, which is solution to (A.4) on $D_{\tau}$. \hfill \Box

**Lemma A.2** Let $\xi = (\xi_1, \xi_2, \xi_3)$ denote the absolutely continuous solution of (A.4) on $D_{\tau}$ from Lemma A.1. Then $\xi$ is Lipschitz in each variable on $D_{\tau}$.

**Proof** As $\xi_1, \xi_2$ and $\xi_3$ are continuous on $D_{\tau}$, there exists a constant $m > 0$ such that $|\xi_1| \wedge |\xi_2| \wedge |\xi_3| \leq m$ on $D_{\tau}$. Thus, $\xi_1$ is Lipschitz with constant $\kappa = m^2\sigma^{-2}$. Let us now show that $\xi_2$ and $\xi_3$ are Lipschitz in the $s$-variable. Fix $t \in [T - d - \tau, T - d]$ and $\eta > 0$. Then, for any $s \in [-d, 0]$ such that $s + \eta \in [-d, 0]$, we have

$$|\xi_2(t, s) - \xi_2(t, s + \eta)| \leq \kappa \eta + \sigma^{-2} \int_{t}^{(T-d)\wedge(t+s+\eta+d)} \xi_2(x, 0)\xi_3(x, t + s + \eta - x, 0) dx$$

$$- \int_{t}^{(T-d)\wedge(t+s+d)} \xi_2(x, 0)\xi_3(x, t + s - x, 0) dx$$

$$\leq \kappa \eta + I(t, s) + II(t, s),$$
Since $|\xi_2| \leq m$, it yields
\[
I(t, s) \leq \int_t^{(T-d)\wedge(t+s+d)} |\xi_2(x, 0)| |\xi_3(x, t + s + \eta - x, 0) - \xi_3(x, t + s - x, 0)| \, dx \\
\leq m \int_t^{(T-d)\wedge(t+s+d)} \epsilon(x) \, dx,
\]
where $\epsilon$ is defined as
\[
\epsilon(x) = \sup_{s, r} |\xi_2(x, s, r) - \xi_3(x, s + \eta, r)| + \sup_{s \in [-d, 0]} |\xi_2(x, s) - \xi_2(x, s + \eta)|.
\]
Furthermore, as $|\xi_2| \wedge |\xi_3| \leq m$ on $D_\tau$, we have
\[
II(t, s) \leq \int_t^{(T-d)\wedge(t+s+\eta+d)} |\xi_2(x, 0)\xi_3(x, t + s + \eta - x, 0)\, dx \\
\leq m^2 \eta.
\]
Consequently, for any $t \in [T - d - \tau, T - d]$, we obtain
\[
\sup_s |\xi_2(t, s) - \xi_2(t, s + \eta)| \leq m^2 \eta + m \int_t^{T-d} \epsilon(r) \, dr. \tag{A.5}
\]
Looking at the equation of $\xi_3$ in system (A.4), we obtain in a similar manner
\[
|\xi_3(t, s, r) - \xi_3(t, s + \eta, r)| \leq |b|I(t, s, r) + \sigma^{-2}II(t, s, r). \tag{A.6}
\]
An application to the triangle inequality combined with (A.5) and the Lipschitzianity of $\xi_1$ leads to
\[
I(t, s, r) \leq |\xi_2((T - d) \wedge (t + (s + \eta) \wedge r + d), (s + \eta - r) \vee (r - (s + \eta)) - d) \\
- \xi_2((T - d) \wedge (t + s + r + d), (s - r) \vee (r - s) - d)| \\
\leq (\kappa + m^2(1 + \sigma^{-2})) \eta + m \int_t^{T-d} \epsilon(x) \, dx \\
\leq (1 + 2\kappa) \eta + m \int_t^{T-d} \epsilon(x) \, dx. \tag{A.7}
\]
Furthermore
\[
II(t, s, r) \leq \left| \int_t^{(T-d)\wedge(t+s+\eta)\wedge r+d} \xi_3(x, t + s + \eta - x, 0)\xi_3(x, 0, t + r - x) \, dx \\
- \int_t^{(T-d)\wedge(t+s+r+d)} \xi_3(x, t + s - x, 0)\xi_3(x, 0, t + r - x) \, dx \right|
\]
\[
\begin{align*}
&\leq \int_t^{(T-d)\wedge(t+s\wedge r+d)} \left| \xi_3(x, 0, t + r - x) \right| \xi_3(x, t + s - x, 0) \\
&\quad - \xi_3(x, t + (s + \eta) - x, 0) \right| \, dx \\
&\quad + \int_t^{(T-d)\wedge(t+(s+\eta))\wedge r+d)} \left| \xi_3(x, t + (s + \eta) - x, 0) \xi_3(r, 0, t + r - x) \right| \, dx \\
&\leq m^2 \eta + \int_t^{T-d} \epsilon(r) \, dr \\
\end{align*}
\tag{A.8}
\]

Thus, inequality (A.7) together with (A.8) and (A.6) yield the existence of a positive constant \(c > 0\), independent of \(\eta\), such that

\[
\sup_{s,r,\in [-d,0]^2} |\xi_3(t, s, r) - \xi_3(t, s + \eta, r)| \leq c \left( \eta + \int_t^{T-d} \epsilon(r) \, dr \right),
\]

which, combined with (A.5) leads, for any \(t \in [T - d - \tau, T - d]\), to

\[
\epsilon(t) \leq c \left( \eta + \int_t^{T-d} \epsilon(r) \, dr \right).
\]

Consequently, an application to Gronwall’s lemma yields \(\epsilon(t) \leq m' \eta\) on \([T - d - \tau, T - d]\), with \(m' > 0\). Thus, \(\xi_2\) and \(\xi_3\) are Lipschitz in the \(s\)-variable. The arguments for showing that \(\xi_2\) and \(\xi_3\) are Lipschitz in the \(t\)-variable and \(\xi_3\) Lipschitz in the \(r\)-variable follow the same line. \(\square\)

**Lemma A.3** There exists a unique absolutely continuous solution \(\xi = (\xi_1, \xi_2, \xi_3)\) of (A.4) on \([T - 2d, T - d] \times [-d, 0]\) such that \(\xi_1 \geq 1 - d \left( \frac{b}{r} \right)^2 > 0\).

**Proof** Let \(\theta \in [T - 2d, T - d]\) denote the lower limit of all \(\tau\)'s such that there exists an absolutely continuous solution \((\xi_1, \xi_2, \xi_3)\) to (A.4) on \([\theta, T - d]\). Assume \(\theta > T - 2d\). From Lemma A.2, \(\xi_2\) and \(\xi_3\) are Lipschitz in each variable and thus admit a limit, when \(t \to \theta\), which is Lipschitz. Therefore, the argument of Lemma A.1 can be repeated to extend the existence and uniqueness of the solution of system (A.4) on \([\xi, T - d]\) for \(T - 2d \leq \xi < \theta\). As a result, we necessarily have \(\theta = T - 2d\). It remains to prove that \(0 < \xi_1\). For this, note that since \(\xi_1\) is solution to (A.4), we have

\[
\|\xi_1 - 1\|_\infty \leq \frac{d}{\sigma^2} \sup_{t \in [T-2d,T-d]} |\xi_2(t, 0)|^2. \tag{A.9}
\]

By injecting the boundary condition (A.2) into the system (A.4), one notes that \(t \in [T - 2d, T - d] \mapsto \xi_2(t, 0)\) is solution to

\[
\xi_2(t, 0) = b - \sigma^{-2} \int_t^{T-d} \xi_2(x, 0) \xi_3(x, t - x, 0) \, dx, \quad T - 2d \leq t \leq T - d.
\]
But for every \( t \in [ T - 2d, T - d ] \), \( f_t : x \in [ t, T - d ] \mapsto f_t(x) := \xi_3(x, t - x, 0) \) takes only positive values as \( f_t \) is solution to the system

\[
f_t(x) = b^2 - \sigma^2 \int_x^{T - d} f_t(u) \xi_3(u, 0, x - u) du, \quad x \in [ t, T - d ],
\]

\[
f_t(T - d) = b^2,
\]

which can be proven to admit, through a contraction proof in the Banach space \( C([t, T - d], \mathbb{R}) \), a unique positive solution since \( \xi \) and its derivatives are bounded. Similarly, we also have \( \xi_2(t, 0) \geq 0 \) for any \( t \in [ T - 2d, T - d ] \). As a result, we have \( \text{sign}(\xi_2) = \text{sign}(b) \) and

\[
\sup_{[T - 2d, T - d]} |\xi_2(t, 0)| \leq |b|.
\]

(A.10)

Consequently, (A.9) and (A.10) yield that for any \( T - 2d \leq t \leq T - d \), we have \( \xi_1 \geq 1 - d \left( \frac{b^2}{\sigma^2} \right)^2 = a_2 > 0 \) as \( \mathcal{N}(d, b, \sigma) \) is assumed to be greater than 2.

Finally, by setting \( P_{11}(t) = \xi_1(t), P_{12}(t, s) = \xi_2(t, s), P_{22}(t, s, r) = \xi_3(t, s, r) \) and \( P_{22}(t, s) = \xi_1(t + s + d) \) for any \( (t, s, r) \in [ T - 2d, T - d ] \times [ -d, 0 ] \), Lemma A.3 yields the existence and uniqueness of a solution \( P \) to (2.9)–(2.10)–(2.11) in the sense of Definition 2.6 on \( [ T - 2d, T - d ] \). The concatenation of the unique solution of (2.9)–(2.10)–(2.11) on \( [ T - d, T ] \) and \( [ T - 2d, T - d ] \) leads to a unique solution on \( [ T - 2d, T ] \).

### A.3 From Slice \([T - nd, T]\) to \([T - (n + 1)d, T]\)

Let \( n \) be an integer such that \( 2 \leq n < \mathcal{N}(d, b, \sigma) \). Assume that there exists a solution \( P \) to (2.9)–(2.10)–(2.11) in the sense of Definition 2.6 on \( [ T - nd, T ] \) such that \( 0 < a_n \leq P_{11}(t) \leq 1 \), for any \( t \geq T - nd \). Recall the Definition 3.7 of \((a_n)_{n \geq 0} \). Consider the following system on \([T - (n + 1)d, T - nd] \times [ -d, 0 ]^2\)

\[
P_{11}(t) = P_{11}(T - nd) - \int_t^{T - nd} \frac{P_{12}(x, 0)^2}{\sigma^2 P_{11}(x + d)} dx,
\]

\[
P_{12}(t, s) = b P_{11}((T - nd) \land (t + s + d)) - \int_t^{(T - nd) \land (t + s + d)} \frac{P_{12}(x, 0) P_{22}(x, t + s - x, 0)}{\sigma^2 P_{11}(x + d)} dx,
\]

\[
P_{22}(t, s, r) = b P_{12}((T - nd) \land (t + s \land r + d), (s - r) \lor (r - s) - d) - \int_t^{(T - nd) \land (t + s \land r + d)} \frac{P_{22}(x, t + s - x, 0) P_{22}(x, 0, t + r - x)}{\sigma^2 P_{11}(x + d)} dx.
\]

(A.11)

Note that this system is the same as (A.4), the only difference being the term \( x \in [ T - (n + 1)d, T - nd ] \mapsto P_{11}(x + d) \) which comes from the previous slice \([ T -
Therefore, it can be considered as a positive continuous coefficient by induction hypothesis. As result, existence and uniqueness on \([T - (n + 1)d, T - nd]\) can be proven in the same fashion as in Lemmas A.1–A.2–A.3. It remains to prove that \(P_{11}(t) \geq a_{n+1}\) for any \(t \in [T - (n + 1)d, T - nd]\). As in Lemma A.3, and by using the induction hypothesis, we have

\[
|P_{12}(t, -d)| \leq |bP_{11}(T - nd)| \leq |b|, \quad t \in [T - (n + 1)d, T - nd]. \tag{A.12}
\]

Furthermore, \(P_{11}\) satisfies (A.11) on \([T - (n + 1)d, T - nd]\), which, combined with \(P_{11} \geq a_n\) on \([T - nd, T - (n - 1)d]\) and (A.12) yields

\[
P_{11}(t) \geq P_{11}(T - nd) - \frac{d}{a_n} \left(\frac{b}{\sigma}\right)^2 \geq a_n - \frac{d}{a_n} \left(\frac{b}{\sigma}\right)^2 = a_{n+1} > 0,
\]

for any \(t \in [T - (n + 1)d, T - nd]\), which ends the proof.

References

1. Alekal, Y., Brunovsky, P., Chyung, D., Lee, E.: The quadratic problem for systems with time delays. IEEE Trans. Autom. Control 16(6), 673–687 (1971)
2. Asea, P.K., Zak, P.J.: Time-to-build and cycles. J. Econ. Dyn. Control 23(8), 1155–1175 (1999)
3. Bambi, M.: Endogenous growth and time-to-build: the AK case. J. Econ. Dyn. Control 32(4), 1015–1040 (2008)
4. Bambi, M., Fabbri, G., Gozzi, F.: Optimal policy and consumption smoothing effects in the time-to-build AK model. Econ. Theory 50(3), 635–669 (2012)
5. Bensoussan, A., Da Prato, G., Delfour, M.C., Mitter, S.: Representation and Control of Infinite Dimensional Systems, 2nd edn. Birkhäuser, Boston, MA (2007)
6. Carmona, R., Fouque, J.-P., Mousavi, M., Sun, L.-H.: Systemic risk and stochastic games with delay. J. Optim. Theory Appl. 179(2), 366–399 (2018)
7. dAlbis, H., Augeraud-Véron, E., Venditti, A.: Business cycle fluctuations and learning-by-doing externalities in a one-sector model. J. Math. Econ. 48(5), 295–308 (2012)
8. El Karoui, N.: Les Aspects Probabilistes Du Controle Stochastique. 9th Saint Flour Probability Summer School-1979. Lecture Notes in Mathematics, vol. 876, pp. 73–238. Springer, Berlin (1981)
9. Fabbri, G., Federico, S.: On the infinite-dimensional representation of stochastic controlled systems with delayed control in the diffusion term. Math. Econ. Lett. 23(3–4), 33–43 (2014)
10. Fouque, J.-P., Zhang, Z.: Deep learning methods for mean field control problems with delay. Front. Appl. Math. Stat. 6, 11 (2020)
11. Gozzi, F., Marinelli, C.: Stochastic optimal control of delay equations arising in advertising models. In: Stochastic Partial Differential Equations and Applications—VII. Lecture Notes in Pure and Applied Mathematics, vol. 245, pp. 133–148. Chapman & Hall/CRC, Boca Raton (2006)
12. Gozzi, F., Marinelli, C., Savin, S.: On controlled linear diffusions with delay in a model of optimal advertising under uncertainty with memory effects. J. Optim. Theory Appl. 142(2), 291–321 (2009)
13. Hall, R.E., Sims, C.A., Modigliani, F., Brainard, W.: Investment, interest rates, and the effects of stabilization policies. Brook. Pap. Econ. Act. 1977(1), 61–121 (1977)
14. Han, J., Hu, R.: Recurrent neural networks for stochastic control problems with delay. Math. Control Signals Syst. (2021). https://doi.org/10.1007/s00498-021-00300-3
15. Huzmezan, M., Gough, W.A., Dumont, G.A., Kovac, S.: Time delay integrating systems: a challenge for process control industries. A practical solution. Control Eng. Pract. 10(10), 1153–1161 (2002)
16. Ichikawa, A.: Quadratic control of evolution equations with delays in control. SIAM J. Control Optim. 20(5), 645–668 (1982)
17. Jarlebring, E., Damm, T.: The Lambert W function and the spectrum of some multidimensional time-delay systems. Automatica 43(12), 2124–2128 (2007)
18. Kydland, F.E., Prescott, E.C.: Time to build and aggregate fluctuations. Econometrica 50(6), 1345–1370 (1982)
19. Markowitz, H.: Portfolio selection. J. Financ. 7(1), 77–91 (1952)
20. Pauwels, W.: Optimal dynamic advertising policies in the presence of continuously distributed time lags. J. Optim. Theory Appl. 22(1), 79–89 (1977)
21. Pham, H.: Continuous-Time Stochastic Control and Optimization with Financial Applications. Springer, Berlin (2009)
22. Raissi, M., Perdikaris, P., Karniadakis, G.E.: Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J. Comput. Phys. 378, 686–707 (2019)
23. Sethi, S.P.: Sufficient conditions for the optimal control of a class of systems with continuous lags. J. Optim. Theory Appl. 13(5), 545–552 (1974)
24. Sipahi, R., Niculescu, S.-I., Abdallah, C.T., Michiels, W., Gu, K.: Stability and stabilization of systems with time delay. IEEE Control Syst. Mag. 31(1), 38–65 (2011)
25. Sirignano, J., Spiliopoulos, K.: DGM: A deep learning algorithm for solving partial differential equations. J. Comput. Phys. 375, 1339–1364 (2018)
26. Tian, Y.-C., Gao, F.: Control of integrator processes with dominant time delay. Ind. Eng. Chem. Res. 38(8), 2979–2983 (1999)
27. Tsoukalas, J.D.: Time to build capital: revisiting investment-cash flow sensitivities. J. Econ. Dyn. Control 35(7), 1000–1016 (2011)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.