



**HAL**  
open science

# Galerkin approximations for the optimal control of nonlinear delay differential equations

Mickaël D Chekroun, Axel Kröner, Honghu Liu

► **To cite this version:**

Mickaël D Chekroun, Axel Kröner, Honghu Liu. Galerkin approximations for the optimal control of nonlinear delay differential equations. [Research Report] INRIA. 2017. hal-01534673

**HAL Id: hal-01534673**

**<https://hal.science/hal-01534673v1>**

Submitted on 15 Jun 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# GALERKIN APPROXIMATIONS FOR THE OPTIMAL CONTROL OF NONLINEAR DELAY DIFFERENTIAL EQUATIONS

MICKAËL D. CHEKROUN, AXEL KRÖNER, AND HONGHU LIU

ABSTRACT. Optimal control problems of nonlinear delay differential equations (DDEs) are considered for which we propose a general Galerkin approximation scheme built from Koornwinder polynomials. Error estimates for the resulting Galerkin-Koornwinder approximations to the optimal control and the value function, are derived for a broad class of cost functionals and nonlinear DDEs. The approach is illustrated on a delayed logistic equation set not far away from its Hopf bifurcation point in the parameter space. In this case, we show that low-dimensional controls for a standard quadratic cost functional can be efficiently computed from Galerkin-Koornwinder approximations to reduce at a nearly optimal cost the oscillation amplitude displayed by the DDE's solution. Optimal controls computed from the Pontryagin's maximum principle (PMP) and the Hamilton-Jacobi-Bellman equation (HJB) associated with the corresponding ODE systems, are shown to provide numerical solutions in good agreement. It is finally argued that the value function computed from the corresponding reduced HJB equation provides a good approximation of that obtained from the full HJB equation.

## CONTENTS

1. Introduction	1
2. Optimal control of DDEs: Galerkin-Koornwinder approximations	3
2.1. Recasting a DDE into an infinite-dimensional evolution equation	3
2.2. Galerkin-Koornwinder approximation of DDEs	5
2.3. Galerkin-Koornwinder approximation of nonlinear optimal control problems associated with DDEs	7
3. Application to an oscillation amplitude reduction problem	13
3.1. Optimal control of the Wright equation	13
3.2. Galerkin-Koornwinder approximation of the optimal control problem	16
3.3. Numerical results	17
4. Towards the approximation of HJB equations associated with DDEs	21
5. Concluding remarks	26
Acknowledgments	26
References	26

## 1. INTRODUCTION

Delay differential equations (DDEs) are widely used in many fields such as biosciences [20, 41, 45, 54], climate dynamics [15, 24, 40, 56, 58], chemistry and engineering [26, 38, 43, 46, 55]. The inclusion of time-lag terms are aimed to account for delayed responses of the modeled systems to either internal or external factors. Examples of such factors include incubation period of infectious diseases [41], wave propagation [10, 56], or time lags arising in engineering [38] to name a few.

---

*Key words and phrases.* Delay differential equations, Galerkin approximations, Hamilton-Jacobi-Bellman equation, Hopf bifurcation, Koornwinder polynomials, Optimal control.

In contrast to ordinary differential equations (ODEs), the phase space associated even with a scalar DDE is infinite-dimensional, due to the presence of delay terms implying thus the knowledge of a function as an initial datum, namely the initial history to solve the Cauchy problem [18, 26]; cf. Section 2. The infinite-dimensional nature of the state space renders the related optimal control problems more challenging to solve compared to the ODE case. It is thus natural to seek for low-dimensional approximations to alleviate the inherent computational burden. The need of such approximations is particularly relevant when (nearly) optimal controls in feedback form are sought, to avoid solving the infinite-dimensional Hamilton-Jacobi-Bellman (HJB) equation associated with the optimal control problem of the original DDE.

For optimal control of *linear* DDEs, averaging methods as well as spline approximations are often used for the design of (sub)optimal controls in either open-loop form or feedback form; see e.g. [1, 4, 34, 42] and references therein. The usage of spline approximations in the case of open-loop control for *nonlinear* DDEs has also been considered [44], but not in a systematic way. Furthermore, due to the locality of the underlying basis functions, the use of e.g. spline approximations for the design of feedback controls leads, especially in the nonlinear case, to surrogate HJB equations of too high dimension to be of practical interest.

In this article, we bring together a recent approach dealing with the Galerkin approximations for optimal control problems of general nonlinear evolution equations in a Hilbert space [11], with techniques for the finite-dimensional and analytic approximations of nonlinear DDEs based on Legendre-type polynomials, namely the Koornwinder polynomials [9]. Within the resulting framework, we adapt ideas from [11, Sect. 2.6] to derive—for a broad class of cost functionals and nonlinear DDEs—error estimates for the approximations of the value function and optimal control, associated with the Galerkin-Koornwinder (GK) systems; see Theorem 2.1 and Corollary 2.1. These error estimates are formulated in terms of residual energy, namely the energy contained in the controlled DDE solutions as projected onto the orthogonal complement of a given Galerkin subspace.

Our approach is then applied to a Wright equation<sup>1</sup> with the purpose of reducing (optimally) the amplitude of the oscillations displayed by the DDE’s solution subject to a quadratic cost; see Sect. 3. For this model, the oscillations emerge through a supercritical Hopf bifurcation as the delay parameter  $\tau$  crosses a critical value,  $\tau_c$ , from below. We show, for  $\tau$  above and close to  $\tau_c$ , that a suboptimal controller at a nearly optimal cost can be synthesized from a 2D projection of the 6D GK system; see Sect. 3.3. The projection is made onto the space spanned by the first two eigenvectors of the linear part of the 6D GK system; the sixth dimension constituting for this example the minimal dimension—using Koornwinder polynomials—to resolve accurately the linear contribution to the oscillation frequency of the DDE’s solution; see (3.29) in Sect. 3.3.

Using the resulting 2D ODE system, the syntheses of (sub)optimal controls obtained either from application of the Pontryagin’s maximum principle (PMP) or by solving the associated HJB equation, are shown to provide nearly identical numerical solutions. Given the good control skills obtained from the controller synthesized from the reduced HJB equation, one can reasonably infer that the corresponding “reduced” value function provides thus a good approximation of the “full” value function associated with the DDE optimal control problem; see Sect. 4.

The article is organized as follows. In Sect. 2, we first introduce the class of nonlinear DDEs considered in this article and recall in Sect. 2.1 how to recast such DDEs into infinite-dimensional evolution equations in a Hilbert space. We summarize then in Sect. 2.2 the main tools from [9] for the analytic determination of GK systems approximating DDEs. In Sect. 2.3, we derive error estimates for the approximations of the value function and optimal control, obtained from the GK systems. Our approach is applied to the Wright equation in Sect. 3. Section 3.1 introduces the optimal control problem

---

<sup>1</sup>Analogue to a logistic equation with delay; see Sect. 3.1.

associated with this equation and its abstract functional formulation following Sect. 2.1. The explicit GK approximations of the corresponding optimal control problem are derived in Sect. 3.2. Numerical results based on PMP from a projected 2D GK system are then presented in Sect. 3.3. In Sect. 4, we show that by solving numerically the associated reduced HJB equation, a (sub)optimal control in a feedback form can be synthesized at a nearly optimal cost. Finally, directions to derive reduced systems of even lower dimension from GK approximations are outlined in Sect. 5.

## 2. OPTIMAL CONTROL OF DDES: GALERKIN-KOORNWINDER APPROXIMATIONS

In this article, we are concerned with optimal control problems associated with nonlinear DDEs of the form

$$(2.1) \quad \frac{dm(t)}{dt} = am(t) + bm(t - \tau) + c \int_{t-\tau}^t m(s)ds + F\left(m(t), m(t - \tau), \int_{t-\tau}^t m(s)ds\right),$$

where  $a$ ,  $b$  and  $c$  are real numbers,  $\tau > 0$  is the delay parameter, and  $F$  is a nonlinear function. We refer to Sect. 2.3 for assumption on  $F$ . To simplify the presentation, we restrict ourselves to this class of scalar DDEs, but nonlinear systems of DDEs involving several delays are allowed by the approximation framework of [9] adopted in this article.

Note that even for scalar DDEs such as given by Eq. (2.1), the associated state space is infinite-dimensional. This is due to the presence of time-delay terms, which require providing initial data over an interval  $[-\tau, 0]$ , where  $\tau > 0$  is the delay. It is often desirable, though, to have low-dimensional ODE systems that capture qualitative features, as well as approximating certain quantitative aspects of the DDE dynamics.

The approach adopted here consists of approximating the infinite-dimensional equation (2.1) by a finite-dimensional ODE system built from Koornwinder polynomials following [9], and then solve reduced-order optimal control problems aimed at approximating a given optimal control problem associated with Eq. (2.1).

To justify this approach, Eq. (2.1) needs first to be recast into an evolution equation, in order to apply in a second step, recent results dealing with the Galerkin approximations to optimal control problems governed by general nonlinear evolution equations in Hilbert spaces, such as presented in [11].

As a cornerstone, rigorous Galerkin approximations to Eq. (2.1) are crucial, and are recalled hereafter. We first recall how a DDE such as Eq. (2.1) can be recast into an infinite-dimensional evolution equation in a Hilbert space.

**2.1. Recasting a DDE into an infinite-dimensional evolution equation.** The reformulation of a system of DDEs into an infinite-dimensional ODE is classical. For this purpose, two types of function spaces are typically used as state space: the space of continuous functions  $C([-\tau, 0]; \mathbb{R}^d)$ , see [26], and the Hilbert space  $L^2([-\tau, 0]; \mathbb{R}^d) \times \mathbb{R}^d$ , see [18]. The Banach space setting of continuous functions has been extensively used in the study of bifurcations arising from DDEs, see e.g. [8, 17, 19, 21, 37, 47, 60], while the Hilbert space setting is typically adopted for the approximation of DDEs or their optimal control; see e.g. [1–3, 9, 22, 28, 33, 35, 36, 42].

Being concerned with the optimal control of scalar DDEs of the form (2.1), we adopt here the Hilbert space setting and consider in that respect our state space to be

$$(2.2) \quad \mathcal{H} := L^2([-\tau, 0]; \mathbb{R}) \times \mathbb{R},$$

endowed with the following inner product, defined for any  $(f_1, \gamma_1)$  and  $(f_2, \gamma_2)$  in  $\mathcal{H}$  by:

$$(2.3) \quad \langle (f_1, \gamma_1), (f_2, \gamma_2) \rangle_{\mathcal{H}} := \frac{1}{\tau} \int_{-\tau}^0 f_1(\theta) f_2(\theta) d\theta + \gamma_1 \gamma_2.$$

We will also make use sometimes of the following subspace of  $\mathcal{H}$ :

$$(2.4) \quad \mathcal{W} := H^1([-\tau, 0]; \mathbb{R}) \times \mathbb{R},$$

where  $H^1([-\tau, 0]; \mathbb{R})$  denotes the standard Sobolev subspace of  $L^2([-\tau, 0]; \mathbb{R})$ .

Let us denote by  $m_t$  the time evolution of the history segments of a solution to Eq. (2.1), namely

$$(2.5) \quad m_t(\theta) := m(t + \theta), \quad t \geq 0, \quad \theta \in [-\tau, 0].$$

Now, by introducing the new variable

$$(2.6) \quad y(t) := (m_t, m(t)) = (m_t, m_t(0)), \quad t \geq 0,$$

Eq. (2.1) can be rewritten as the following abstract ODE in the Hilbert space  $\mathcal{H}$ :

$$(2.7) \quad \frac{dy}{dt} = \mathcal{A}y + \mathcal{F}(y),$$

where the linear operator  $\mathcal{A}: D(\mathcal{A}) \subset \mathcal{W} \rightarrow \mathcal{H}$  is defined as

$$(2.8) \quad [\mathcal{A}\Psi](\theta) := \begin{cases} \frac{d^+ \Psi^D}{d\theta}, & \theta \in [-\tau, 0), \\ a\Psi^S + b\Psi^D(-\tau) + c \int_{-\tau}^0 \Psi^D(s) ds, & \theta = 0, \end{cases}$$

with the domain  $D(\mathcal{A})$  of  $\mathcal{A}$  given by (cf. e.g. [33, Prop. 2.6])

$$(2.9) \quad D(\mathcal{A}) = \left\{ (\Psi^D, \Psi^S) \in \mathcal{H} : \Psi^D \in H^1([-\tau, 0]; \mathbb{R}), \lim_{\theta \rightarrow 0^-} \Psi^D(\theta) = \Psi^S \right\}.$$

The nonlinearity  $\mathcal{F}: \mathcal{H} \rightarrow \mathcal{H}$  is here given, for all  $\Psi = (\Psi^D, \Psi^S)$  in  $\mathcal{H}$ , by

$$(2.10) \quad [\mathcal{F}(\Psi)](\theta) := \begin{cases} 0, & \theta \in [-\tau, 0), \\ F(\Psi^S, \Psi^D(-\tau), \int_{-\tau}^0 \Psi^D(s) ds), & \theta = 0, \end{cases}$$

With  $D(\mathcal{A})$  such as given in (2.9), the operator  $\mathcal{A}$  generates a linear  $C_0$ -semigroup on  $\mathcal{H}$  so that the Cauchy problem associated with the linear equation  $\dot{y} = \mathcal{A}y$  is well-posed in the Hadamard's sense; see e.g. [18, Thm. 2.4.6]. The well-posedness problem for the nonlinear equation depends obviously on the nonlinear term  $\mathcal{F}$  and we refer to [9] and references therein for a discussion of this problem; see also [59].

For later usage, we recall that a *mild solution* to Eq. (2.7) over  $[0, T]$  with initial datum  $y(0) = x$  in  $\mathcal{H}$ , is an element  $y$  in  $C([0, T], \mathcal{H})$  that satisfies the integral equation

$$(2.11) \quad y(t) = T(t)x + \int_0^t T(t-s)\mathcal{F}(y(s))ds, \quad \forall t \in [0, T],$$

where  $(T(t))_{t \geq 0}$  denotes the  $C_0$ -semigroup generated by the operator  $\mathcal{A}: D(\mathcal{A}) \rightarrow \mathcal{H}$ . This notion of mild solutions extend naturally when a control term  $\mathfrak{C}(u(t))$  is added to the RHS of Eq. (2.7); in such a case the definition of a mild solution is amended by the presence of the integral term  $\int_0^t T(t-s)\mathfrak{C}(u(s))ds$  to the RHS of Eq. (2.11).

**2.2. Galerkin-Koornwinder approximation of DDEs.** Once a DDE is reframed into an infinite-dimensional ODE in  $\mathcal{H}$ , the approximation problem by finite-dimensional ODEs can be addressed in various ways. In that respect, different basis functions have been proposed to decompose the state space  $\mathcal{H}$ ; these include, among others, step functions [1,36], splines [2,3], and orthogonal polynomial functions, such as Legendre polynomials [28,33]. Compared with step functions or splines, the use of orthogonal polynomials leads typically to ODE approximations with lower dimensions, for a given precision [2,28]. On the other hand, classical polynomial basis functions do not live in the domain of the linear operator underlying the DDE, which leads to technical complications in establishing convergence results [28,33]; see [9, Remark 2.1-(iii)].

One of the main contributions of [9] consisted in identifying that Koornwinder polynomials [39] lie in the domain  $D(\mathcal{A})$  of linear operators such as  $\mathcal{A}$  given in (2.8), allowing in turn for adopting a classical Trotter-Kato (TK) approximation approach from the  $C_0$ -semigroup theory [25,49] to deal with the ODE approximation of DDEs such as given by Eq. (2.1). The TK approximation approach can be viewed as the functional analysis operator version of the Lax equivalence principle.<sup>2</sup> The work [9] allows thus for positioning the approximation problem of DDEs within a well defined territory. In particular, as pointed out in [11] and discussed in Sect. 2.3 below, the optimal control of DDEs benefits from the TK approximation approach.

In this section, we focus on another important feature pointed out in [9] for applications, namely, Galerkin approximations of DDEs built from Koornwinder polynomials can be efficiently computed via simple analytic formulas; see [9, Sections 5-6 and Appendix C]. We recall below the main elements to do so referring to [9] for more details.

First, let us recall that Koornwinder polynomials  $K_n$  are obtained from Legendre polynomials  $L_n$  according to the relation

$$(2.12) \quad K_n(s) := -(1+s) \frac{d}{ds} L_n(s) + (n^2 + n + 1)L_n(s), \quad s \in [-1, 1], \quad n \in \mathbb{N};$$

see [39, Eq. (2.1)].

Koornwinder polynomials are known to form an orthogonal set for the following weighted inner product with a point-mass on  $[-1, 1]$ ,

$$(2.13) \quad \mu(dx) = \frac{1}{2}dx + \delta_1,$$

where  $\delta_1$  denotes the Dirac point-mass at the right endpoint  $x = 1$ ; see [39]. In other words,

$$(2.14) \quad \begin{aligned} \int_{-1}^1 K_n(s)K_p(s)d\mu(s) &= \frac{1}{2} \int_{-1}^1 K_n(s)K_p(s)ds + K_n(1)K_p(1) \\ &= 0, \quad \text{if } p \neq n. \end{aligned}$$

It is also worthwhile noting that the sequence given by

$$(2.15) \quad \{\mathcal{K}_n := (K_n, K_n(1)) : n \in \mathbb{N}\}$$

forms an orthogonal basis of the product space

$$(2.16) \quad \mathcal{E} := L^2([-1, 1]; \mathbb{R}) \times \mathbb{R},$$

where  $\mathcal{E}$  is endowed with the following inner product:

$$(2.17) \quad \langle (f, a), (g, b) \rangle_{\mathcal{E}} = \frac{1}{2} \int_{-1}^1 f(s)g(s)ds + ab, \quad (f, a), (g, b) \in \mathcal{E}.$$

The norm induced by this inner product will be denoted hereafter by  $\|\cdot\|_{\mathcal{E}}$ .

<sup>2</sup>i.e., if “consistency” and “stability” are satisfied, then “convergence” holds, and reciprocally.

From the original Koornwinder basis given on the interval  $[-1, 1]$ , orthogonal polynomials on the interval  $[-\tau, 0]$  for the inner product (2.3) can now be easily obtained by using a simple linear transformation  $\mathcal{T}$  defined by:

$$(2.18) \quad \mathcal{T}: [-\tau, 0] \rightarrow [-1, 1], \quad \theta \mapsto 1 + \frac{2\theta}{\tau}.$$

Indeed, for  $K_n$  given by (2.12), let us define the rescaled polynomial  $K_n^\tau$  by

$$(2.19) \quad \begin{aligned} K_n^\tau: [-\tau, 0] &\rightarrow \mathbb{R}, \\ \theta &\mapsto K_n\left(1 + \frac{2\theta}{\tau}\right), \quad n \in \mathbb{N}. \end{aligned}$$

As shown in [9], the sequence

$$(2.20) \quad \{\mathcal{K}_n^\tau := (K_n^\tau, K_n^\tau(0)) : n \in \mathbb{N}\}$$

forms an orthogonal basis for the space  $\mathcal{H} = L^2([-\tau, 0]; \mathbb{R}) \times \mathbb{R}$  endowed with the inner product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$  given in (2.3). Note that since  $K_n(1) = 1$  [9, Prop. 3.1], we have

$$(2.21) \quad K_n^\tau(0) = 1.$$

As shown in [9, Sect. 5 & Appendix C], (rescaled) Koornwinder polynomials allow for analytical Galerkin approximations of general nonlinear systems of DDEs. In the case of a nonlinear scalar DDE such as Eq. (2.1), the  $N^{\text{th}}$  Galerkin-Koornwinder (GK) approximation,  $y_N$ , is obtained as

$$(2.22) \quad y_N(t) = \sum_{j=0}^{N-1} \tilde{\xi}_j^N(t) \mathcal{K}_j^\tau, \quad t \geq 0,$$

where the  $\tilde{\xi}_j^N(t)$  solve the  $N$ -dimensional ODE system

$$(2.23) \quad \begin{aligned} \frac{d\tilde{\xi}_j^N}{dt} &= \frac{1}{\|\mathcal{K}_j\|_{\mathcal{E}}^2} \sum_{n=0}^{N-1} \left( a + bK_n(-1) + c\tau(2\delta_{n,0} - 1) \right. \\ &\quad \left. + \frac{2}{\tau} \sum_{k=0}^{n-1} a_{n,k} (\delta_{j,k} \|\mathcal{K}_j\|_{\mathcal{E}}^2 - 1) \right) \tilde{\xi}_n^N(t) \\ &\quad + \frac{1}{\|\mathcal{K}_j\|_{\mathcal{E}}^2} F \left( \sum_{n=0}^{N-1} \tilde{\xi}_n^N(t), \sum_{n=0}^{N-1} \tilde{\xi}_n^N(t) K_n(-1), \tau \tilde{\xi}_0^N(t) - \tau \sum_{n=1}^{N-1} \tilde{\xi}_n^N(t) \right), \\ &0 \leq j \leq N-1. \end{aligned}$$

Here the Kronecker symbol  $\delta_{j,k}$  has been used, and the coefficients  $a_{n,k}$  are obtained by solving a triangular linear system in which the right hand side has explicit coefficients depending on  $n$ ; see [9, Prop. 5.1].

In practice, an approximation  $m_N(t)$  of  $m(t)$  solving Eq. (2.1) is obtained as the state part (at time  $t$ ) of  $y_N$  given by (2.22) which, thanks to the normalization property  $K_n^\tau(0) = 1$  given in (2.21), reduces to

$$(2.24) \quad m_N(t) = \sum_{j=0}^{N-1} \tilde{\xi}_j^N(t).$$

For later usage, we rewrite the above GK system into the following compact form:

$$(2.25) \quad \boxed{\frac{d\tilde{\xi}_N}{dt} = M\tilde{\xi}_N + G(\tilde{\xi}_N),}$$

where  $M\xi_N$  denotes the linear part of Eq. (2.23), and  $G(\xi_N)$  the nonlinear part. Namely,  $M$  is the  $N \times N$  matrix whose elements are given by

$$(2.26) \quad (M)_{j,n} = \frac{1}{\|\mathcal{K}_j\|_{\mathcal{E}}^2} \left( a + bK_n(-1) + c\tau(2\delta_{n,0} - 1) + \frac{2}{\tau} \sum_{k=0}^{n-1} a_{n,k} (\delta_{j,k} \|\mathcal{K}_j\|_{\mathcal{E}}^2 - 1) \right),$$

where  $j, n = 0, \dots, N-1$ , and the nonlinear vector field  $G: \mathbb{R}^N \rightarrow \mathbb{R}^N$ , is given component-wisely by

$$(2.27) \quad G_j(\xi_N) = \frac{1}{\|\mathcal{K}_j\|_{\mathcal{E}}^2} F \left( \sum_{n=0}^{N-1} \xi_n^N(t), \sum_{n=0}^{N-1} \xi_n^N(t) K_n(-1), \tau \xi_0^N(t) - \tau \sum_{n=1}^{N-1} \xi_n^N(t) \right),$$

for each  $0 \leq j \leq N-1$ .

For rigorous convergence results of GK systems (2.25) for Eq. (2.1) when  $F$  does not depend on the delay term  $m(t-\tau)$ , we refer to [9, Sect. 4]. When  $F$  does depend on  $m(t-\tau)$ , convergence is also believed to hold, as supported by the striking numerical results shown in [9, Sect. 6].

### 2.3. Galerkin-Koornwinder approximation of nonlinear optimal control problems associated with DDEs.

2.3.1. *Preliminaries.* In this section, we outline how the recent results of [11] concerned with the Galerkin approximations to optimal control problems governed by nonlinear evolution equations in Hilbert spaces, apply to the context of nonlinear DDEs when these approximations are built from the Koornwinder polynomials. We provide thus here further elements regarding the research program outlined in [11, Sect. 4].

We consider here, given a finite horizon  $T > 0$ , the following controlled version of Eq. (2.7),

$$(2.28) \quad \begin{aligned} \frac{dy}{ds} &= \mathcal{A}y + \mathcal{F}(y) + \mathfrak{C}(u(s)), \quad s \in (t, T], \quad u \in \mathcal{U}_{ad}[t, T], \\ y(t) &= x \in \mathcal{H}. \end{aligned}$$

The (possibly nonlinear) mapping  $\mathfrak{C}: V \rightarrow \mathcal{H}$  is assumed to be such that  $\mathfrak{C}(0) = 0$ , and the control  $u(s)$  lies in a bounded subset  $U$  of a separable Hilbert space  $V$  possibly different from  $\mathcal{H}$ . Assumptions about the set,  $\mathcal{U}_{ad}[t, T]$ , of admissible controls,  $u$ , is made precise below; see (2.36).

Here our Galerkin subspaces are spanned by the rescaled Koornwinder polynomials, namely

$$(2.29) \quad \mathcal{H}_N = \text{span}\{\mathcal{K}_0^\tau, \dots, \mathcal{K}_{N-1}^\tau\}, \quad N = 1, 2, \dots,$$

where  $\mathcal{K}_N^\tau$  is defined in (2.20). Denoting by  $\Pi_N$  the projector of  $\mathcal{H}$  onto  $\mathcal{H}_N$ , the corresponding GK approximation of Eq. (2.28) is then given by

$$(2.30) \quad \begin{aligned} \frac{dy_N}{ds} &= \mathcal{A}_N y_N + \Pi_N \mathcal{F}(y_N) + \Pi_N \mathfrak{C}(u(s)), \quad s \in (t, T], \quad u \in \mathcal{U}_{ad}[t, T], \\ y_N(t) &= \Pi_N x \in \mathcal{H}_N, \end{aligned}$$

where  $\mathcal{A}_N := \Pi_N \mathcal{A} \Pi_N$ .

Given a cost functional assessed along a solution  $y_{t,x}(\cdot; u)$  of Eq. (2.28) driven by  $u$

$$(2.31) \quad J_{t,x}(u) := \int_t^T [\mathcal{G}(y_{t,x}(s; u)) + \mathcal{E}(u(s))] ds, \quad t \in [0, T], \quad u \in \mathcal{U}_{ad}[t, T],$$



(with  $\mathcal{G}$  and  $\mathcal{E}$  to be defined) the focus of [11] was to identify simple checkable conditions that guarantee in—but not limited to—such a context, the convergence of the corresponding value functions, namely

$$(2.32) \quad \lim_{N \rightarrow \infty} \sup_{t \in [0, T]} |v_N(t, \Pi_N x) - v(t, x)| = 0,$$

where

$$(2.33a) \quad v(t, x) := \inf_{u \in \mathcal{U}_{ad}[t, T]} J_{t, x}(u), \quad (t, x) \in [0, T] \times \mathcal{H} \text{ and } v(T, x) = 0,$$

$$(2.33b) \quad v_N(t, x_N) := \inf_{u \in \mathcal{U}_{ad}[t, T]} J_{t, x_N}^N(u), \quad (t, x_N) \in [0, T] \times \mathcal{H}_N \text{ and } v_N(T, x_N) = 0.$$

and  $J_{t, x_N}^N(u)$  denotes the cost functional  $J$  assessed along the Galerkin approximation  $y_{t, x_N}^N(\cdot; u)$  to (2.28) driven by  $u$  and emanating at time  $t$  from

$$(2.34) \quad x_N := \Pi_N x.$$

As shown in [11], conditions ensuring the convergence (2.32) can be grouped into three categories. The first set of conditions deal with the operator  $\mathcal{A}$  and its Galerkin approximation  $\mathcal{A}_N$ . Essentially, it boils down to show that  $\mathcal{A}$  generates a  $C_0$ -semigroup on  $\mathcal{H}$ , and that the Trotter-Kato approximation conditions are satisfied; see Assumptions **(A0)**–**(A2)** in [11, Section 2.1]. The latter conditions are, as mentioned earlier, satisfied in the case of DDEs and GK approximations as shown by [9, Lemmas 4.2 and 4.3].

The second group of conditions identified in [11] is also not restrictive in the sense that only local Lipschitz conditions<sup>3</sup> on  $\mathcal{F}$ ,  $\mathcal{G}$ , and  $\mathcal{C}$  are required as well as continuity of  $\mathcal{E}$  and compactness of  $U$  where  $\mathcal{U}_{ad}[t, T]$  is taken to be

$$(2.35) \quad \mathcal{U}_{ad}[t, T] := \{u|_{[t, T]} : u \in \mathcal{U}_{ad}\},$$

with

$$(2.36) \quad \mathcal{U}_{ad} := \{u \in L^q(0, T; V) : u(s) \in U \text{ for a.e. } s \in [0, T]\}, \quad q \geq 1.$$

Also required to hold, are a priori bounds—uniform in  $u$  in  $\mathcal{U}_{ad}$ —for the solutions of Eq. (2.28) as well as of their Galerkin approximations. Depending on the specific form of Eq. (2.28) such bounds can be derived for a broad class of DDEs; in that respect the proofs of [9, Estimates (4.75)] and [9, Corollary 4.3] can be easily adapted to the case of controlled DDEs.

Finally, the last condition to ensure (2.32), requires that the residual energy of the solution  $s \mapsto y(s; x, u)$  of (2.28) (with  $y(0) = x$ ) driven by  $u$ , satisfies

$$(2.37) \quad \|\Pi_N^\perp y(s; x, u)\|_{\mathcal{H}} \longrightarrow 0,$$

uniformly with respect to both the control  $u$  in  $\mathcal{U}_{ad}$  and the time  $s$  in  $[0, T]$ ; see Assumption **(A7)** in [11, Section 2.4]. Here,

$$(2.38) \quad \Pi_N^\perp := \text{Id}_{\mathcal{H}} - \Pi_N.$$

Easily checkable conditions ensuring such a double uniform convergence are identified in [11, Section 2.7] for a broad class of evolution equations and their Galerkin approximations but unfortunately do not apply to the case of the GK approximations considered here. The scope of this article does not allow for an investigation of such conditions in the case of DDEs, and results along this direction will be communicated elsewhere.

<sup>3</sup>Note however that in the case of DDEs, nonlinearities  $\mathcal{F}$  that include discrete delays may complicate the verification of a local Lipschitz property in  $\mathcal{H}$  as given by (2.2). The case of nonlinearities  $\mathcal{F}$  depending on distributed delays and/or the current state can be handled in general more easily; see [9, Sect. 4.3.2].

Nevertheless, error estimates can be still derived in the case of DDEs by adapting elements provided in [11, Section 2.6]. We turn now to the formulation of such error estimates.

2.3.2. *Error estimates.* Our aim is to derive error estimates for GK approximations to the following type of optimal control problem associated with DDEs (2.1) framed into the abstract form (2.7),

$$(P) \quad \min J(x, u) \quad \text{s.t.} \quad (y, u) \in L^2(0, T; \mathcal{H}) \times \mathcal{U}_{ad} \text{ solves (2.28),}$$

in which  $J$  is given by (2.31) with  $t = 0$ , and  $y$  solves (2.28) with  $y(0) = x$  in  $\mathcal{H}$ .

To do so, we consider the following set of assumptions

- (i) The mappings  $\mathcal{F} : \mathcal{H} \rightarrow \mathcal{H}$ ,  $\mathcal{G} : \mathcal{H} \rightarrow \mathbb{R}^+$  are locally Lipschitz, and  $\mathcal{E} : V \rightarrow \mathbb{R}^+$  is continuous.
- (ii) The linear operator  $\mathcal{A}_N : \mathcal{H}_N \rightarrow \mathcal{H}_N$  is diagonalizable over  $\mathbb{C}$  and there exists  $\alpha$  in  $\mathbb{R}$  such that for each  $N$

$$(2.39) \quad \text{Re } \lambda_j^N \leq \alpha, \quad \forall j \in \{1, \dots, N\},$$

where  $\{\lambda_j^N\}$  denotes the set of (complex) eigenvalues of  $\mathcal{A}_N$ .

- (iii) Let  $\mathcal{U}_{ad}$  be given by (2.36) with  $U$  bounded in  $V$ . For each  $T > 0$ ,  $(x, u)$  in  $\mathcal{H} \times \mathcal{U}_{ad}$ , the problem (2.28) (with  $t = 0$ ) admits a unique mild solution  $y(\cdot; x, u)$  in  $C([0, T], \mathcal{H})$ , and for each  $N \geq 0$ , its GK approximation (2.30) admits a unique solution  $y_N(\cdot; x_N, u)$  in  $C([0, T], \mathcal{H})$ . Moreover, there exists a constant  $\mathcal{C} := \mathcal{C}(T, x)$  such that

$$(2.40) \quad \begin{aligned} \|y(t; x, u)\|_{\mathcal{H}} &\leq \mathcal{C}, & \forall t \in [0, T], u \in \mathcal{U}_{ad}, \\ \|y_N(t; x_N, u)\|_{\mathcal{H}} &\leq \mathcal{C}, & \forall t \in [0, T], N \in \mathbb{N}^*, u \in \mathcal{U}_{ad}. \end{aligned}$$

- (iv) The mild solution  $y(\cdot; x, u)$  belongs to  $C([0, T], D(\mathcal{A}))$  if  $x$  lives in  $D(\mathcal{A})$ .

**Remark 2.1.**

- (i) Compared with the case of eigen-subspaces considered in [11, Section 2.6], the main difference here lies in the regularity assumption (iv) above. This assumption is used to handle the term  $\Pi_N \mathcal{A} \Pi_N^\perp y$  arising in the error estimates; see e.g. (2.48) below. Note that this term is zero when  $\mathcal{A}$  is self-adjoint and  $\mathcal{H}_N$  is an eigen-subspace of  $\mathcal{A}$ , which is the setting considered in [11, Section 2.6].
- (ii) A way to ensure Condition (iv) to hold consists of proving that the mild solution belongs to  $C^1([0, T], \mathcal{H})$ . For conditions on  $\mathcal{F}$  ensuring such a regularity see e.g. [6, Theorem 3.2].

Finally, we introduce the cost functional  $J_N$  to be the cost functional  $J$  assessed along the Galerkin approximation  $y_N$  solving (2.30) with  $y_N(0) = x_N \in \mathcal{H}_N$ , namely

$$(2.41) \quad J_N(x_N, u) := \int_0^T [\mathcal{G}(y_N(s; x_N, u)) + \mathcal{E}(u(s))] ds,$$

along with the optimal control problem

$$(P_N) \quad \min J_N(x_N, u) \quad \text{s.t.} \quad (y_N, u) \in L^2(0, T; \mathcal{H}_N) \times \mathcal{U}_{ad} \text{ solves (2.30),}$$

$$\text{with } y_N(0) = x_N \in \mathcal{H}_N.$$

We are now in position to derive error estimates as formulated in Theorem 2.1 below. For that purpose Table 1 provides a list of the main symbols necessary to a good reading of the proof and statement of this theorem. The proof is largely inspired from that of [11, Theorem 2.3]; the main amendments being specified within this proof.

**Theorem 2.1.** *Assume that assumptions (i)-(iv) hold. Assume furthermore that for each  $(t, x)$  in  $[0, T] \times D(\mathcal{A})$ , there exists a minimizer  $u_{t,x}^*$  (resp.  $u_{t,x_N}^{N,*}$ ) for the value function  $v$  (resp.  $v_N$ ) defined in (2.33).*

TABLE 1. Glossary of principal symbols

Symbol	Terminology
$(y^*, u^*)$	An optimal pair of the optimal control problem ( $\mathcal{P}$ )
$u_{t,x}^*$	Minimizer of the value function $v$ defined in (2.33)
$u_{t,x_N}^{N,*}$	Minimizer of the value function $v_N$ defined in (2.33)
$y_{t,x}(\cdot; u_{t,x}^*)$	Mild solution to (2.28) driven by $u_{t,x}^*$
$y_{t,x}(\cdot; u_{t,x_N}^{N,*})$	Mild solution to (2.28) driven by $u_{t,x_N}^{N,*}$
$\mathfrak{B}$	The closed ball in $\mathcal{H}$ centered at $0_{\mathcal{H}}$ with radius $\mathcal{C}$ given by (2.40)
$\ \mathcal{A}\ _{\text{op}}$	Norm of $\mathcal{A}$ as a linear bounded operator from $D(\mathcal{A})$ to $\mathcal{H}$

Then for any  $(t, x)$  in  $[0, T] \times D(\mathcal{A})$  it holds that

$$(2.42) \quad \begin{aligned} & |v(t, x) - v_N(t, x_N)| \\ & \leq \text{Lip}(\mathcal{G}|_{\mathfrak{B}}) \left[ \sqrt{T-t} + (T-t)\sqrt{f(T)} \right] \mathcal{R}(u_{t,x}^*, u_{t,x_N}^{N,*}), \end{aligned}$$

with

$$(2.43) \quad f(T) := \left( \text{Lip}(\mathcal{F}|_{\mathfrak{B}}) + \|\mathcal{A}\|_{\text{op}}^2 \right) \exp \left( 2 \left[ \alpha + \frac{3}{2} \text{Lip}(\mathcal{F}|_{\mathfrak{B}}) + \frac{1}{2} \right] T \right),$$

and

$$(2.44) \quad \mathcal{R}(u_{t,x}^*, u_{t,x_N}^{N,*}) := \|\Pi_N^\perp y_{t,x}(\cdot; u_{t,x}^*)\|_{L^2(t,T;D(\mathcal{A}))} + \|\Pi_N^\perp y_{t,x}(\cdot; u_{t,x_N}^{N,*})\|_{L^2(t,T;D(\mathcal{A}))}.$$

*Proof.* We provide here the main elements of the proof that needs to be amended from [11, Theorem 2.3]. First note that mild solutions to (2.28) lie in  $C([0, T], D(\mathcal{A}))$  due to Assumption (iv) and since the initial datum  $x$  is taken in  $D(\mathcal{A})$ .

We want to prove that for each  $N$ , there exists a constant  $\beta > 0$  such that for any  $(t, x) \in [0, T] \times D(\mathcal{A})$ , and  $u \in \mathcal{U}_{\text{ad}}[t, T]$ ,

$$(2.45) \quad \|\Pi_N y_{t,x}(s; u) - y_{t,x_N}^N(s; u)\|_{\mathcal{H}}^2 \leq f(T) \int_t^s \|\Pi_N^\perp y_{t,x}(s'; u)\|_{D(\mathcal{A})}^2 ds', \quad s \in [t, T].$$

Let us introduce

$$(2.46) \quad w(s) := \Pi_N y_{t,x}(s; u) - y_{t,x_N}^N(s; u).$$

By applying  $\Pi_N$  to both sides of Eq. (2.28), we obtain that  $\Pi_N y_{t,x}(\cdot; u)$  (denoted by  $\Pi_N y$  to simplify), satisfies:

$$\frac{d\Pi_N y}{ds} = \mathcal{A}_N \Pi_N y + \Pi_N \mathcal{A} \Pi_N^\perp y + \Pi_N \mathcal{F}(\Pi_N y + \Pi_N^\perp y) + \Pi_N \mathfrak{C}(u(s)).$$

This together with (2.30) implies that  $w$  satisfies the following problem:

$$(2.47) \quad \begin{aligned} \frac{dw}{ds} &= \mathcal{A}_N w + \Pi_N \mathcal{A} \Pi_N^\perp y + \Pi_N \mathcal{F}(\Pi_N y + \Pi_N^\perp y) - \Pi_N F(y_N), \\ w(t) &= 0. \end{aligned}$$

By taking the  $\mathcal{H}$ -inner product on both sides of (2.47) with  $w$ , we obtain:

$$(2.48) \quad \begin{aligned} \frac{1}{2} \frac{d\|w\|_{\mathcal{H}}^2}{ds} &= \langle \mathcal{A}_N w, w \rangle_{\mathcal{H}} + \langle \Pi_N \mathcal{A} \Pi_N^\perp y, w \rangle_{\mathcal{H}} \\ &\quad + \langle \Pi_N (\mathcal{F}(\Pi_N y + \Pi_N^\perp y) - \mathcal{F}(y_N)), w \rangle_{\mathcal{H}}. \end{aligned}$$

We estimate now the term  $\langle \mathcal{A}_N w, w \rangle_{\mathcal{H}}$  in the above equation using Assumption (ii). For this purpose, note that for any  $p$  in  $\mathcal{H}_N$ , the following identity holds:

$$(2.49) \quad \langle \mathcal{A}_N p, p \rangle_{\mathcal{H}} = \langle M \tilde{\zeta}, \tilde{\zeta} \rangle_{\mathbb{R}^n},$$

where  $M$  is the matrix representation of  $\mathcal{A}_N$  under the Koornwinder basis given by (2.26), and  $\tilde{\zeta}$  denotes the column vector  $(\tilde{\zeta}_0, \dots, \tilde{\zeta}_{N-1})^{\text{tr}}$  whose entries are given by

$$\tilde{\zeta}_j := \frac{1}{\|\mathcal{K}_j\|_{\mathcal{E}}} \langle p, \mathcal{K}_j^{\text{r}} \rangle_{\mathcal{H}}, \quad j = 0, \dots, N-1.$$

Note that the  $N \times N$  matrix  $M$  has the same eigenvalues  $\{\lambda_j^N : j = 1, N\}$  as  $\mathcal{A}_N$ . Thanks to Assumption (ii),  $M$  is diagonalizable over  $\mathbb{C}$ . Denoting by  $f_j$  the normalized eigenvector of  $M$  corresponding to each  $\lambda_j^N$ , we have

$$M \tilde{\zeta} = M \sum_{j=1}^N \tilde{\zeta}_j f_j = \sum_{j=1}^N \lambda_j^N \tilde{\zeta}_j f_j,$$

where  $\tilde{\zeta}_j = \langle \tilde{\zeta}, f_j \rangle_{\mathbb{C}^n}$ . It follows then

$$\langle M \tilde{\zeta}, \tilde{\zeta} \rangle_{\mathbb{R}^n} = \langle \sum_{j=1}^N \lambda_j^N \tilde{\zeta}_j f_j, \tilde{\zeta} \rangle_{\mathbb{C}^n} = \sum_{j=1}^N \lambda_j^N |\tilde{\zeta}_j|^2 = \sum_{j=1}^N \text{Re } \lambda_j^N |\tilde{\zeta}_j|^2,$$

for which the latter equality holds since  $\tilde{\zeta} \mapsto \langle M \tilde{\zeta}, \tilde{\zeta} \rangle_{\mathbb{R}^n}$  is real-valued. Due to the bound (2.39), we have thus

$$(2.50) \quad \langle M \tilde{\zeta}, \tilde{\zeta} \rangle_{\mathbb{R}^n} \leq \alpha \sum_{j=1}^N |\tilde{\zeta}_j|^2.$$

On the other hand, by noting that

$$(2.51) \quad |\tilde{\zeta}_j|^2 = \langle \tilde{\zeta}, \tilde{\zeta} \rangle_{\mathbb{C}^n} = \langle \tilde{\zeta}, \tilde{\zeta} \rangle_{\mathbb{R}^n} = \langle p, p \rangle.$$

we obtain from (2.49)–(2.51), by taking  $p = w$ , that

$$(2.52) \quad \langle \mathcal{A}_N w(s), w(s) \rangle_{\mathcal{H}} \leq \alpha \|w(s)\|_{\mathcal{H}}^2, \quad s \in [t, T].$$

We infer from (2.48) that

$$\begin{aligned} \frac{1}{2} \frac{\text{d} \|w(s)\|_{\mathcal{H}}^2}{\text{d}s} &\leq \left( \alpha + \frac{3}{2} \text{Lip}(\mathcal{F}|_{\mathfrak{B}}) + \frac{1}{2} \right) \|w(s)\|_{\mathcal{H}}^2 \\ &\quad + \frac{1}{2} \|\mathcal{A}\|_{\text{op}}^2 \|\Pi_N^{\perp} y\|_{D(\mathcal{A})}^2 + \frac{1}{2} \text{Lip}(\mathcal{F}|_{\mathfrak{B}}) \|\Pi_N^{\perp} y\|_{\mathcal{H}}^2, \end{aligned}$$

in which we have also used Assumption (iii) and the local Lipschitz property of  $\mathcal{F}$  on the closed ball  $\mathfrak{B}$  in  $\mathcal{H}$  centered at  $0_{\mathcal{H}}$  with radius  $\mathcal{C}$  given by (2.40).

By Gronwall's lemma and recalling that  $w(t) = 0$ , we obtain thus

$$\|\Pi_N y_{t,x}(s; u) - y_{t,x_N}^N(s; u)\|_{\mathcal{H}}^2 \leq f(T) \int_t^s \|\Pi_N^{\perp} y_{t,x}(s'; u)\|_{D(\mathcal{A})}^2 \text{d}s',$$

with

$$f(T) = \left( \text{Lip}(\mathcal{F}|_{\mathfrak{B}}) + \|\mathcal{A}\|_{\text{op}}^2 \right) \exp \left( 2 \left[ \alpha + \frac{3}{2} \text{Lip}(\mathcal{F}|_{\mathfrak{B}}) + \frac{1}{2} \right] T \right).$$

Then by noting that

$$|J_{t,x}(u) - J_{t,x_N}^N(u)| \leq \text{Lip}(\mathcal{G}|_{\mathfrak{B}}) \int_t^T \left( \|\Pi_N y_{t,x}(s; u) - y_{t,x_N}^N(s; u)\|_{D(\mathcal{A})} + \|\Pi_N^\perp y_{t,x}(s; u)\|_{D(\mathcal{A})} \right) ds,$$

we have

$$(2.53) \quad |J_{t,x}(u) - J_{t,x_N}^N(u)| \leq \text{Lip}(\mathcal{G}|_{\mathfrak{B}}) \left[ \sqrt{T-t} + (T-t)\sqrt{f(T)} \right] \|\Pi_N^\perp y_{t,x}(\cdot; u)\|_{L^2(t,T;D(\mathcal{A}))},$$

for all  $u \in \mathcal{U}_{ad}[t, T]$ .

Finally by noting that

$$v(t, x) = J_{t,x}(u_{t,x}^*) \leq J_{t,x}(u_{t,x_N}^{N,*}) \quad \text{and} \quad v_N(t, x_N) = J_{t,x_N}^N(u_{t,x_N}^{N,*}),$$

we obtain

$$v(t, x) - v_N(t, x_N) \leq J_{t,x}(u_{t,x_N}^{N,*}) - J_{t,x_N}^N(u_{t,x_N}^{N,*}).$$

It follows then from (2.53) that

$$(2.54) \quad v(t, x) - v_N(t, x_N) \leq \text{Lip}(\mathcal{G}|_{\mathfrak{B}}) \left[ \sqrt{T-t} + (T-t)\sqrt{f(T)} \right] \|\Pi_N^\perp y_{t,x}(\cdot; u_{t,x_N}^{N,*})\|_{L^2(t,T;D(\mathcal{A}))}.$$

Similarly,

$$(2.55) \quad v_N(t, x_N) - v(t, x) \leq \text{Lip}(\mathcal{G}|_{\mathfrak{B}}) \left[ \sqrt{T-t} + (T-t)\sqrt{f(T)} \right] \|\Pi_N^\perp y_{t,x}(\cdot; u_{t,x}^*)\|_{L^2(t,T;D(\mathcal{A}))}.$$

The estimate (2.42) results then from (2.54) and (2.55). □

We conclude this section with the following corollary providing the error estimates between the optimal control and that obtained from a GK approximation.

**Corollary 2.1.** *Assume that the conditions given in Theorem 2.1 hold. Given  $x$  in  $\mathcal{H}$ , let us introduce the notations,  $u^* := u_{0,x}^*$  and  $u_N^* := u_{0,x_N}^{N,*}$ . Assume furthermore that there exists  $\sigma > 0$  such that the following local growth condition is satisfied for the cost functional  $J$  (with  $t = 0$ ) given by (2.31):*

$$(2.56) \quad \sigma \|u^* - v\|_{L^q(0,T;V)}^q \leq J(x, v) - J(x, u^*),$$

for all  $v$  in some neighborhood  $\mathcal{O} \subset \mathcal{U}_{ad}$  of  $u^*$ , with  $\mathcal{U}_{ad}$  given by (2.36). Assume finally that  $u_N^*$  lies in  $\mathcal{O}$ . Then,

$$(2.57) \quad \begin{aligned} & \|u^* - u_N^*\|_{L^q(0,T;V)}^q \\ & \leq \frac{1}{\sigma} \text{Lip}(\mathcal{G}|_{\mathfrak{B}}) \left[ \sqrt{T} + T\sqrt{f(T)} \right] \left( \|\Pi_N^\perp y^*(\cdot; u^*)\|_{L^2(0,T;D(\mathcal{A}))} \right. \\ & \quad \left. + 2\|\Pi_N^\perp y(\cdot; u_N^*)\|_{L^2(0,T;D(\mathcal{A}))} \right), \end{aligned}$$

where the constant  $f(T)$  is given by (2.43).

*Proof.* By the assumptions, we have

$$\|u^* - u_N^*\|_{L^q(0,T;V)}^q \leq \frac{1}{\sigma} (J(x, u_N^*) - J(x, u^*)).$$

Note also that

$$\begin{aligned} J(x, u_N^*) - J(x, u^*) &= J(x, u_N^*) - J_N(x_N, u_N^*) + J_N(x_N, u_N^*) - J(x, u^*) \\ &= J_{0,x}(u_N^*) - J_{0,x_N}^N(u_N^*) + v_N(0, x_N) - v(0, x), \end{aligned}$$

where we used the fact that

$$\begin{aligned} J(x, u_N^*) &= J_{0,x}(u_N^*), & J_N(x_N, u_N^*) &= J_{0,x_N}^N(u_N^*), \\ J_N(x_N, u_N^*) &= v_N(0, x_N) & \text{and} & \quad J(x, u^*) = v(0, x). \end{aligned}$$

The result follows by applying the estimate (2.53) to  $J_{0,x}(u_N^*) - J_{0,x_N}^N(u_N^*)$  and the estimate (2.42) to  $v_N(0, x_N) - v(0, x)$ . □

### 3. APPLICATION TO AN OSCILLATION AMPLITUDE REDUCTION PROBLEM

**3.1. Optimal control of the Wright equation.** As an application, we consider the Wright equation

$$(3.1) \quad \frac{dm}{dt} = -m(t - \tau)(1 + m(t)),$$

where  $m$  is the unknown scalar function, and  $\tau$  is a nonnegative delay parameter. This equation has been studied by numerous authors, and notably among the first ones, are Jones [29, 30], Kakutani and Markus [31], and Wright [61]. This equation can also be transformed via a simple change of variable [53] into the well-known Hutchinson equation [27] arising in population dynamics. It corresponds then to the logistic equation with a delay effect introduced into the *intraspecific competition term*, namely with the change of variable  $m = -1 + p$ , Eq. (3.1) becomes

$$(3.2) \quad \frac{dp}{dt} = p(t)(1 - p(t - \tau)).$$

Essentially the idea of Hutchinson [27] was to point out that negative effects that high population density  $p$  has on the environment, influences birth rates at later times due to developmental and maturation delays, justifying thus Eq. (3.2).

As a result, the solutions of Eq. (3.1) are obtained from a simple shift of the solutions of Eq. (3.2), and reciprocally. Since Eq. (3.2) benefits of a more intuitive interpretation, we will often prefer to think about Eq. (3.1) in terms of Eq. (3.2).

Anyway, it is known that Eq. (3.1) (and thus Eq. (3.2)) experiences a supercritical Hopf bifurcation as the delay parameter is varied and crosses a critical value,  $\tau_c$ , from below; see e.g. [17, Sect. 9] and [57].<sup>4</sup>

Figure 1 shows—in the embedded reduced phase space  $(m(t), m(t - \tau))$ —the corresponding bifurcation of the limit cycle unfolding from the trivial steady state, as  $\tau$  is varied<sup>5</sup>. As Fig. 1 illustrates, the amplitude of the oscillation that takes place via the Hopf bifurcation, increases as  $\tau$  is increased away from  $\tau_c$ .

<sup>4</sup>The critical value  $\tau_c$  at which the trivial steady state of the Wright equation (3.1) loses its linear stability is given by  $\tau_c = \frac{\pi}{2}$ . This can be found by analyzing the associated linear eigenvalue problem  $-\phi(\theta - \tau) = \beta\phi(\theta)$ ,  $\theta \in [-\tau, 0]$ . By using the ansatz  $\phi(\theta) = e^{\beta\theta}$ , we obtain  $-e^{-\beta\tau} = \beta$ . Assuming that  $\beta = \pm i\omega$ , we get  $-\cos(\omega\tau) + i\sin(\omega\tau) = i\omega$ , leading thus to  $\omega = 1$  and  $\tau = \frac{(2n+1)\pi}{2}$ ,  $n = 0, 1, 2, \dots$ . Consequently, the critical delay parameter is  $\tau_c = \frac{\pi}{2}$ .

<sup>5</sup>Such an embedded phase space is classically used to visualize attractors associated with DDEs; see e.g. [15].

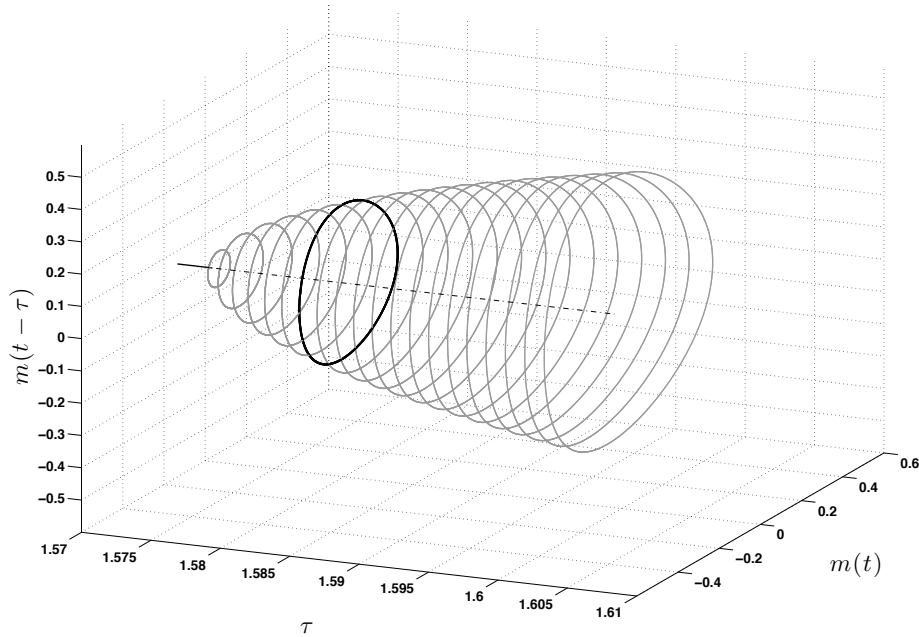


FIGURE 1. Unfolding of the Hopf bifurcation taking place for Eq. (3.1) as  $\tau$  is varied. The limit cycles are represented in the embedded reduced phase space  $(m(t), m(t - \tau))$ . The limit cycle in heavy black indicates the “ $\tau$ -section” considered for the numerical results of Sectns. 3.3 and 4 below. The trivial steady state is represented via a dotted line when unstable and a plain one when stable.

Since fluctuating populations are susceptible to extinction due to sudden and unforeseen environmental disturbances [51], a knowledge of the conditions for which the population is resilient to these disturbances is of great interest in planning and designing control as well as management strategies; see e.g. [14, 50, 52].

Given a convex functional depending on  $m$  and a control to be designed, our goal is to show (for Eq. (3.1)) that relatively close to the criticality (i.e.  $\tau \approx \tau_c$ ), the oscillation amplitude can be reduced at a nearly optimal cost, by solving efficiently low-dimensional optimal control problems.

For that purpose, we consider, given  $u$  in  $L^2(0, T; \mathbb{R})$  and  $T > 0$ , the following forced version of Eq. (3.1)

$$(3.3) \quad \frac{dm}{dt} = -m(t - \tau)(1 + m(t)) + u(t), \quad t \in (0, T),$$

supplemented by the initial history  $m(t) = \phi(t)$  (with  $\phi \in L^2(-\tau, 0; \mathbb{R})$ ) for  $t$  in the interval  $[-\tau, 0)$ , and by  $m(0) = m_0$  (in  $\mathbb{R}$ ) at  $t = 0$ . In Eq. (3.3),  $u(t)$  can be thought as an environmental management strategy. Our goal is to understand the management strategies that may lead to a reduction of the oscillation amplitude of the population (and possibly a stabilization); i.e. to determine  $u$  for which  $m \approx 0$  and thus  $p \approx 1$ , while dealing with limited resources to plan such strategies.

Naturally, we introduce thus the following cost functional

$$(3.4) \quad J(m, v) := \int_0^T \left[ \frac{1}{2} m(t)^2 + \frac{\mu}{2} u(t)^2 \right] dt, \quad \mu > 0.$$

We are interested in solving optimal control problem associated with Eq. (3.3) for this cost functional. To address this problem, we recast first this optimal control problem into the abstract framework of Sect. 2.1.

In that respect, we introduce the variables

$$(3.5) \quad y := (m_t, m_t(0)), \quad \text{with } m \text{ solving (3.3),}$$

and we take  $\mathcal{U}_{ad}$  defined in (2.36) with  $q = 2$  and with  $U$  to be (possibly) a bounded subset of  $V := \mathbb{R}$ .

The control operator  $\mathfrak{C}: V \rightarrow \mathcal{H}$  appearing in Eq. (2.28) is taken here to be linear and given by

$$(3.6) \quad [\mathfrak{C}v](\theta) := \begin{cases} 0, & \theta \in [-\tau, 0) \\ v, & \theta = 0 \end{cases}, \quad v \in V.$$

As explained in Sect. 2.1, Eq. (3.3) can be recast into the following abstract ODE posed on  $\mathcal{H}$

$$(3.7) \quad \frac{dy}{dt} = \mathcal{A}y + \mathcal{F}(y) + \mathfrak{C}u(t),$$

with the linear operator  $\mathcal{A}$  given in (2.8) that, in the case of Eq. (3.3), takes the form

$$(3.8) \quad [\mathcal{A}\Psi](\theta) := \begin{cases} \frac{d^+\Psi^D}{d\theta}, & \theta \in [-\tau, 0), \\ -\Psi^D(-\tau), & \theta = 0, \end{cases}$$

with domain given in (2.9). The nonlinearity  $\mathcal{F}$  given in (2.10) takes here the form,

$$(3.9) \quad [\mathcal{F}(\Psi)](\theta) := \begin{cases} 0, & \theta \in [-\tau, 0), \\ -\Psi^D(-\tau)\Psi^S, & \theta = 0, \end{cases} \quad \forall \Psi = (\Psi^D, \Psi^S) \in \mathcal{H}.$$

Let  $\Phi$  be in  $\mathcal{H}$  given by

$$(3.10) \quad \Phi := (\phi, m_0), \quad \text{with } \phi \in L^2(-\tau, 0; \mathbb{R}), \quad m_0 \in \mathbb{R}.$$

We rewrite now the cost functional defined in (3.4) by using the state-variable  $y$  given in (3.5), and define thus

$$(3.11) \quad J(y, u) := \int_0^T \left[ \frac{1}{2} (m_t(0))^2 + \frac{\mu}{2} u(t)^2 \right] dt;$$

still denoted by  $J$ .

The optimal control problem associated with Eq. (3.7) becomes then

$$(3.12) \quad \min J(\Phi, u) \text{ s.t. } (y, u) \in L^2(0, T; \mathcal{H}) \times \mathcal{U}_{ad} \text{ solves Eq. (3.7) with } y(0) = \Phi \in \mathcal{H},$$

with  $\mathcal{U}_{ad}$  defined in (2.36) with  $q = 2$  and with either  $U = \mathbb{R}$  or  $U$  to be a bounded subset of  $V = \mathbb{R}$ .

**Remark 3.1.** Note that in the case  $u(t)$  in Eq. (3.3) is replaced by  $u(t-r)$  (with  $0 < r$ ), one can still recast Eq. (3.3) into an abstract ODE of the form (3.7); i.e. to deal with control with delay. In that case by taking  $V := H^1(-r, 0; \mathbb{R}) \times \mathbb{R}$ , the operator  $\mathfrak{C}$  becomes

$$(3.13) \quad [\mathfrak{C}(\phi^D, \phi^S)](\theta) := \begin{cases} 0, & \theta \in [-\tau, 0) \\ \phi^D(-r), & \theta = 0 \end{cases}, \quad (\phi^D, \phi^S) \in V.$$

This reformulation is compatible with the framework of [9] and thus allow us to derive GK approximations of such abstract ODEs. We leave the associated optimal control problems to the interested reader.



**3.2. Galerkin-Koornwinder approximation of the optimal control problem.** First, we specify for Eq. (3.3), the (concrete) ODE version (2.23) of the (abstract)  $N^{\text{th}}$  GK approximation (2.30). For this purpose, we first compute  $\Pi_N \mathfrak{C}u(s)$ .

Note that since  $\{\mathcal{K}_j^\tau : j \in \mathbb{N}\}$  forms a Hilbert basis of  $\mathcal{H}$ , for any  $\Psi = (\Psi^D, \Psi^S)$  in  $\mathcal{H}$ , we have (see also [9, Eq. (3.25)])

$$\begin{aligned} \Psi &= \sum_{j=0}^{\infty} \frac{\langle \Psi, \mathcal{K}_j^\tau \rangle_{\mathcal{H}}}{\|\mathcal{K}_j^\tau\|_{\mathcal{H}}^2} \mathcal{K}_j^\tau \\ (3.14) \quad &= \sum_{j=0}^{\infty} \left( \frac{1}{\tau} \langle \Psi^D, K_j^\tau \rangle_{L^2} + \Psi^S K_j^\tau(0) \right) \frac{\mathcal{K}_j^\tau}{\|\mathcal{K}_j^\tau\|_{\mathcal{H}}^2}. \end{aligned}$$

By taking now  $\Psi = \mathfrak{C}v$  with  $\mathfrak{C}$  given by (3.6) and  $v$  in  $\mathbb{R}$ , we have

$$(3.15) \quad \mathfrak{C}v = \sum_{j=0}^{\infty} \left( v K_j^\tau(0) \right) \frac{\mathcal{K}_j^\tau}{\|\mathcal{K}_j^\tau\|_{\mathcal{H}}^2},$$

which leads to

$$(3.16) \quad \Pi_N \mathfrak{C}v = v \sum_{j=0}^{N-1} \frac{\mathcal{K}_j^\tau}{\|\mathcal{K}_j^\tau\|_{\mathcal{H}}^2},$$

recalling that  $K_j^\tau(0) = 1$ ; see (2.21).

Then by defining  $\mathfrak{C}_N: V \rightarrow \mathbb{R}^N$  to be

$$(3.17) \quad \mathfrak{C}_N v := \left( \frac{1}{\|\mathcal{K}_0\|_{\mathcal{E}}^2}, \dots, \frac{1}{\|\mathcal{K}_{N-1}\|_{\mathcal{E}}^2} \right)^{\text{tr}} v,$$

one can write the Galerkin approximation (2.30) in the Koornwinder basis as the ODE system

$$\begin{aligned} (3.18) \quad \frac{d\boldsymbol{\zeta}_N}{dt} &= M\boldsymbol{\zeta}_N + G(\boldsymbol{\zeta}_N) + \mathfrak{C}_N u(t), \\ \boldsymbol{\zeta}_N(0) &= \boldsymbol{\zeta}_N \in \mathbb{R}^N. \end{aligned}$$

Here  $\boldsymbol{\zeta}_N = (\zeta_0^N, \dots, \zeta_{N-1}^N)^{\text{tr}}$ ; the  $N \times N$ -matrix  $M$  given in (2.26) has its elements given here by

$$(3.19) \quad (M)_{j,n} = \frac{1}{\|\mathcal{K}_j\|_{\mathcal{E}}^2} \left( -K_n(-1) + \frac{2}{\tau} \sum_{k=0}^{n-1} a_{n,k} (\delta_{j,k} \|\mathcal{K}_j\|_{\mathcal{E}}^2 - 1) \right), \quad j, n = 0, \dots, N-1;$$

Recall that here the Kronecker symbol  $\delta_{j,k}$  has been used, and the coefficients  $a_{n,k}$  are obtained by solving a triangular linear system in which the right hand side has explicit coefficients depending on  $n$ ; see [9, Prop. 5.1].

The initial datum  $\boldsymbol{\zeta}_N := (\zeta_0^N, \dots, \zeta_{N-1}^N)^{\text{tr}}$  has its components defined by

$$(3.20) \quad \zeta_j^N = \frac{1}{\|\mathcal{K}_j\|_{\mathcal{E}}^2} \langle \Phi, \mathcal{K}_j^\tau \rangle_{\mathcal{H}}, \quad j = 0, \dots, N-1,$$

where  $\Phi$  in  $\mathcal{H}$  denotes the initial data given by (3.10) for the abstract ODE (3.7).

The nonlinearity  $G$  given in (2.27) takes, in the case of Eq. (3.1), the form

$$(3.21) \quad G_j(\boldsymbol{\zeta}_N) = -\frac{1}{\|\mathcal{K}_j\|_{\mathcal{E}}^2} \left[ \sum_{n=0}^{N-1} \zeta_n^N(t) \right] \left[ \sum_{n=0}^{N-1} \zeta_n^N(t) K_n(-1) \right], \quad 0 \leq j \leq N-1.$$

We turn next to the rewriting—using the variables  $\zeta_n^N$ —of the corresponding cost functional  $J_N$  given in its abstract version by (2.41). To do so, it is sufficient to recall from (2.24) that the  $N$ -dim GK approximation  $m_N$  is given by

$$m_N(t) = \sum_{n=0}^{N-1} \zeta_n^N(t) K_n^\tau(0) = \sum_{n=0}^{N-1} \zeta_n^N(t),$$

which, denoting still by  $J_N$  the rewritten cost functional, gives

$$(3.22) \quad J_N(\zeta_N, u) := \int_0^T \left[ \frac{1}{2} \left( \sum_{n=0}^{N-1} \zeta_n^N(t; \zeta_N, u) \right)^2 + \frac{\mu}{2} u(t)^2 \right] dt, \quad \zeta_N \in \mathbb{R}^N.$$

We are now in position to write the corresponding Galerkin approximations of the initial optimal control problem (3.12). More precisely, the (reduced) optimal control problem associated with the  $N$ -dimensional GK approximation (3.18) of Eq. (3.3), is

$$(3.23) \quad \boxed{\begin{array}{l} \min J_N(\zeta_N, u) \text{ subject to } (\zeta_N(\cdot; \zeta_N, u), u) \in L^2(0, T; \mathbb{R}^N) \times \mathcal{U}_{ad} \\ \text{solves the ODE system (3.18) supplemented with } \zeta_N(0) = \zeta_N, \end{array}}$$

with  $\mathcal{U}_{ad}$  defined in (2.36) with  $q = 2$  and with either  $U = \mathbb{R}$  or  $U$  to be a bounded subset of  $V = \mathbb{R}$ ; see Sect. 4.

**3.3. Numerical results.** As mentioned earlier, the delay parameter  $\tau$  in Eq. (3.1) is taken to be slightly above its critical value  $\tau_c$  at which the supercritical Hopf bifurcation takes place; see again Figure 1. In particular, the  $\tau$ -value is selected so that there is only one conjugate pair of unstable eigenmodes for the linearized DDE. We can thus hope for relatively low-dimensional Galerkin systems aimed at the synthesis of (sub)optimal controls at a nearly optimal cost [11, Sect. 4], and obtained in a feedback form from the associated reduced HJB equation if the reduced dimension is low enough; see Sect. 4 below.

Indeed, for this choice of  $\tau$ -value, a very good approximation of the optimal control can be already synthesized from the 6D-GK system via a PMP approach, as checked by comparison with higher-dimensional GK systems. To further reduce the dimension, we will show that a 2D projected GK system obtained from the 6D-GK system allows for the synthesis of a sub-optimal control at a nearly optimal cost (see Table 2) whereas, the 2D-GK system fails to do so.

We describe next the precise setup of the optimal control problem. After that, we provide the explicit forms of the 2D- and 6D-GK systems, and explain how the 2D projected GK system is obtained. The control synthesis results by using the PMP approach are then presented at the end of this section. As a comparison, the results by using the HJB approach are presented in Sect. 4.

**3.3.1. The numerical setup.** For the delay parameter, we take  $\tau = 1.58$ . Although this  $\tau$  value represents less than 1% increase from the critical value  $\tau_c = \pi/2$ , the amplitude of the emerged stable periodic oscillation is already not so small, as can be observed from Figure 1 (the thick black limit cycle). The parameter  $\mu$  in the cost functional  $J$  given by (3.11) is taken to be  $1/2$ . For the selection of initial history segment and the time horizon of the control, the uncontrolled Wright equation (3.1) is integrated over a sufficiently long time interval  $[0, t^*]$  so that the solution has already settled down to the attracting periodic orbit. The control horizon  $T$  is chosen such that a half of the oscillation period has developed<sup>6</sup> for the uncontrolled problem (see the dashed curve in Figure 2), and we take then a “snippet” of this periodic orbit over  $[t^* - T - \tau, t^* - T]$  as the initial history for the controlled problem (3.3). In other words, the DDE (3.3) is initiated from a snippet of the attracting periodic orbit,  $\Phi$ ,

<sup>6</sup>In our experiments it corresponds to  $T = 4$ .

from which we want to reduce its amplitude at an optimal cost by solving (3.12). We turn hereafter to the determination of the relevant GK systems in order to propose an approximate solution to this problem.

3.3.2. *The concrete Galerkin-Koornwinder approximation systems.* Recall that the  $N$ -dim GK system of Eq. (3.3) is given by Eq. (3.18)–(3.21). To get access to the numerical values of the coefficients involved in Eq. (3.18), we only need to specify the matrix  $M$ , the square of the norm of the basis function  $\|\mathcal{K}_j\|_{\mathcal{E}}^2$  as well as the value  $K_j(-1)$  of the Koornwinder polynomials at the left endpoint for  $j = 0, \dots, N-1$ .

In our case, we have (cf. [9, Proposition 3.1])

$$(3.24) \quad \begin{aligned} (\|\mathcal{K}_0\|_{\mathcal{E}}^2, \dots, \|\mathcal{K}_5\|_{\mathcal{E}}^2) &= (2, 3.333, 10, 24.2857, 49.1111, 87.4545), \\ (K_0(-1), \dots, K_5(-1)) &= (1, -3, 7, -13, 21, -31). \end{aligned}$$

With  $N = 6$ , we have from (3.19) that the matrix  $M$  appearing in Eq. (3.18) is given after computation by

$$(3.25) \quad M_6 = \begin{pmatrix} -0.5000 & 2.7658 & -5.3987 & 10.9304 & -16.8291 & 25.6266 \\ -0.3000 & 0.1405 & 2.8367 & -2.5557 & 6.6114 & -7.4089 \\ -0.1000 & 0.0468 & -2.2190 & 8.6418 & -9.8215 & 18.4165 \\ -0.0412 & 0.0193 & -0.9137 & -1.6538 & 8.4652 & -7.0109 \\ -0.0204 & 0.0095 & -0.4518 & -0.8178 & -3.2628 & 11.8690 \\ -0.0114 & 0.0054 & -0.2537 & -0.4593 & -1.8323 & -3.1193 \end{pmatrix},$$

up to four digits precision after the decimal place.

The initial data for the 6D-GK system is simply taken to be the projection of the aforementioned DDE initial history segment  $\Phi$  onto the first 6 Koornwinder basis functions; cf. (3.20). We have thus

$$(3.26) \quad \zeta_6 = (0.0590, 0.0827, 0.0014, -0.0006, 0.0, 0.0)^{\text{tr}}.$$

For the 2D-GK system, the matrix  $M_2$  consists of the  $2 \times 2$  block in the upper left corner of  $M_6$  given by (3.25). Namely,

$$(3.27) \quad M_2 = \begin{pmatrix} -0.5000 & 2.7658 \\ -0.3000 & 0.1405 \end{pmatrix}.$$

The initial data for the 2D-GK system is the same as the first two component of  $\zeta_6$  given by (3.26).

As reported in Table 2 below, the 6D-GK system produces a control strategy that is already optimal as compared with the cost from the 12D-GK system. On the other hand, the 2D-GK system leads to a control strategy that significantly inflates the control cost. This failure of the 2D-GK system can be understood by simply inspecting its linear part.

Indeed, although the first two Koornwinder modes already capture more than 98% of the  $L^2$ -energy contained in the uncontrolled DDE solution, the linear part of the corresponding 2D-GK system does not resolve well the pair of unstable eigenvalues of the DDE's linear part. More precisely, while the eigenvalues of  $M_2$  are given by

$$(3.28) \quad \beta_{2,1} = -0.1798 + 0.8527i, \quad \beta_{2,2} = -0.1798 - 0.8527i,$$

the first pair of eigenvalues of the DDE's linear part is given up to four significant digits by

$$(3.29) \quad \beta_1 = 0.0026 + 0.9958i, \quad \beta_2 = 0.0026 - 0.9958i.$$

This unstable pair of eigenvalues can nevertheless be resolved up to the given precision from  $M_6$  given in (3.25), which actually corresponds to the lowest dimension to do so.

Although, the 2D-GK system lacks to resolve important dynamical features such as the dominant eigenpair of the DDE's linear part, we provide for later usage the explicit expression of the 2D-GK system:

$$(3.30) \quad \frac{d\tilde{\zeta}_2}{dt} = M_2\tilde{\zeta}_2 - \begin{pmatrix} 0.5(\tilde{\zeta}_{2,1})^2 - \tilde{\zeta}_{2,1}\tilde{\zeta}_{2,2} - 1.5(\tilde{\zeta}_{2,2})^2 \\ 0.3(\tilde{\zeta}_{2,1})^2 - 0.6\tilde{\zeta}_{2,1}\tilde{\zeta}_{2,2} - 0.9(\tilde{\zeta}_{2,2})^2 \end{pmatrix} + \begin{pmatrix} 0.5 \\ 0.3 \end{pmatrix} u(t),$$

where  $\tilde{\zeta}_2 := (\tilde{\zeta}_{2,1}, \tilde{\zeta}_{2,2})^{\text{tr}}$ , and the initial data is  $\tilde{\zeta}_2(0) = (0.0590, 0.0827)^{\text{tr}}$ .

Let  $J_2$  be the cost functional given by (3.22) with  $N = 2$ , the control problem of the form (3.23) associated with (3.30) reads

$$(3.31) \quad \boxed{\begin{array}{l} \min J_2(\tilde{\zeta}_2, u) \text{ subject to } (\tilde{\zeta}_2(\cdot; \tilde{\zeta}_2, u), u) \in L^2(0, T; \mathbb{R}^2) \times \mathcal{U}_{ad} \\ \text{solves the ODE system (3.30) supplemented with } \tilde{\zeta}_2(0) = \tilde{\zeta}_2, \end{array}}$$

where the initial data  $\tilde{\zeta}_2$  is taken to be  $(0.0590, 0.0827)^{\text{tr}}$ .

**3.3.3. The projected GK system.** In this section we propose a natural way to cope with the resolution issue of the dominant eigenpair (3.29), while keeping the dimension of the reduced system as small as possible. It is based on the projection onto the space  $\hat{\mathcal{H}}_2$  spanned by the first two eigenvectors<sup>7</sup> of the matrix  $M_6$  given by (3.25); recalling that the sixth dimension constituting for this example the minimal dimension to resolve accurately (3.29).

Note that since the eigenvalues of  $M_6$  are all complex-valued, the resulting 2D projected GK system has complex-valued coefficients. However, the eigenvectors of  $M_6$  appear in conjugate pairs because  $M_6$  is a real-valued matrix. As a result, the two components of the projected system, denoted by  $z_2 := (z_{2,1}, z_{2,2})^{\text{tr}}$ , are complex conjugate to each other. We can thus rewrite the system under a new real-valued variable  $\tilde{\zeta}_2 := (\tilde{\zeta}_{2,1}, \tilde{\zeta}_{2,2})^{\text{tr}}$  defined by

$$(3.32) \quad \tilde{\zeta}_{2,1} = \frac{z_{2,1} + z_{2,2}}{2}, \quad \tilde{\zeta}_{2,2} = \frac{z_{2,1} - z_{2,2}}{2i}.$$

Using the MATLAB built-in function `eig` to compute the eigenbasis of  $M_6$  and the MATLAB Symbolic Math Toolbox, the resulting transformed equations for the  $\tilde{\zeta}_2$ -variable are given by

$$(3.33) \quad \frac{d\tilde{\zeta}_2}{dt} = \begin{pmatrix} 0.0026 & -0.9958 \\ 0.9958 & 0.0026 \end{pmatrix} \tilde{\zeta}_2 + \begin{pmatrix} b_{20}^1(\tilde{\zeta}_{2,1})^2 + b_{11}^1\tilde{\zeta}_{2,1}\tilde{\zeta}_{2,2} + b_{02}^1(\tilde{\zeta}_{2,2})^2 \\ b_{20}^2(\tilde{\zeta}_{2,1})^2 + b_{11}^2\tilde{\zeta}_{2,1}\tilde{\zeta}_{2,2} + b_{02}^2(\tilde{\zeta}_{2,2})^2 \end{pmatrix} + \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} u(t),$$

with the coefficients given by  $\alpha_1 = 0.0608$ ,  $\alpha_2 = 0.1133$ , and

$$\begin{aligned} b_{2,0}^1 &= -1.7887, & b_{1,1}^1 &= 2.1915, & b_{0,2}^1 &= 1.7996, \\ b_{2,0}^2 &= -3.3320, & b_{1,1}^2 &= 4.0824, & b_{0,2}^2 &= 3.3524. \end{aligned}$$

The initial data is given by

$$(3.34) \quad \tilde{\zeta}_2(0) = (0.0107, 0.0253)^{\text{tr}},$$

obtained by projecting the initial data of the 6D Galerkin system given by (3.26) onto  $\hat{\mathcal{H}}_2$ , followed by the transformation (3.32).

Associated with the cost functional  $J_6$  defined by (3.22) for the 6D GK system, we have the following cost functional  $\tilde{J}_2$  associated with the 2D projected GK system (3.33):

$$(3.35) \quad \tilde{J}_2(\tilde{\zeta}_2, u) := \int_0^T \left[ \frac{1}{2} \left( d_1 \tilde{\zeta}_{2,1}(t; \tilde{\zeta}_2, u) + d_2 \tilde{\zeta}_{2,1}(t; \tilde{\zeta}_2, u) \right)^2 + \frac{\mu}{2} u(t)^2 \right] dt, \quad \tilde{\zeta}_2 \in \mathbb{R}^2,$$

<sup>7</sup>Note that the eigenvalues are labeled in the lexicographical order. Namely, for  $1 \leq n < n' \leq 6$ , we have either  $\text{Re}\beta_n > \text{Re}\beta_{n'}$ , or  $\text{Re}\beta_n = \text{Re}\beta_{n'}$  and  $\text{Im}\beta_n \geq \text{Im}\beta_{n'}$ .

with  $d_1 = -4.0668$  and  $d_2 = 7.2710$ .

Then, corresponding to the optimal control problem (3.23), the optimal control problem associated with (3.33)–(3.34) reads:

$$(3.36) \quad \boxed{\begin{array}{l} \min \tilde{J}_2(\tilde{\zeta}_2, u) \text{ subject to } (\tilde{\zeta}_2(\cdot; \tilde{\zeta}_2, u), u) \in L^2(0, T; \mathbb{R}^2) \times \mathcal{U}_{ad} \\ \text{solves the ODE system (3.33) supplemented with } \tilde{\zeta}_2(0) = \tilde{\zeta}_2, \end{array}}$$

where the initial data  $\tilde{\zeta}_2$  is given by (3.34).

**3.3.4. Numerical results obtained from the PMP.** We present now the numerical results obtained by applying the PMP to the the optimal control problems associated with the three ODE systems specified above, namely, the problem (3.23) associated with the GK system for  $N = 2$  and  $N = 6$ , and the problem (3.36) associated with the 2D projected GK system (3.33).

In each case, the PMP approach leads to a boundary value problem (BVP) for the corresponding state and co-state variables, which is solved using the MATLAB built-in function `bvp4c`; see e.g. [12, Sect. 5] for more details. The results are shown in Fig. 2. As already mentioned earlier, the 6D-GK system allows for a control synthesis that is nearly optimal by comparison with that synthesized from the 12D-GK system; cf. Table 2. In contrast, the control synthesized from the 2D-GK system leads to a substantially higher cost value. Remarkably, the 2D projected GK system (3.33) allows for a control synthesis at a cost value very close to that obtained from higher-dimensional GK systems, i.e. at a nearly optimal cost. This success is due simply to the fact that the latter system resolves the unstable pair of eigenvalues of the DDE's linear part, while the former does not.

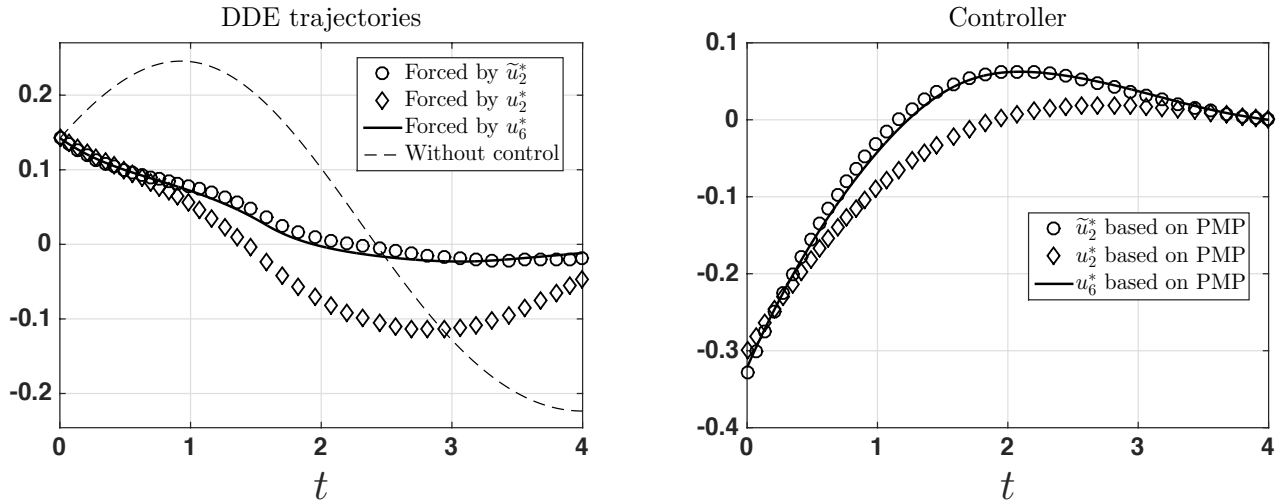


FIGURE 2. **Left panel:** DDE trajectories by solving Eq. (3.3) driven respectively by the (sub)optimal control  $\tilde{u}_2^*$ ,  $u_2^*$ , and  $u_6^*$ ; the uncontrolled trajectory is also plotted as a comparison. **Right panel:** The controller  $\tilde{u}_2^*$  synthesized by (3.36) based on the 2D projected GK system, and the controllers  $u_2^*$  and  $u_6^*$  synthesized by (3.23) based respectively on the 2D- and 6D-GK systems. We have taken  $\mu = 0.5$  in the cost functional (3.4). The maximal time of control is  $T = 4$  and the delay parameter is taken to be  $\tau = 1.58$ .

TABLE 2. Cost values and relative errors

	12D-GK	6D-GK	2D-GK	2D projected GK
Cost values	0.0163	0.0163	0.0253	0.0166
Relative error			54.93%	1.4655%

#### 4. TOWARDS THE APPROXIMATION OF HJB EQUATIONS ASSOCIATED WITH DDES

In this section we compute optimal controls by solving the HJB equations associated with the control problems (3.31) and (3.36). This allows to obtain (numerically) the controls in feedback form. However, the approach becomes infeasible in higher dimensions by the curse of dimensions (i.e. the exponential increase of complexity with the number of unknowns), since the dimension of the underlying Galerkin system is equal to the dimension of the space the HJB equation is defined on (with respect to the spatial variable).

Let us introduce for  $(\boldsymbol{\eta}, u) \in \mathbb{R}^2 \times V$  the functionals

$$(4.1a) \quad \mathcal{L}(\boldsymbol{\eta}, u) := \frac{1}{2}|\boldsymbol{\eta}|^2 + \frac{\mu}{2}u^2,$$

$$(4.1b) \quad \tilde{\mathcal{L}}(\boldsymbol{\eta}, u) := \frac{1}{2}(-4.0668\eta_1 + 7.2710\eta_2)^2 + \frac{\mu}{2}u^2.$$

Here the functional  $\mathcal{L}$  is associated with the 2D-GK system (3.30), while the functional  $\tilde{\mathcal{L}}$  is associated with the 2D projected GK system (3.33).

Given a control  $u$  in  $\mathcal{U}_{ad}[t, T]$ , we define then the following cost functionals over the interval  $[t, T]$

$$(4.2) \quad J_{t,\boldsymbol{\eta}}(u) := \int_t^T \mathcal{L}(\boldsymbol{\xi}_2(s; \boldsymbol{\eta}, u), u(s)) ds, \quad \tilde{J}_{t,\boldsymbol{\eta}}(u) := \int_t^T \tilde{\mathcal{L}}(\boldsymbol{\xi}_2(s; \boldsymbol{\eta}, u), u(s)) ds$$

where  $\boldsymbol{\xi}_2(s; \boldsymbol{\eta}, u)$  denotes the solution of the 2D-GK system (3.30), and  $\tilde{\boldsymbol{\xi}}_2(s; \boldsymbol{\eta}, u)$  the solution of the 2D projected GK system (3.33); solutions emanating in each case from the initial datum  $\boldsymbol{\eta}$  in  $\mathbb{R}^2$ . The associated value functions are given by

$$(4.3) \quad v(t, \boldsymbol{\eta}) := \inf_{u \in \mathcal{U}_{ad}[t, T]} J_{t,\boldsymbol{\eta}}(u), \quad \tilde{v}(t, \boldsymbol{\eta}) := \inf_{u \in \mathcal{U}_{ad}[t, T]} \tilde{J}_{t,\boldsymbol{\eta}}(u).$$

Formally, the value functions are associated with the instationary HJB equation of the following type

$$(4.4) \quad \partial_t v(t, \boldsymbol{\eta}) + H(\boldsymbol{\eta}, \nabla_{\boldsymbol{\eta}} v(t, \boldsymbol{\eta})) = 0 \quad \text{in } (0, T) \times \mathbb{R}^2, \quad v(T, \boldsymbol{\eta}) = 0 \quad \text{in } \mathbb{R}^2;$$

with Hamiltonian

$$(4.5) \quad H(\boldsymbol{\eta}, \boldsymbol{p}) := \inf_{u \in U} (\langle f(\boldsymbol{\eta}, u), \boldsymbol{p} \rangle_{\mathbb{R}^2} + \mathcal{L}(\boldsymbol{\eta}, u)), \quad (\boldsymbol{\eta}, \boldsymbol{p}) \in \mathbb{R}^2 \times \mathbb{R}^2;$$

where  $f$  denotes either the RHS of (3.30) if  $\mathcal{L}$  is given by (4.1a), or the RHS of (3.33) if  $\mathcal{L} = \tilde{\mathcal{L}}$  given by (4.1b). To be more specific, we denote by  $\boldsymbol{\alpha}$  the vector of coefficients in front of  $u(t)$  in either Eq. (3.30) or Eq. (3.33). We can then write  $f(\boldsymbol{\eta}, u)$  under the following generic compact form

$$(4.6) \quad f(\boldsymbol{\eta}, u) := K\boldsymbol{\eta} + N(\boldsymbol{\eta}) + \boldsymbol{\alpha}u,$$

with  $K = M_2$  and  $N(\boldsymbol{\eta})$  corresponding to the quadratic terms in Eq. (3.30), or with  $K\boldsymbol{\eta}$  and  $N(\boldsymbol{\eta})$  corresponding respectively to the linear and quadratic terms in Eq. (3.33).

In each case, by noting that the terms depending on  $u$  in  $\langle f(\boldsymbol{\eta}, u), \boldsymbol{p} \rangle_{\mathbb{R}^2} + \mathcal{L}(\boldsymbol{\eta}, u)$ , can be rewritten as

$$(4.7) \quad \begin{aligned} \frac{\mu}{2}u^2 + \langle \boldsymbol{\alpha}, \boldsymbol{p} \rangle_{\mathbb{R}^2}u &= \frac{\mu}{2} \left( u^2 + \frac{2}{\mu} \langle \boldsymbol{\alpha}, \boldsymbol{p} \rangle_{\mathbb{R}^2}u \right) \\ &= \frac{\mu}{2} \left( \left( u + \frac{1}{\mu} \langle \boldsymbol{\alpha}, \boldsymbol{p} \rangle_{\mathbb{R}^2} \right)^2 - \frac{1}{\mu^2} \langle \boldsymbol{\alpha}, \boldsymbol{p} \rangle_{\mathbb{R}^2}^2 \right). \end{aligned}$$

we obtain, for  $U = \mathbb{R}$ , that

$$(4.8) \quad H(\boldsymbol{\eta}, \boldsymbol{p}) = \frac{1}{2} \langle \boldsymbol{\eta}, Q\boldsymbol{\eta} \rangle_{\mathbb{R}^2} + \langle K\boldsymbol{\eta} + N(\boldsymbol{\eta}), \boldsymbol{p} \rangle_{\mathbb{R}^2} - \frac{\langle \boldsymbol{\alpha}, \boldsymbol{p} \rangle_{\mathbb{R}^2}^2}{2\mu},$$

with  $Q = \text{Id}_2$  in the case of Eq. (3.30), and  $\langle \boldsymbol{\eta}, Q\boldsymbol{\eta} \rangle_{\mathbb{R}^2} = (-4.0668\eta_1 + 7.2710\eta_2)^2$  in the case of Eq. (3.33).

To prove that the value functions are the unique viscosity solutions of (4.4) goes beyond the scope of this paper; here, we formulate the HJB equation and the feedback law formally; we also assume that the value function is smooth. For an analysis of HJB equations associated with infinite horizon optimal control problems for differential equations with distributed delays we refer to [22].

4.0.1. *Discretization.* For solving the HJB equation (4.4) there exist various schemes. In the presented setting we are in particular interested in efficient solvers for high dimensional HJB equations to allow also the treatment of reduced systems of more than two dimensions. Recently, different approaches have been considered to address this challenge arising from the curse of dimensions when solving HJB equations; see e.g. [32] for a polynomial approximation of high-dimensional HJB equations. In the context of suboptimal control of PDEs an approach to solve the associated HJB equation based on sparse grids was considered in [23]. Here, we use a finite difference scheme method; more precisely, an *essentially non-oscillatory* (ENO) scheme [48] for space discretization is coupled with a Runge-Kutta time discretization scheme of second order following [7, 48].

We give a brief description of the scheme for a given continuous Hamiltonian  $H: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ . For temporal mesh parameter  $\Delta t > 0$  and spatial mesh parameter  $h = (h_1, h_2) \in \mathbb{R}_{>0}^2$ , we define a spatial mesh

$$(4.9) \quad \mathcal{T} := \{(kh_1, lh_2) \in \mathbb{R}^2 : (k, l) \in \mathbb{Z}^2\}$$

and temporal mesh

$$(4.10) \quad t = t_0 < t_1 < \dots < t_N = T, \quad \Delta t = t_{j+1} - t_j,$$

for  $j = 0, \dots, N-1$  and  $N$  in  $\mathbb{N}^*$ . Given an approximation of the value function on the mesh  $\mathcal{T}$  at time step  $t_j$  denoted by  $\boldsymbol{v}: \mathbb{Z}^2 \rightarrow \mathbb{R}$ , where  $\boldsymbol{v}_{kl}$  for  $(k, l) \in \mathbb{Z}^2$  gives an approximation of the value function at grid point  $\boldsymbol{\eta}_{kl} := (kh_1, lh_2)$ , we define the difference quotients

$$(4.11) \quad (D^+ \boldsymbol{v})_{kl} := ((D_1^+ \boldsymbol{v})_{kl}, (D_2^+ \boldsymbol{v})_{kl}), \quad (D^- \boldsymbol{v})_{kl} := ((D_1^- \boldsymbol{v})_{kl}, (D_2^- \boldsymbol{v})_{kl})$$

with

$$(4.12) \quad (D_1^\pm \boldsymbol{v})_{kl} := \pm \frac{(\boldsymbol{v}_{k\pm 1, l} - \boldsymbol{v}_{kl})}{h_1}, \quad (D_2^\pm \boldsymbol{v})_{kl} := \pm \frac{(\boldsymbol{v}_{k, l\pm 1} - \boldsymbol{v}_{kl})}{h_2}.$$

Then the Lax-Friedrichs scheme for the HJB equation reads as follows at mesh node  $\boldsymbol{\eta}_{kl}$  by:

$$(4.13) \quad \begin{cases} \boldsymbol{v}_{kl}^N = 0, \\ \boldsymbol{v}_{kl}^{j-1} = \boldsymbol{v}_{kl}^j - \Delta t H_{LF}(\boldsymbol{\eta}_{kl}, (D^+ \boldsymbol{v}^j)_{kl}, (D^- \boldsymbol{v}^j)_{kl}) \quad \text{for } j = N, \dots, 1, \end{cases}$$

with  $H$  given in (4.8) and the Lax-Friedrichs Hamiltonian  $H_{\text{LF}}: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$  given by

$$(4.14) \quad H_{\text{LF}}(\hat{\boldsymbol{\eta}}, \mathbf{p}^+, \mathbf{p}^-) := H\left(\hat{\boldsymbol{\eta}}, \frac{\mathbf{p}^+ + \mathbf{p}^-}{2}\right) - \sum_{i=1}^2 \frac{v_i(\hat{\boldsymbol{\eta}})}{2}(p_i^+ - p_i^-),$$

and the stabilizing functions  $v_i(\hat{\boldsymbol{\eta}})$  satisfying

$$(4.15) \quad \max_{\mathbf{p} \in \mathbb{R}^2} |\partial_{\mathbf{p}_i} H(\hat{\boldsymbol{\eta}}, \mathbf{p})| \leq v_i(\hat{\boldsymbol{\eta}}),$$

for all  $\hat{\boldsymbol{\eta}} \in \mathbb{R}^2$  and  $i = 1$  to 2. The Lax-Friedrichs scheme is monotone and the convergence is of first order as far as the following Courant-Friedrichs-Lewy (CFL) condition is satisfied,

$$(4.16) \quad \Delta t \max_{\hat{\boldsymbol{\eta}} \in \mathcal{T}} \left( \frac{v_1(\hat{\boldsymbol{\eta}})}{h_1} + \frac{v_2(\hat{\boldsymbol{\eta}})}{h_2} \right) \leq 1.$$

An ENO scheme can be obtained by considering a variant of the Lax-Friedrichs scheme, namely

$$(4.17) \quad \mathbf{v}_{kl}^{j-1} = \mathbf{v}_{kl}^j - \Delta t H_{\text{LF}}(\boldsymbol{\eta}_{kl}, (\tilde{D}^+ \mathbf{v}^j)_{kl}, (\tilde{D}^- \mathbf{v}^j)_{kl}),$$

where  $\tilde{D}^\pm \mathbf{v}^j$  are higher order approximations of the gradient of the value function at time step  $t_j$  coupled with a Runge-Kutta time discretization scheme of second order, see [7] and [48].

The HJB equation is posed on the full space  $\mathbb{R}^2$  for the spatial variable. For the numerical computation we restrict it to a bounded domain

$$(4.18) \quad D := [c_1, d_1] \times [c_2, d_2],$$

for some  $\mathbf{c}, \mathbf{d}$  in  $\mathbb{R}^2$ , such that  $D$  contains the trajectory obtained from the PMP approach in Sect. 3.3. Choosing numbers of grid points  $p_1$  and  $p_2$  for each of the spatial dimension, we set

$$(4.19) \quad h_i := \frac{d_i - c_i}{p_i - 1}, \quad i = 1, 2.$$

and obtain the spatial meshing

$$(4.20) \quad \{(c_1 + kh_1, c_2 + lh_2) : (k, l) \in [0, \dots, p_1 - 1] \times [0, \dots, p_2 - 1]\}.$$

An appropriate choice for the boundary condition is required to get a well defined numerical scheme. At time steps  $t_j$  for  $j$  from  $N$  to 1 and grid point  $\boldsymbol{\eta}_{kl}$  we define the upwind derivatives on the boundary for all  $l \in [0, \dots, p_2 - 1]$  by

$$(4.21) \quad \begin{aligned} k = 0 : \quad (D_{\eta_1}^+ \mathbf{v})_{kl} &= \frac{\mathbf{v}_{k+1,l} - \mathbf{v}_{kl}}{h_1}, \\ k = p_1 - 1 : \quad (D_{\eta_1}^- \mathbf{v})_{kl} &= -\frac{\mathbf{v}_{k-1,l} - \mathbf{v}_{kl}}{h_1}, \end{aligned}$$

and for all  $k \in [0, \dots, p_1 - 1]$  by

$$(4.22) \quad \begin{aligned} l = 0 : \quad (D_{\eta_2}^+ \mathbf{v})_{kl} &= \frac{\mathbf{v}_{k,l+1} - \mathbf{v}_{kl}}{h_2}, \\ l = p_2 - 1 : \quad (D_{\eta_2}^- \mathbf{v})_{kl} &= -\frac{\mathbf{v}_{k,l-1} - \mathbf{v}_{kl}}{h_2}. \end{aligned}$$

This choice of boundary condition corresponds to a vanishing second order derivative boundary condition.



4.0.2. *Computation of optimal trajectories.* For the computation of optimal trajectories we formulate a feedback law based on the solution of the HJB equation (4.4); the solution is assumed to be smooth here. Following e.g. [5, p. 11] a control  $u^*$  is optimal for the initial condition  $\boldsymbol{\eta}$  if and only if

$$(4.23) \quad u^*(s) = S(s, \boldsymbol{\zeta}^*(s; \boldsymbol{\eta}, u)) \quad \text{for a.e. } s \in (0, T),$$

with

$$(4.24) \quad S(s, \hat{\boldsymbol{\eta}}) := \operatorname{argmin}_{u \in U} (\langle f(\hat{\boldsymbol{\eta}}, u) \nabla_{\boldsymbol{\eta}} v(s, \hat{\boldsymbol{\eta}}) \rangle_{\mathbb{R}^2} + \mathcal{L}(\hat{\boldsymbol{\eta}}, u)) \quad \text{for all } \hat{\boldsymbol{\eta}} \in \mathbb{R}^2,$$

where, here again,  $f$  denotes either the RHS of (3.30) if  $\mathcal{L}$  is given by (4.1a), or the RHS of (3.33) if  $\mathcal{L} = \tilde{\mathcal{L}}$  given by (4.1b). To obtain an optimal trajectory one then solves the system

$$(4.25) \quad \frac{d\boldsymbol{\zeta}(s; \boldsymbol{\eta}, u)}{ds} = f(\boldsymbol{\zeta}(s; \boldsymbol{\eta}, u), S(s, \boldsymbol{\zeta}(s; \boldsymbol{\eta}, u))), \quad s > 0, \quad \boldsymbol{\zeta}(0; \boldsymbol{\eta}, u) = \boldsymbol{\eta}.$$

From (4.7), the element  $u^*$  that leads after minimization to the Hamiltonian  $H$  given in (4.8) (and thus to (4.24)), is given by

$$(4.26) \quad u^* = -\frac{1}{\mu} \langle \boldsymbol{\alpha}, \boldsymbol{p} \rangle_{\mathbb{R}^2}.$$

which leads thus to the following *feedback law*:

$$(4.27) \quad S(s, \hat{\boldsymbol{\eta}}) = -\frac{1}{\mu} \langle \boldsymbol{\alpha}, \nabla_{\boldsymbol{\eta}} v(s, \hat{\boldsymbol{\eta}}) \rangle_{\mathbb{R}^2}.$$

For the numerical computations, the gradient  $\nabla_{\boldsymbol{\eta}} v$  is approximated component-by-component by a central difference quotient. Once the feedback operator is defined, the optimal trajectory is derived from (4.25).

4.0.3. *Numerical results.* Here, we compute optimal controls for (3.31) and (3.36) by solving the corresponding HJB equation as described above. We choose the domain  $D$  given in (4.18) in such a way that it contains the optimal trajectory by using the numerical results obtained from the PMP approach.

For (3.31) we set  $c_1 = c_2 = -0.02$  and  $d_1 = d_2 = 0.1$ , the temporal mesh parameter to  $\Delta t = 1.543 \cdot 10^{-4}$  and the spatial one to  $h_1 = h_2 = 8.5 \cdot 10^{-3}$ . The CFL constants in (4.16) are chosen to be  $\nu_1 = 5$  and  $\nu_2 = 2$ .

For (3.36), we set  $c_1 = c_2 = -0.04$  and  $d_1 = d_2 = 0.04$ ,  $\Delta t = 1.543 \cdot 10^{-4}$ , and  $h_1 = h_2 = 1.33 \cdot 10^{-3}$ . The CFL constants are here  $\nu_1 = 1.5$  and  $\nu_2 = 1.5$ .

For the computation of the optimal control from (4.27) for (3.31) or (3.36), we use the corresponding spatial and temporal mesh parameters given above.

Figure 3 shows the optimal controls for problems (3.31) and (3.36) derived from the corresponding HJB equations as described above, and also computed from the PMP approach; see Sect. 3.3. In each case, the syntheses of (sub)optimal controls obtained either from application of the PMP or by solving the associated HJB equations, provide nearly identical numerical solutions.

Figure 4 shows an approximation<sup>8</sup> of the value function  $v(0, \boldsymbol{\eta})$  solving, at  $t = 0$ , the HJB equation (4.4) associated with (3.36). We observe that the value function is smooth and convex, which justifies numerically the above assumption concerning the smoothness of the value function. Given the good control skills obtained from the controller synthesized from the reduced HJB equation (4.4) associated with (3.36), one can reasonably infer that the corresponding “reduced” value function  $v(0, \boldsymbol{\eta})$  shown in Fig. 4 provides thus a good approximation of the “full” value function associated with the DDE

<sup>8</sup>According to the ENO scheme (4.17) over the domain (4.18) with parameters specified as above and supplemented with the boundary conditions (4.21)-(4.22).

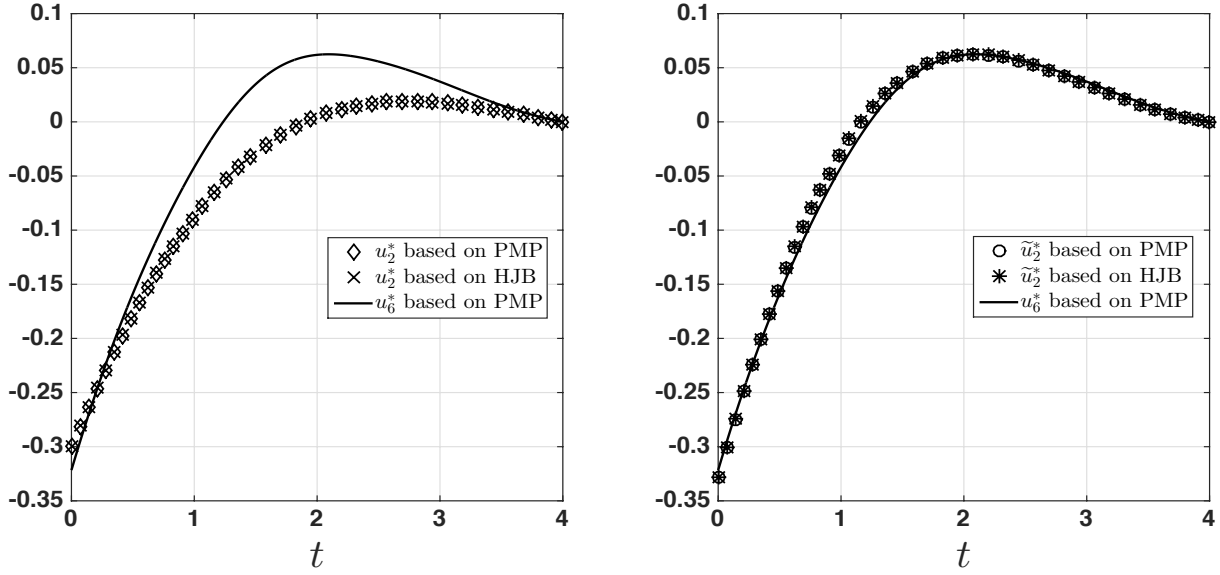


FIGURE 3. **Left panel:** Control obtained without basis transformation (HJB vs PMP), i.e. by solving (3.31). **Right panel:** Control obtained after basis transformation (HJB vs PMP), i.e. by solving (3.36).

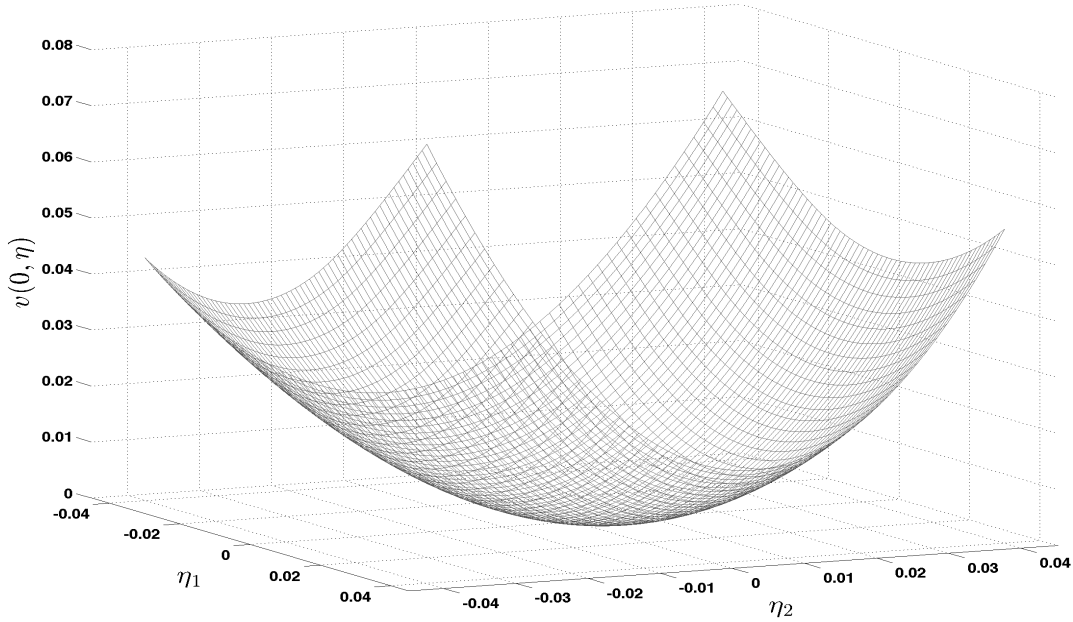


FIGURE 4. Value function solving, at  $t = 0$ , the reduced HJB equation (4.4) associated with (3.36).

optimal control problem. Due to its simple shape, a polynomial approximation of  $v(0, \eta)$  can be easily computed and gives (up to a small residual<sup>9</sup>)

$$(4.28) \quad v(0, \eta) = 28.0025\eta_1^2 - 7.8733\eta_1\eta_2 + 0.0046\eta_1 + 10.6258\eta_2^2 + 0.0525\eta_2,$$

<sup>9</sup>The root mean square error between  $v(0, \eta)$  and the polynomial approximation is  $12 \times 10^{-4}$ .

i.e. a quadratic polynomial asymmetric in the  $\eta_1$ - and  $\eta_2$ -directions. Besides the quadratic cost functional used here, such a low-order approximation of the value function is probably due to the proximity to the first criticality and not limited to time  $t = 0$  shown here. This observation deserves of course more understanding which although not pursued here, is, we believe, worth mentioning.

## 5. CONCLUDING REMARKS

As mentioned earlier, the good numerical skills shown in Sections 3.3 and 4, as obtained from low-dimensional Galerkin systems, are conditioned to the distance from the first criticality, here the Hopf bifurcation. In fact a numerical estimation of the RHS in the error estimate (2.42) of Theorem 2.1 reveals that the corresponding residual energies are actually small (not shown), in agreement with the numerical results.

As one gets further away from the first criticality, a larger number of Koornwinder polynomials is typically required to dispose of good GK approximations of, already, the uncontrolled dynamics; see [9]. The numerical burden of the synthesis of controls at a nearly optimal cost—by solving the HJB equations corresponding to the relevant GK systems—becomes then quickly prohibitive, especially for the case of locally distributed controls<sup>10</sup>. One avenue to work within reduced state spaces of further reduced dimension compared to what would be required by a GK approximation, is to search for high-mode parameterizations that help reduce the residual energy contained in the unresolved modes, i.e. to reduce a quantity like  $\mathcal{R}(u_{t,x}^*, u_{t,x_N}^{N,*})$  in (2.42) that involves the residual energies,  $\|\Pi_N^\perp y_{t,x}(\cdot; u_{t,x}^*)\|_{L^2(t,T;D(\mathcal{A}))}$  and  $\|\Pi_N^\perp y_{t,x}(\cdot; u_{t,x_N}^{N,*})\|_{L^2(t,T;D(\mathcal{A}))}$ .

The theory of parameterizing manifolds (PM) [12, 13, 16] allows for such a reduction leading typically to approximate controls coming essentially with error estimates that introduce a multiplying factor  $0 \leq Q < 1$  (related to the PM) in an RHS similar to that of (2.42); see [12, Theorem 1 & Corollary 2]. The combination of the GK framework of [9] with the PM reduction techniques of [12] constitutes thus an idea that is worth pursuing for solving efficiently optimal control problems of nonlinear DDEs.

## ACKNOWLEDGMENTS

This work has been partially supported by the Office of Naval Research (ONR) Multidisciplinary University Research Initiative (MURI) grant N00014-12-1-0911 and N00014-16-1-2073 (MDC), and by the National Science Foundation grants DMS-1616981 (MDC) and DMS-1616450 (HL). MDC and AK were supported by the project “Optimal control of partial differential equations using parameterizing manifolds, model reduction, and dynamic programming” funded by the *Foundation Hadamard/Gaspard Monge Program for Optimization and Operations Research (PGMO)*.

## REFERENCES

- [1] H. T. Banks and J. A. Burns, Hereditary control problems: Numerical methods based on averaging approximations, *SIAM J. Control Optim.* 16 (1978), 169–208.
- [2] H. T. Banks and F. Kappel, Spline approximations for functional differential equations, *Journal of Differential Equations* 34 (1979), 496–522.
- [3] H. T. Banks and K. Kunisch, The linear regulator problem for parabolic systems, *SIAM J. Control Optim.* 22 (1984), 684–698.
- [4] H. T. Banks, I. G. Rosen and K. Ito, A spline based technique for computing Riccati operators and feedback controls in regulator problems for delay equations, *SIAM J. Sci. Stat. Comput.* 5 (1984), 830–855.
- [5] M. Bardi and I. Capuzzo-Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, Birkhäuser, Boston, 2008.

<sup>10</sup>See [12, Sect. 7] for such a situation in the context of a viscous Burgers equation with instability.

- [6] A. Bensoussan, G. Da Prato, M. C. Delfour and S. K. Mitter, *Representation and Control of Infinite Dimensional Systems*, second ed, Birkhäuser Boston, Inc., Boston, MA, 2007.
- [7] O. Bokanowski, N. Forcadel and H. Zidani, A general Hamilton-Jacobi framework for non-linear state constrained control problems, *SIAM J. Control and Optim.* 48 (2010), 4292–4316.
- [8] A. Casal and M. Freedman, A Poincaré-Lindstedt approach to bifurcation problems for differential-delay equations, *Automatic Control, IEEE Transactions on* 25 (1980), 967–973.
- [9] M. D. Chekroun, M. Ghil, H. Liu and S. Wang, Low-dimensional Galerkin approximations of nonlinear delay differential equations, *Disc. Cont. Dyn. Sys. A* 36 (2016), 4133–4177.
- [10] M. D. Chekroun and N. E. Glatt-Holtz, Invariant measures for dissipative dynamical systems: Abstract results and applications, *Commun. Math. Phys.* 316 (2012), 723–761.
- [11] M. D. Chekroun, A. Kröner and H. Liu, Galerkin approximations of nonlinear optimal control problems in Hilbert spaces, *arXiv preprint* (2017), [arXiv:1704.00427](https://arxiv.org/abs/1704.00427).
- [12] M. D. Chekroun and H. Liu, Finite-horizon parameterizing manifolds, and applications to suboptimal control of nonlinear parabolic PDEs, *Acta Applicandae Mathematicae* 135 (2015), 81–144.
- [13] ———, Post-processing finite-horizon parameterizing manifolds for optimal control of nonlinear parabolic PDEs, in: *2016 IEEE 55th Conference on Decision and Control (CDC)*, IEEE, pp. 1411–1416, 2016.
- [14] M. D. Chekroun and L. J. Roques, Models of population dynamics under the influence of external perturbations: Mathematical results, *C. R. Math. Acad. Sci. Paris* 343 (2006), 307–310.
- [15] M.D. Chekroun, M. Ghil and J. D. Neelin, *Pullback attractor crisis in a delay differential ENSO model*, Advances in Non-linear Geosciences, to appear (A. Tsonis, ed.), Springer, 2017.
- [16] M.D. Chekroun, H. Liu and J.C. McWilliams, The emergence of fast oscillations in a reduced primitive equation model and its implications for closure theories, *Computers & Fluids* (2016).
- [17] S. N. Chow and J. Mallet-Paret, Integral averaging and bifurcation, *J. Differential Equations* 26 (1977), 112–159.
- [18] R.F. Curtain and H. Zwart, *An Introduction to Infinite-Dimensional Linear Systems Theory*, 21, Springer, 1995.
- [19] S. L. Das and A. Chatterjee, Multiple scales without center manifold reductions for delay differential equations near Hopf bifurcations, *Nonlinear Dynamics* 30 (2002), 323–335.
- [20] O. Diekmann, S. A. van Gils, S. M. Verduyn Lunel and H.-O. Walther, *Delay Equations: Functional, Complex, and Non-linear Analysis*, Applied Mathematical Sciences 110, Springer-Verlag, New York, 1995.
- [21] T. Faria and L. T. Magalhães, Normal forms for retarded functional differential equations with parameters and applications to Hopf bifurcation, *J. differential equations* 122 (1995), 181–200.
- [22] S. Federico, B. Goldys and F. Gozzi, HJB equations for the optimal control of differential equations with delays and state constraints, I: regularity of viscosity solutions, *SIAM J. Control Optim.* 48 (2010), 4910–4937.
- [23] J. Garcke and A. Kröner, Suboptimal Feedback Control of PDEs by Solving HJB Equations on Adaptive Sparse Grids, *Journal of Scientific Computing* (2016), 1–28.
- [24] M. Ghil, M. D. Chekroun and G. Stepan, A collection on ‘Climate Dynamics: Multiple Scales and Memory Effects’, Editorial, *R. Soc. Proc. A* 471 (2015).
- [25] J. A. Goldstein, *Semigroups of Linear Operators and Applications*, Oxford University Press, 1985.
- [26] J. K. Hale and S. M. Verduyn Lunel, *Introduction to Functional-Differential Equations*, Applied Mathematical Sciences 99, Springer-Verlag, New York, 1993.
- [27] G. E. Hutchinson, Circular causal systems in ecology, *Ann. N. Y. Acad. Sci.* 50 (1948), 221–246.
- [28] K. Ito and R. Teglás, Legendre-tau approximations for functional-differential equations, *SIAM J. Control Optim.* 24 (1986), 737–759.
- [29] G.S. Jones, The existence of periodic solutions  $\dot{f}(x) = -\alpha f(x-1) \{1 + f(x)\}$ , *Journal of Mathematical Analysis and Applications* 5 (1962), 435–450.
- [30] ———, On the nonlinear differential-difference equation  $\dot{f}(x) = -\alpha f(x-1) \{1 + f(x)\}$ , *Journal of Mathematical Analysis and Applications* 4 (1962), 440–469.
- [31] S. Kakutani and L. Markus, *On the non-linear difference-differential equation  $\dot{y}(t) = \{A - By(t-\tau)\} y(t)$* , Contributions to the Theory of Nonlinear Oscillations IV, Annals of Mathematical Studies No. 41, Princeton Univ. Press, Princeton, N. J., 1958, p. 1–18.
- [32] D. Kalise and K. Kunisch, *Polynomial approximation of high-dimensional Hamilton-Jacobi-Bellman equations and applications to feedback control of semilinear parabolic PDEs*, Report, 2017, preprint.
- [33] F. Kappel, *Semigroups and delay equations*, Semigroups, Theory and Applications, Vol. II (Trieste, 1984), Pitman Res. Notes Math. Ser. 152, Longman Sci. Tech., Harlow, 1986, pp. 136–176.
- [34] F. Kappel and D. Salamon, Spline approximation for retarded systems and the Riccati equation, *SIAM J. Control Optim.* 25 (1987), 1082–1117.

- [35] ———, Spline approximation for retarded systems and the Riccati equation, *SIAM J. Control Optim.* 25 (1987), 1082–1117.
- [36] F. Kappel and W. Schappacher, Autonomous nonlinear functional differential equations and averaging approximations, *Nonlinear Analysis: Theory, Methods & Applications* 2 (1978), 391–422.
- [37] N. D. Kazarinoff, Y.-H. Wan and P. Van den Driessche, Hopf bifurcation and stability of periodic solutions of differential-difference and integro-differential equations, *IMA J. Appl. Math.* 21 (1978), 461–477.
- [38] V. B. Kolmanovskii and V. R. Nosov, *Stability of Functional Differential Equations*, Mathematics in Science and Engineering 180, Academic Press, Inc., London, 1986.
- [39] T. H. Koornwinder, Orthogonal polynomials with weight function  $(1-x)^\alpha(1+x)^\beta + M\delta(x+1) + N\delta(x-1)$ , *Canad. Math. Bull.* 27 (1984), 205–214.
- [40] B. Krauskopf and J. Sieber, Bifurcation analysis of delay-induced resonances of the El-Niño Southern Oscillation, *Proc. R. Soc. A* 470 (2014).
- [41] Y. Kuang, *Delay Differential Equations with Applications in Population Dynamics*, Academic Press, 1993.
- [42] K. Kunisch, Approximation schemes for the linear-quadratic optimal control problem associated with delay equations, *SIAM J. Control Optim.* 20 (1982), 506–540.
- [43] M. Lakshmanan and D. V. Senthikumar, *Dynamics of Nonlinear Time-delay Systems*, Springer, Heidelberg, 2010.
- [44] P. K. Lamm, Spline approximations for nonlinear hereditary control systems, *Journal of optimization theory and applications* 44 (1984), 585–624.
- [45] N. MacDonald, *Biological Delay Systems: Linear Stability Theory*, Cambridge Studies in Mathematical Biology 8, Cambridge University Press, Cambridge, 1989.
- [46] W. Michiels and S.-I. Niculescu, *Stability and Stabilization of Time-Delay Systems: An Eigenvalue-based Approach*, Advances in Design and Control 12, SIAM, Philadelphia, PA, 2007.
- [47] A. H. Nayfeh, Order reduction of retarded nonlinear systems—the method of multiple scales versus center-manifold reduction, *Nonlinear Dyn.* 51 (2008), 483–500.
- [48] S. Osher and C.-W. Shu, Essentially nonoscillatory schemes for Hamilton-Jacobi equations, *SIAM J. Numer. Anal.* 28 (1991), 907–922.
- [49] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Applied Mathematical Sciences 44, Springer-Verlag, New York, 1983.
- [50] L. Roques and M. D. Chekroun, On population resilience to external perturbations, *SIAM J. Appl. Math.* 68 (2007), 133–153.
- [51] ———, Probing chaos and biodiversity in a simple competition model, *Ecological Complexity* 8 (2011), 98–104.
- [52] L. Roques and M.D. Chekroun, Does reaction-diffusion support the duality of fragmentation effect?, *Ecological Complexity* 7 (2010), 100–106.
- [53] S. Ruan, *Delay differential equations in single species dynamics*, Delay Differential Equations and Applications, NATO Sci. Ser. II Math. Phys. Chem. 205, Springer, Dordrecht, 2006, pp. 477–517.
- [54] H. Smith, *An Introduction to Delay Differential Equations with Applications to the Life Sciences*, Texts in Applied Mathematics 57, Springer, New York, 2011.
- [55] G. Stepan, *Retarded Dynamical Systems: Stability and Characteristic Functions*, Pitman Research Notes in Mathematics Series 210, Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1989.
- [56] M. J. Suarez and P. S. Schopf, A delayed action oscillator for ENSO, *J. atmos. Sci.* 45 (1988), 3283–3287.
- [57] C. Sun, M. Han and Y. Lin, Analysis of stability and Hopf bifurcation for a delayed logistic equation, *Chaos, Solitons & Fractals* 31 (2007), 672–682.
- [58] E. Tziperman, L. Stone, M. A. Cane and H. Jarosh, El Niño chaos: Overlapping of resonances between the seasonal cycle and the Pacific ocean-atmosphere oscillator, *Science* 264 (1994), 72–74.
- [59] G. F. Webb, Functional differential equations and nonlinear semigroups in  $L^p$ -spaces, *J. Diff. Equat.* 20 (1976), 71–89.
- [60] W. Wischert, A. Wunderlin, A. Pelster, M. Olivier and J. Gros Lambert, Delay-induced instabilities in nonlinear feedback systems, *Physical Review E* 49 (1994), 203.
- [61] E. M. Wright, A non-linear difference-differential equation, *J. Reine Angew. Math.* 194 (1955), 66–87.

DEPARTMENT OF ATMOSPHERIC & OCEANIC SCIENCES, UNIVERSITY OF CALIFORNIA, LOS ANGELES, CA 90095-1565, USA

*E-mail address:* mchekroun@atmos.ucla.edu

INSTITUT FÜR MATHEMATIK, HUMBOLDT-UNIVERSITÄT BERLIN, 10099 BERLIN, GERMANY; INRIA AND CMAP, ÉCOLE POLYTECHNIQUE, CNRS, UNIVERSITÉ PARIS SACLAY, 91128 PALAISEAU, FRANCE

*E-mail address:* axel.kroener@inria.fr

DEPARTMENT OF MATHEMATICS, VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY, BLACKSBURG, VIRGINIA 24061, USA

*E-mail address:* hhliu@vt.edu