UCLA
UCLA Previously Published Works

Title
Low-dimensional galerkin approximations of nonlinear delay differential equations

Permalink
https://escholarship.org/uc/item/5xj2h1mn

Authors
Wang, Shouhong
Liu, Honghu
Ghil, Michael
et al.

Peer reviewed

LOW-DIMENSIONAL GALERKIN APPROXIMATIONS OF NONLINEAR DELAY DIFFERENTIAL EQUATIONS

MICKAËL D. CHEKROUN, MICHAEL GHIL, HONGHU LIU, AND SHOUHONG WANG

ABSTRACT. This article revisits the approximation problem of systems of nonlinear delay differential equations (DDEs) by a set of ordinary differential equations (ODEs). We work in Hilbert spaces endowed with a natural inner product including a point mass, and introduce polynomials orthogonal with respect to such an inner product that live in the domain of the linear operator associated with the underlying DDE. These polynomials are then used to design a general Galerkin scheme for which we derive rigorous convergence results and show that it can be numerically implemented via simple analytic formulas. The scheme so obtained is applied to three nonlinear DDEs, two autonomous and one forced: (i) a simple DDE with distributed delays whose solutions recall Brownian motion; (ii) a DDE with a discrete delay that exhibits bimodal and chaotic dynamics; and (iii) a periodically forced DDE with two discrete delays arising in climate dynamics. In all three cases, the Galerkin scheme introduced in this article provides a good approximation by low-dimensional ODE systems of the DDE's strange attractor, as well as of the statistical features that characterize its nonlinear dynamics.

1. INTRODUCTUION

Systems of delay differential equations (DDEs) are widely used in many fields such as the biosciences, climate dynamics, chemistry, control theory, economics, and engineering [BGV82,DHL14,DvGVLW95, GCS15,GC87,GZT08,HVL93,KS14,Kua93,LS10,Mac89,MN07,RCC+14,Smi11,Ste89]. In particular, certain DDEs or more general differential equations with retarded arguments can be derived from hyperbolic partial differential equations that support wave propagation [CGH12,GT00,HVL93].

In contrast to ordinary differential equations (ODEs), the state space associated even with a scalar DDE is infinite-dimensional, 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 derivation of ODE approximations of DDEs involves, in general, two types of function spaces as state space: that of continuous functions $C([-\tau, 0]; \mathbb{R}^d)$, and the Hilbert space $L^2([-\tau, 0); \mathbb{R}^d)$. The former spaces have been extensively used in the case of bifurcation analysis [CF80, CMP77, DC02,FM95,KWVdD78,Nay08,WWP+94], while the latter are typically adopted in situations where quantitative accuracy is an important factor, such as in optimal control [BB78, KS78, BRI84, Kap86, KS87, Kun82, IT86, BK79].

Within the Hilbert space setting, different basis functions have been proposed to decompose the state space; these include, among others, step functions [BB78, KS78], splines [BK79, BRI84], and orthogonal polynomial functions, such as Legendre polynomials [Kap86, IT86]. Compared with step functions or splines, the use of orthogonal polynomials leads typically to ODE approximations with lower dimensions, for a given precision [BK79, IT86]. 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 [Kap86, IT86]; see Remark 2.1(iii) below.

In the present article, we propose to avoid these technical difficulties in approximating DDEs as systems of ODEs by using an alternative polynomial basis: the elements of this basis belong naturally to the domain of the underlying linear operator, but they have not been used in the DDE literature so far. The polynomials we shall use are named after Koornwinder [Koo84], who investigated polynomials that are orthogonal with respect to weight functions adjoining point masses, as discussed in Section 3 below. This polynomial basis turns out to be particularly useful not only for the rigorous analysis of polynomial-based Galerkin approximations of nonlinear systems of DDEs, as shown in Section 4, but also for their numerical treatment, cf. Section 6.

Useful new properties of the Koornwinder polynomials are identified in Lemma 3.1 for the scalar case, and a generalization of these polynomials to the vector case is given in Section 3.3; the latter includes the multi-dimensional extension of Lemma 3.1, namely Lemma 3.2. We show that these properties are essential for checking key stability and convergence conditions in Lemmas 4.1 and 4.3. Standard Galerkin approximation results for abstract nonlinear ODEs in Hilbert spaces are recalled in Theorem 4.1 and the rest of Section 4.1. They are then applied, with the help of Lemmas 4.1 and 4.3, to nonlinear systems of DDEs in Section 4.2.

Finite-time uniform convergence results are then derived for the proposed Galerkin approximations of nonlinear systems of DDEs, subject to simple and checkable conditions on the nonlinear term. These conditions are identified in Section 4.2; see Corollaries 4.2 and 4.3. The results apply to a broad class of nonlinear systems of DDEs, as discussed in Section 4.3.

The proposed framework yields a simple numerical calculation of the corresponding Galerkin approximations. Their coefficients are easily computable from the original system of DDEs by relying on simple recurrence formulas, cf. Proposition 3.1, and by solving upper triangular systems of linear equations, cf. Proposition 5.1; see Section 5 and Appendix C.

Finally, we outline here a useful interpretation of our proposed scheme regarding the finite-dimensional approximation of the linear part $\mathcal{A}$ of general systems of DDEs, when considered in the framework of Hilbert spaces, cf. (4.24). This interpretation relies on a formulation of the evolution in time of the initial state $\{x(\theta) : \theta \in [-\tau, 0]\}$ as the solution of a partial differential equation (PDE); see also Remark 4.1.

To do so, we first distinguish between the *historic part* of the evolving state, $\{x(t+\theta) : \theta \in [-\tau, 0)\}$, and the *state part*, $\{x(t)\}$. Denoting by $u(t, \theta)$ the historic part, one can then rewrite, for instance, the simple linear DDE

$$(1.1) \qquad\qquad \dot{x} = x(t - \tau), \quad \tau > 0,$$

as the linear PDE

$$(1.2) \qquad\qquad \partial_t u = \partial_\theta u, \quad -\tau \leq \theta < 0,$$

with the boundary condition

$$(1.3) \qquad\qquad \partial_t u|_{\theta=0} = u(t, -\tau), \quad t \geq 0.$$

The key point is that, roughly speaking, the local differential operator $v \mapsto \partial_\theta v$ — obtained as the history component of $\mathcal{A}$, and written out explicitly in Eq. (2.8), for instance — is approximated here by the nonlocal differential operator

$$(1.4) \qquad\qquad v \mapsto \partial_\theta v + b_N(\theta)\Big(v(-\tau) - \partial_\theta v\big|_{\theta=0}\Big), \quad \theta \in [-\tau, 0),$$

and that $b_N(\theta)$ — expressed by means of Koornwinder polynomials, cf. (4.46) — is a bounded oscillatory coefficient that vanishes in L^2 as $N \to \infty$; see Lemma 3.1. This nonlocal operator is the PDE

representation for the history component of our Koornwinder-based Galerkin approximation $\mathcal{A}_N$ given in (4.29).

Note that terms such as $v(-\tau) - \partial_\theta v\big|_{\theta=0}$, which is responsible for the nonlocal aspect of (1.4), play an important role in the theory of numerical approximation of DDEs; see, for instance, [GZT08, Appendix A] and references therein. In particular, this term provides the exact value of the jump associated with the boundary condition (1.3). The first such jump occurs at $t = 0$ in the derivative of solutions of Eq. (1.1) that emanate from a constant history[1]; this discontinuity propagates to higher-order derivatives at subsequent, integer multiples of the delay, $t = k\tau$, $k \in \mathbb{Z}^+$.

The fact that this jump term is weighted by a vanishing term suggests that, for a given degree of accuracy, good approximation can be expected when using relatively low-dimensional Koornwinder-based approximations, as long as b_N vanishes sufficiently quickly. We do not address such numerical considerations here; see, however, Table 1 in Remark 4.1 for results in the case of Eq. (1.1).

Instead, in Section 6, we provide several applications that show the proposed approximation to be not only rigorously justified, but very effective in nonlinear cases that yield quasi-periodic and chaotic, as well as nearly Brownian dynamics. In each case, low-dimensional ODE systems succeed in approximating important topological as well as statistical features of the corresponding DDE's nonlinear dynamics.

The article is organized as follows. In Section 2, we introduce the functional framework that will be adopted in Section 4.2 to recast a system of nonlinear DDEs into an abstract ODE. This framework relies on Hilbert spaces endowed with a natural inner product with a point mass. Koornwinder polynomials are then introduced in Section 3. The convergence of the Galerkin ODE systems built by projecting onto these polynomials to the original DDEs is proven in Section 4.

We provide explicit expressions of the Galerkin approximation in Section 5 for the scalar case, and in Appendix C for nonlinear systems of DDEs. Finally, numerical applications to three nonlinear DDEs are provided in Section 6. These applications involve: (i) a simple DDE with distributed delays whose solutions recall Brownian motion [Spr07]; (ii) a DDE with a discrete delay that illustrates bimodal, as well as chaotic dynamics [Spr07]; and (iii) a periodically forced DDE with two discrete delays as a highly idealized model of the El Niño-Southern Oscillation (ENSO: [GZT08, and references therein]). In all three examples, it is shown that our Galerkin scheme provides a good approximation by low-dimensional ODE systems of the DDE's strange attractor, as well as the statistical features that characterize the associated nonlinear dynamics.

2. Background and motivation

We introduce in this section the functional framework that will be adopted in Section 4.2 for the derivation of Galerkin approximations of a given nonlinear system of DDEs. Several function spaces can be used as a state space for the reformulation of a system of DDEs into an abstract ODE where among the most standard ones, those built-up out of the space of continuous functions on the interval $[-\tau, 0]$ play an important role in the DDE theory; e.g. [DvGVLW95, HVL93].

In this article, we adopt instead the use of Hilbert spaces which are more classically used in control or approximation theory of DDEs; see e.g., [BB78, BHS83, CZ95, Kap86, KZ86, KS87, NY89]. For a didactic expository of the associated theory of semigroups for (linear) systems of DDEs in this functional setting we refer to [CZ95, Sect. 2.4]; see also [BHS83].

More precisely, the following Hilbert product space

$$(2.1) \qquad\qquad \mathcal{H} := L^2([-\tau, 0); \mathbb{R}^d) \times \mathbb{R}^d,$$

---

[1] For instance, if $\tau = 1$ and the history is given by $\{\psi(\theta) \equiv c, \ -\tau \le \theta < 0\}$, then the solution $x(t)$ to Eq. (1.1) is equal to $c(t+1)$ on $[0,1]$. This discontinuity leads to a jump $\psi(-1) - \partial_\theta \psi\big|_{\theta=0} = c$ in its time derivative at $t = 0$.

will serve as our state space, and will be endowed with the inner product defined for any $(f_1, \gamma_1), (f_2, \gamma_2) \in \mathcal{H}$, as:

$$(2.2) \qquad \langle (f_1, \gamma_1), (f_2, \gamma_2) \rangle_{\mathcal{H}} := \frac{1}{\tau} \int_{-\tau}^{0} \langle f_1(\theta), f_2(\theta) \rangle \mathrm{d}\theta + \langle \gamma_1, \gamma_2 \rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product of $\mathbb{R}^d$.

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

$$(2.3) \qquad \mathcal{V} := H^1([-\tau, 0); \mathbb{R}^d) \times \mathbb{R}^d,$$

where $H^1([-\tau, 0); \mathbb{R}^d)$ denotes the standard Sobolev subspace of $L^2([-\tau, 0); \mathbb{R}^d)$; see, e.g. [Bré10, Chap. 8]. This space consists of functions that are square integrable and whose first-order weak derivatives exist in a distributional sense and are also square integrable.

Instead of presenting the general nonlinear systems of DDEs considered in this article (see Section 4.2), we introduce below a class of scalar DDEs that will serve us to identify within a simple context, the issues inherent to the Galerkin approximation of DDEs; see Remark 2.1 hereafter.

Example 2.1. In this example, we recall how a scalar DDE can be recast into an abstract ODE. For simplicity, we will focus on the following autonomous scalar DDE $(d = 1)$:

$$(2.4) \qquad \frac{\mathrm{d}x(t)}{\mathrm{d}t} = ax(t) + bx(t - \tau) + c \int_{t-\tau}^{t} x(s)\mathrm{d}s + F\Big(x(t), \int_{t-\tau}^{t} x(s)\mathrm{d}s\Big),$$

where a, b and c are real numbers, $\tau > 0$ is the delay parameter, and F is a given scalar nonlinear function. The case of nonlinear systems of DDEs will be dealt with in Section 4.2.

The reformulation of Eq. (2.4) into an abstract ODE is classical and proceeds as follows. Let us denote by x_t the time evolution of the history segments of a solution to Eq. (2.4), namely

$$(2.5) \qquad x_t(\theta) := x(t + \theta), \qquad t \geq 0, \qquad \theta \in [-\tau, 0].$$

Now, by introducing the new variable

$$(2.6) \qquad u(t) := (x_t, x(t)) = (x_t, x_t(0)),$$

Eq. (2.4) can be rewritten as the following abstract ODE:

$$(2.7) \qquad \frac{\mathrm{d}u}{\mathrm{d}t} = \mathcal{A}u + \mathcal{F}(u),$$

where the linear operator $\mathcal{A} \colon D(\mathcal{A}) \subset \mathcal{V} \to \mathcal{H}$ is defined by

$$(2.8) \qquad [\mathcal{A}\Psi](\theta) := \begin{cases} \dfrac{\mathrm{d}^+\Psi^D}{\mathrm{d}\theta}, & \theta \in [-\tau, 0), \\[2ex] a\Psi^S + b\Psi^D(-\tau) + c \displaystyle\int_{-\tau}^{0} \Psi^D(s)\mathrm{d}s, & \theta = 0, \end{cases}$$

with the domain $\mathcal{A}$ given by (cf. [Kap86, Prop. 2.6])

$$(2.9) \qquad D(\mathcal{A}) = \Big\{ (\Psi^D, \Psi^S) \in L^2([-\tau, 0); \mathbb{R}) \times \mathbb{R} : \Psi^D \in H^1([-\tau, 0); \mathbb{R}), \lim_{\theta \to 0^-} \Psi^D(\theta) = \Psi^S \Big\};$$

and with the nonlinearity $\mathcal{F} \colon \mathcal{H} \to \mathcal{H}$ defined by

$$(2.10) \qquad [\mathcal{F}(\Psi)](\theta) := \begin{cases} 0, & \theta \in [-\tau, 0), \\[1ex] F\Big(\Psi^S, \int_{-\tau}^{0} \Psi^D(s)\mathrm{d}s\Big), & \theta = 0, \end{cases} \qquad \forall\, \Psi = (\Psi^D, \Psi^S) \in \mathcal{H}.$$

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 $\dot{u} = \mathcal{A}u$ is well-posed in the Hadamard's sense; see

e.g [CZ95, Thm. 2.4.6]. The well-posedness problem for the nonlinear equation depends obviously on the nonlinear term $\mathcal{F}$ and we refer to Section 4.2 for a solution to this problem within our functional framework; see also [Web76].

Remark 2.1.

(i) It is important to note that when instead of $L^2([-\tau, 0]; \mathbb{R}^d)$, the space of continuous functions $X = C([-\tau, 0]; \mathbb{R}^d)$ endowed with the supremum norm is retained [HVL93], the continuity requirement at 0 in (2.9) is naturally satisfied. On the other hand, X is not a Hilbert space and the analysis of the *adjoint eigenvalue problem* [HVL93, Sect. 7.5] is required for the derivation of low-dimensional ODE systems which no longer contain memory terms [WWP$^+$94]. By working within the framework of Hilbert spaces we avoid technicalities inherent to the analysis of this adjoint problem.

(ii) When we consider the Hilbert space $\mathcal{H}$, a natural choice of set of functions to decompose the solutions of (2.7) is constituted by the eigenfunctions of the operator $\mathcal{A}$ with domain $D(\mathcal{A})$. When $\mathcal{A}$ does not contain distributed delay terms, these eigenfunctions are well-known and can be found in e.g. [CZ95, Thm. 2.4.6]. In case where the eigenvalues of $\mathcal{A}$ are all simple, this set of eigenfunctions actually correspond to the set $\mathcal{E}$ of eigenfunctions in $C([-\tau, 0]; \mathbb{R}^d)$ [HVL93, Thm. 4.2, p. 207]. The latter set may fail however in approximating continuous functions [HVL93, Cor. 8.1, p. 222] and can be even finite-dimensional [HVL93, p. 220] which limits seriously its usage in practice[2] if for instance snippets of solutions to Eq. (2.7) are spanned by elements outside of $\mathcal{E}$.[3]

(iii) Due to the aforementioned limitations of the eigenfunctions, other basis functions are often used for the derivation of ODE systems to approximate the dynamics of the underlying DDE. Choices proposed in the literature include step functions [BB78, KS78], splines [BK79, BRI84], and orthogonal polynomial functions such as Legendre polynomials [Kap86, IT86].[4]

In most of the cases, a version of the Trotter-Kato theorem (see e.g. [Paz83, Thm. 4.5, p. 88]) is typically used to obtain finite-time uniform approximation results of the semigroup generated by $\mathcal{A}$. In the cases of step functions and splines, the conditions required in the Trotter-Kato theorem (see e.g. Conditions **(A1)** and **(A2)** in Theorem 4.1 below) have been analyzed in [BB78] and [BK79].

For the case of Legendre polynomials, technical complications have been encountered to check these conditions either in the setting of Galerkin approximation [Kap86] or in the setting of tau-method [IT86] largely due to the fact that the basis functions do not live in the domain of $\mathcal{A}$. As noted in [Kap86, p. 168] or in [IT86, Sect. 5], either $X_N \not\subset D(\mathcal{A})$ or Π_N is not orthogonal for the polynomial functions considered in [Kap86] and [IT86], respectively.

On the other hand, at a given precision, the use of polynomial basis leads typically to ODE approximations with lower dimensions when compared with those built out of step functions or splines [BK79, IT86]. □

The problems discussed in (iii) of the above remark already encountered in the linear case, have limited the applications of polynomial basis for the approximation of nonlinear systems of DDEs. The above discussion leads naturally to the question whether there exists an orthogonal polynomial basis

---

[2]for the purpose of low-dimensional approximations.

[3]See however [CZ95, Thm. 2.5.10] for a sufficient condition for the set of (generalized) eigenfunctions to be dense in $\mathcal{H}$ still for the case when $\mathcal{A}$ does not contain distributed delay terms.

[4]It is also worth mentioning the more recent works [Vya12, WC05], in which interesting approximation schemes based on linear and sine functions have been proposed for the case of state dependent delays, and for which successful numerical performances have been reported although rigorous convergence results seem still to be lacking, within this approach.

for which standard approximation results for abstract nonlinear systems such as recalled in Theorem 4.1 below, could be applied to the case of nonlinear systems of DDEs.

The next section introduces orthogonal polynomials that will allow us to answer this question by the affirmative, leading to direct and explicit formulas for the rigorous Galerkin approximations of a broad class of nonlinear systems of DDEs; see Sections 4.2 and 5. As explained next, the key is to seek for polynomials that live in the domain of $\mathcal{A}$, which is achieved here by seeking for polynomials to be orthogonal for the inner product (2.2) with a point mass.

3. ORTHOGONAL POLYNOMIALS FOR INNER PRODUCTS WITH A POINT MASS

The inner product given in (2.2) is naturally associated with the measure

$$(3.1) \qquad \nu(\mathrm{d}\theta) = \mathrm{d}\theta + \delta_0,$$

where δ_0 denotes the Dirac measure concentrated at $\theta = 0$.

Orthogonal polynomials with respect to the Lebesgue measure $\mathrm{d}\theta$ or measures having a smooth density with respect to it, has a long history [Sze75]. The study of orthogonal polynomials with respect to a measure including a point mass such as given by (3.1) has been studied only lately [Ism05, Chap. 2.9]. It was in particular noticed that orthogonal polynomials with respect to such a measure can be expressed in terms of polynomials orthogonal with respect to the smooth part of the measure; see [Uva69] for an early contribution on the topic.

Koornwinder in [Koo84] dealt with the case of orthogonal polynomials on $[-1, 1]$ associated with measures given by

$$(3.2) \qquad \nu(\mathrm{d}x) = \frac{\Gamma(\alpha + \beta + 2)}{2^{\alpha+\beta+1}\Gamma(\alpha + 1)\Gamma(\beta + 1)}(1 - x)^{\alpha}(1 + x)^{\beta}\mathrm{d}x + M\delta_{-1} + N\delta_1, \ \alpha, \beta > -1,$$

i.e. associated with measures having a Jacobi weight on $[-1, 1]$ with two point-masses added to the extremities of the interval.

Although many properties — such as three-term recurrence relationships or differential equations satisfied by such polynomials — remain valid in the case of a measure with a point mass, subtle but important qualitative and quantitative differences arise. For instance, [AMR97, Thm. 3 c)] shows that the zeroes closest to 1 of polynomials orthogonal with respect to the measure ν given in (3.2) converge to 1 faster than those associated with the standard Jacobi polynomials.

It is our goal to show that orthogonal polynomials with respect to the measure ν in (3.2) allows us to work within a simple and more direct framework than those found in the literature, for building Galerkin approximations of DDEs. Indeed, the approximation of DDEs by systems of ODEs built from orthogonal polynomials were not, so far, relying on classical Galerkin schemes as noted in [Kap86, p. 168] or in [IT86, Sect. 5].

The results given below correspond to the case $\alpha = \beta = M = 0$, and $N = 1$ considered in [Koo84].

3.1. **Koornwinder polynomials.** We recall next from [Koo84, Eq. (2.1)] the following sequence of Koornwinder polynomials $\{K_n\}$ that can be built from the Legendre polynomials L_n according to

$$(3.3) \qquad K_n(s) := -(1 + s)\frac{\mathrm{d}}{\mathrm{d}s}L_n(s) + (n^2 + n + 1)L_n(s), \ s \in [-1, 1], \ n \in \mathbb{N}.$$

As recalled above, this polynomial sequence is known to be orthogonal when a Dirac point-mass at the right endpoint, δ_1, is adjoined to the Lebesgue measure [Koo84], in other words

$$(3.4) \qquad \int_{-1}^{1} K_n(s)K_m(s)\mathrm{d}\mu(s) = \frac{1}{2}\int_{-1}^{1} K_n(s)K_m(s)\mathrm{d}x + K_n(1)K_m(1)$$
$$= 0, \ \text{if } m \neq n.$$

This orthogonality property and the main properties satisfied by $\{K_n\}$ on which we will rely on, are summarized from [Koo84] in the proposition below.

Proposition 3.1. The polynomial K_n defined in (3.3) is of degree n and admits the following expansion in terms of the Legendre polynomials:

$$(3.5) \qquad K_n(s) = -\sum_{j=0}^{n-1}(2j+1)L_j(s) + (n^2+1)L_n(s), \qquad n \in \mathbb{N};$$

and the following normalization property holds:

$$(3.6) \qquad K_n(1) = 1, \qquad n \in \mathbb{N}.$$

Moreover, the sequence given by

$$(3.7) \qquad \{\mathcal{K}_n := (K_n, K_n(1)) : n \in \mathbb{N}\}$$

forms an orthogonal basis of the product space

$$(3.8) \qquad \mathcal{E} := L^2([-1,1);\mathbb{R}) \times \mathbb{R},$$

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

$$(3.9) \qquad \langle (f,a),(g,b)\rangle_{\mathcal{E}} = \frac{1}{2}\int_{-1}^{1} f(s)g(s)\mathrm{d}s + ab, \quad (f,a),(g,b) \in \mathcal{E}.$$

Moreover $\left\{\frac{\mathcal{K}_n}{\|\mathcal{K}_n\|_{\mathcal{E}}}\right\}$ forms a Hilbert basis of $\mathcal{E}$ where the norm $\|\mathcal{K}_n\|_{\mathcal{E}}$ of $\mathcal{K}_n$ induced by $\langle \cdot,\cdot\rangle_{\mathcal{E}}$ possesses the following analytic expression:

$$(3.10) \qquad \|\mathcal{K}_n\|_{\mathcal{E}} = \sqrt{\frac{(n^2+1)((n+1)^2+1)}{2n+1}}, \qquad n \in \mathbb{N}.$$

Proof. Based on (3.3), the proof consists essentially of algebraic manipulations relying on the following standard properties of the Legendre polynomials [STW11, Sect. 3.3]:

- Orthogonality:

$$(3.11) \qquad \int_{-1}^{1} L_m(s)L_n(s)\mathrm{d}x = \frac{2}{2n+1}\delta_{mn}, \qquad m,n \in \mathbb{N},$$

  where δ_{mn} denotes the Kronecker delta.

- Normalization:

$$(3.12) \qquad L_n(1) = 1, \qquad n \in \mathbb{N}.$$

- Three-term recurrence relation:

$$(3.13) \qquad (n+1)L_{n+1}(s) = (2n+1)sL_n(s) - nL_{n-1}(s), \ s \in [-1,1], \ n \geq 1,$$

  where the first two Legendre polynomials are given by

$$(3.14) \qquad L_0 \equiv 1 \quad \text{and} \quad L_1(s) = s.$$

- First order derivative recurrence relation:

$$(3.15) \qquad \frac{\mathrm{d}L_n}{\mathrm{d}s}(s) = \sum_{k \in I_n}(2k+1)L_k(s),$$

  where

$$(3.16) \qquad I_n := \{k \in \mathbb{N} : 0 \leq k \leq n-1, k+n \text{ is odd}\}.$$

Standard density arguments, outlined in Appendix A, allow us then to conclude the proof. $\qquad \square$

3.2. **Rescaled Koornwinder basis.** From the original Koornwinder basis given on the interval $[-1, 1]$, orthogonal polynomials on the interval $[-\tau, 0]$ for the inner product (2.2) can now be easily obtained by using a simple linear transformation $\mathcal{T}$ defined by:

$$(3.17) \qquad \mathcal{T} \colon [-\tau, 0] \to [-1, 1], \qquad \theta \mapsto 1 + \frac{2\theta}{\tau}.$$

Indeed, for K_n given in (3.5), let us define the polynomial K_n^τ by

$$(3.18) \qquad \begin{aligned} K_n^\tau &\colon [-\tau, 0] \to \mathbb{R}, \\ \theta &\mapsto K_n\Big(1 + \frac{2\theta}{\tau}\Big), \qquad n \in \mathbb{N}. \end{aligned}$$

Since the sequence $\{\mathcal{K}_n = (K_n, K_n(1)) : n \in \mathbb{N}\}$ forms an orthogonal basis for $\mathcal{E}$ (cf. Proposition 3.1), it follows then that the polynomial sequence

$$(3.19) \qquad \{\mathcal{K}_n^\tau := (K_n^\tau, K_n^\tau(0)) : n \in \mathbb{N}\}$$

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 (2.2) for $d = 1$.

Since $K_n(1) = 1$ from (3.6), we have

$$(3.20) \qquad K_n^\tau(0) = 1.$$

Moreover, by applying the transformation $\mathcal{T}$, we get trivially that

$$(3.21) \qquad \|\mathcal{K}_n^\tau\|_{\mathcal{H}} = \|\mathcal{K}_n\|_{\mathcal{E}}.$$

We have then the following fundamental lemma.

Lemma 3.1. The rescaled Koornwinder polynomials $\{K_j^\tau\}_{j \geq 0}$ satisfy the following properties:

$$(3.22) \qquad \boxed{\sum_{j=0}^{\infty} \frac{K_j^\tau}{\|\mathcal{K}_j^\tau\|_{\mathcal{H}}^2} = 0, \quad \text{in the } L^2 \text{ sense,}}$$

and

$$(3.23) \qquad \boxed{\sum_{j=0}^{\infty} \frac{1}{\|\mathcal{K}_j^\tau\|_{\mathcal{H}}^2} = 1.}$$

Moreover, each function in $L^2([-\tau, 0); \mathbb{R})$ enjoys the following decomposition in terms of the Koornwinder polynomials K_j^τ:

$$(3.24) \qquad \boxed{f = \sum_{j=0}^{\infty} \frac{\langle f, K_j^\tau \rangle_{L^2}}{\tau \|\mathcal{K}_j^\tau\|_{\mathcal{H}}^2} K_j^\tau, \qquad \forall \, f \in L^2([-\tau, 0); \mathbb{R}).}$$

Proof. For any $\Psi = (\Psi^D, \Psi^S) \in \mathcal{H}$, we have[5]

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

Now, let Ψ^D to be the zero-function on $[-\tau, 0]$ and Ψ^S to be 1. For such a Ψ, by equalizing respectively the D-components and S-components of the RHS and LHS of (3.25), one then obtains from (3.20) that

$$
(3.26) \qquad \sum_{j=0}^{\infty} \frac{K_j^\tau}{\|\mathcal{K}_j^\tau\|_{\mathcal{H}}^2} = 0, \quad \text{in the } L^2 \text{ sense,}
$$

and

$$
(3.27) \qquad \sum_{j=0}^{\infty} \frac{1}{\|\mathcal{K}_j^\tau\|_{\mathcal{H}}^2} = 1.
$$

The decomposition of L^2 functions given in (3.24) follows directly from (3.25) by considering $\Psi := (f, 0) \in \mathcal{H}$. Again, the equality holds in the L^2 sense. $\qquad \square$

As an illustration of the identities (3.22) and (3.23), Figure 1 displays numerical computations of the partial sum $\sum_{j=0}^{N-1} \frac{K_j^\tau}{\|\mathcal{K}_j^\tau\|_{\mathcal{H}}^2}$ for $N = 20$ and $N = 60$, in the case $\tau = 0.5$.
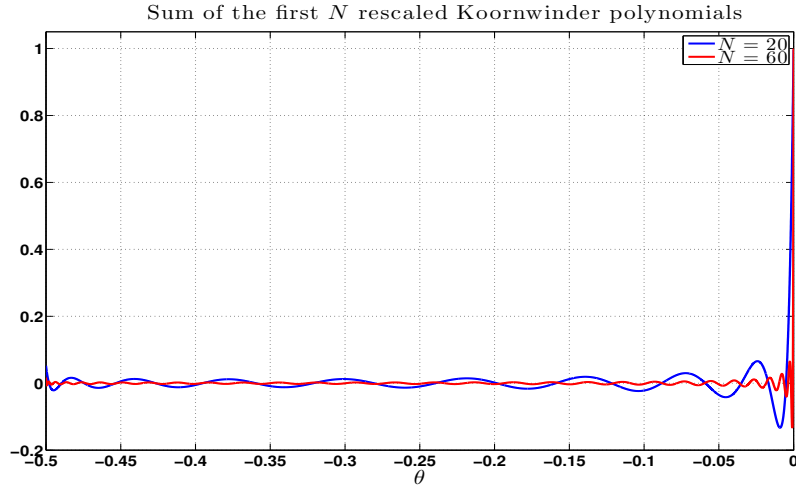

FIGURE 1. Sum of the first N rescaled Koornwinder polynomials: blue curve corresponds to $N = 20$, and red curve to $N = 60$.

---

[5]Note that the equality in (3.25) holds in the sense that $\left\| \Psi - \sum_{j=0}^{\infty} \frac{\langle \Psi, \mathcal{K}_j^\tau \rangle_{\mathcal{H}}}{\|\mathcal{K}_j^\tau\|_{\mathcal{H}}^2} \mathcal{K}_j^\tau \right\|_{\mathcal{H}} = 0$, which is equivalent to $\left\| \Psi^D - \sum_{j=0}^{\infty} \frac{\langle \Psi, \mathcal{K}_j^\tau \rangle_{\mathcal{H}}}{\|\mathcal{K}_j^\tau\|_{\mathcal{H}}^2} K_j^\tau \right\|_{L^2} = 0$ and $\left| \Psi^S - \sum_{j=0}^{\infty} \frac{\langle \Psi, \mathcal{K}_j^\tau \rangle_{\mathcal{H}}}{\|\mathcal{K}_j^\tau\|_{\mathcal{H}}^2} \right| = 0$.

Remark 3.1.

(i) Note that the continuity condition, $\lim_{\theta \to 0^-} \Psi^D(\theta) = \Psi^S$, required in (2.9) in order that $\Psi \in D(\mathcal{A})$, is here naturally satisfied by the Koornwinder basis function $\mathcal{K}_n^\tau = (K_n^\tau, K_n^\tau(0))$. It constitutes thus, for the inner product (2.2), an orthogonal polynomial basis whose elements live in the domain of the linear operator $\mathcal{A}$ given in (2.8).

As explained at the beginning of Section 3, the key element for such a construction relies on the inclusion of a point mass adjoined to the continuous part of the measure. When this point mass is absent, the corresponding orthogonal polynomials are (rescaled) Legendre polynomials. The associated basis in this latter case is given by (cf. [Kap86])

$$(3.28) \qquad \mathfrak{B} := \{\psi_1 := (0_\mathcal{H}, 1)\} \cup \{\psi_n := (L_{n-2}^\tau, 0) \mid n = 2, 3, \cdots \},$$

where $0_\mathcal{H}$ denotes the zero function on $\mathcal{H}$, and L_n^τ is the Legendre polynomial of degree n defined on the interval $[-\tau, 0]$. Clearly, none of the elements in $\mathfrak{B}$ belongs to $D(\mathcal{A})$ since $\lim_{\theta \to 0^-} L_n^\tau(\theta) \neq 0$.

(ii) The fact that the Koornwinder basis functions live in $D(\mathcal{A})$ allows us to construct *standard* Galerkin approximations; whereas, extra correction terms are required in the Galerkin approximation built from the Legendre-based basis given in (3.28) (see e.g. [Kap86, p. 169]). Moreover, technical complications such as pointed out in Remark 2.1 iii) do not take place for the case of Koornwinder basis. Indeed, it will be shown in Section 4.2 that the properties of the Koornwinder polynomials such as summarized in Corollary 3.1 as well as the vectorized version given by Corollary 3.2 below, turn out to be sufficient to obtain finite-time uniform approximation results of the semigroup generated by the linear operator $\mathcal{A}$; see Lemmas 4.1 and 4.3.

(iii) It is also worth mentioning that Corollary 3.1 and Corollary 3.2 as well as the rigorous convergence results presented in Section 4.2 are not limited to the case of Koornwinder basis constructed here. Given any polynomial basis on $[-\tau, 0]$ that are orthogonal with respect to a measure of the form $\nu(\mathrm{d}\theta) = \mathrm{d}\rho(\theta) + \delta_0$ with ρ being a positive non-decreasing function on $[-\tau, 0]$, the aforementioned results would still hold. We refer to [Uva69] for the construction of such polynomials when orthogonal polynomials with respect to $\widetilde{\nu}(\mathrm{d}\theta) = \mathrm{d}\rho(\theta)$ are known. $\square$

3.3. **Vectorization of Koornwinder polynomials.**
We introduce here a generalization of the Koornwinder polynomials that will turn out to be useful for the approximation of nonlinear DDE systems.

The purpose is here to build from the Koornwinder polynomials introduced above, linear subspaces $\mathcal{H}_N$ that approximate $\mathcal{H} := L^2([-\tau, 0); \mathbb{R}^d) \times \mathbb{R}^d$, for $d > 1$.

Each function Ψ in $\mathcal{H}$ has here d-components that can be, each, approximated by a series of Koornwinder polynomials as in (3.25). If we restrict such an approximation to the first N Koornwinder polynomials, $\mathcal{H}_N$ becomes then an $N \times d$-dimensional subspace; see (4.26) below.

Our goal is also here to introduce a vectorization of Koornwinder polynomials which allows for a natural extension of Lemma 3.1 in the case $d > 1$. This extension of Lemma 3.1 will be particularly useful to provide finite-dimensional approximation of the linear part of systems of DDEs; see Lemma 4.1.

To do so, given $j \in \{1, \cdots, Nd\}$, we can associate a Koornwinder polynomial of degree $j_q \in \{0, \cdots, N-1\}$, as follows

$$(3.29) \qquad j = dj_q + j_r,$$

where $j_r \in \{1, \cdots, d\}$ is given by

$$(3.30) \qquad j_r := \begin{cases} \mathrm{mod}(j, d), & \text{if } \mathrm{mod}(j, d) \neq 0, \\ d, & \text{otherwise.} \end{cases}$$

Let us introduce now the following d-dimensional mapping from $[-\tau, 0]$ to $\mathbb{R}^d$

$$(3.31) \qquad \mathbf{K}_j^\tau(\theta) := (\underbrace{0, \cdots, 0}_{j_r - 1 \text{ entries}}, K_{j_q}^\tau(\theta), \underbrace{0, \cdots, 0}_{d - j_r \text{ entries}})^{\mathrm{tr}}, \qquad \theta \in [-\tau, 0].$$

The vector $\mathbf{K}_j^\tau(\theta)$ is then nothing else than a d-dimensional canonical vector whose j_r^{th}-entry is given by the value at θ of the (rescaled) Koornwinder polynomial of degree j_q; the integers j_q and j_r being related to j according to (3.29)-(3.30). Based on these vectorized (rescaled) Koornwinder polynomials $\mathbf{K}_j^\tau$, we also introduce

$$(3.32) \qquad \mathbb{K}_j^\tau := \left(\mathbf{K}_j^\tau, \mathbf{K}_j^\tau(0) \right), \qquad j \geq 1.$$

In the remaining part of this section, we summarize some key properties of $\mathbf{K}_j^\tau$ and $\mathbb{K}_j^\tau$ for later usage. Hereafter, we use $\mathcal{H}_1$ to denote the space $\mathcal{H}$ defined in (2.1) for the case $d = 1$, i.e.,

$$(3.33) \qquad \mathcal{H}_1 = L^2([-\tau, 0); \mathbb{R}) \times \mathbb{R},$$

which is again endowed with the inner product given in (2.2) (still with $d = 1$).

Since the sequence $\{\mathcal{K}_j^\tau = (K_j^\tau, K_j^\tau(0)) : j \in \mathbb{N}\}$ forms an orthogonal basis for $\mathcal{H}_1$ (cf. Section 3.2), one can readily check that $\{\mathbb{K}_j^\tau : j \in \mathbb{N}^*\}$ forms an orthogonal basis for the space $\mathcal{H}$.

Note also that

$$(3.34) \qquad \|\mathbb{K}_j^\tau\|_{\mathcal{H}} = \|\mathcal{K}_{j_q}^\tau\|_{\mathcal{H}_1}, \qquad j \in \mathbb{N}^*.$$

Given this vectorization of Koornwinder polynomials, we can now formulate the following extension of Lemma 3.1 that summarizes the key properties of the $\mathbf{K}_j$'s which will be used for the rigorous approximation of nonlinear systems of DDEs such as described in Section 4.2.

Lemma 3.2. The vectorized rescaled Koornwinder polynomials $\{\mathbf{K}_j^\tau\}_{j \geq 1}$ satisfy the following properties:

$$(3.35) \qquad \boxed{\sum_{j=1}^\infty \frac{\langle \alpha, \mathbf{K}_j^\tau(0) \rangle}{\|\mathcal{K}_{j_q}^\tau\|_{\mathcal{H}_1}^2} \mathbf{K}_j^\tau = 0 \quad \text{in the } L^2([-\tau, 0); \mathbb{R}^d) \text{ sense}, \quad \forall\, \alpha \in \mathbb{R}^d,}$$

and

$$(3.36) \qquad \boxed{\sum_{j=1}^\infty \frac{\langle \alpha, \mathbf{K}_j^\tau(0) \rangle}{\|\mathcal{K}_{j_q}^\tau\|_{\mathcal{H}_1}^2} \mathbf{K}_j^\tau(0) = \alpha, \quad \forall\, \alpha \in \mathbb{R}^d.}$$

Moreover, each function in $L^2([-\tau, 0); \mathbb{R}^d)$ enjoys the following decomposition in terms of the vectorized Koornwinder polynomials $\mathbf{K}_j^\tau$:

$$(3.37) \qquad \boxed{f = \sum_{j=1}^\infty \frac{\langle f, \mathbf{K}_j^\tau \rangle_{L^2}}{\tau \|\mathcal{K}_{j_q}^\tau\|_{\mathcal{H}_1}^2} \mathbf{K}_j^\tau, \qquad \forall\, f \in L^2([-\tau, 0); \mathbb{R}^d);}$$

and the following identity holds:

$$(3.38) \qquad \boxed{\sum_{j=1}^\infty \frac{\langle f, \mathbf{K}_j^\tau \rangle_{L^2}}{\tau \|\mathcal{K}_{j_q}^\tau\|_{\mathcal{H}_1}^2} \mathbf{K}_j^\tau(0) = 0, \qquad \forall\, f \in L^2([-\tau, 0); \mathbb{R}^d).}$$

Proof. The above identities can be obtained by using the same type of reasoning as given in the proof of Lemma 3.1 for the scalar case.

Indeed, by noting that $\{\mathbb{K}_j^\tau : j \in \mathbb{N}^*\}$ forms an orthogonal basis of $\mathcal{H}$, any $\Psi \in \mathcal{H}$ admits the following decomposition:

$$
(3.39) \qquad
\begin{aligned}
\Psi &= \sum_{j=1}^\infty \frac{\langle \Psi, \mathbb{K}_j^\tau \rangle_{\mathcal{H}}}{\|\mathbb{K}_j^\tau\|_{\mathcal{H}}^2} \mathbb{K}_j^\tau \\
&= \sum_{j=1}^\infty \left(\frac{1}{\tau} \langle \Psi^D, \mathbf{K}_j^\tau \rangle_{L^2} + \langle \Psi^S, \mathbf{K}_j^\tau(0) \rangle \right) \frac{\mathbb{K}_j^\tau}{\|\mathcal{K}_{j_q}^\tau\|_{\mathcal{H}_1}^2},
\end{aligned}
$$

where we have used the identity (3.34) in the last equality above.

Now, let $\Psi^D \in L^2([-\tau, 0]; \mathbb{R}^d)$ to be the zero-function and Ψ^S to be an arbitrary vector $\alpha \in \mathbb{R}^d$. For such a Ψ, by equalizing respectively the D-components and S-components of the RHS and LHS of (3.39), we obtain respectively (3.35) and (3.36).

The identities (3.37) and (3.38) also follow directly from (3.39) by considering $\Psi^D = f$ and $\Psi^S = 0 \in \mathbb{R}^d$.

$\square$

4. Galerkin approximation: Rigorous results

In this section, we establish the convergence of the Galerkin scheme based on the rescaled and vectorized Koornwinder polynomials of Section 3.3. These convergence results apply, as we shall see, to a broad class of nonlinear systems of DDEs.

As mentioned in the Introduction and in Remark 2.1-(iii), the advantage of the Koornwinder basis relies on the facts that the constitutive basis functions are orthogonal and belong each to the domain of the linear operator associated with a given DDE. In particular, there is no discontinuity at the right end point for each basis element, by construction; see Section 3. Thanks to these properties of the basis functions, convergence results for the associated Galerkin systems can be derived in a straightforward fashion (see Corollary 4.1) and under useful criteria on the nonlinear terms (see Corollaries 4.2 and 4.3), compared to other Galerkin schemes built from other bases; see, e.g., [Kap86, KS87, Vya12, WC05] and references therein.

In the following, we first present in Section 4.1 a general convergence result for Galerkin approximations of abstract nonlinear ODEs in Hilbert spaces by relying essentially on the theory of semigroups and the Trotter-Kato theorem [Paz83, Thm. 4.5, p. 88]. The result is then applied to the DDE context in Section 4.2. General examples are provided in Section 4.3.

4.1. Galerkin approximations of nonlinear ODEs in Hilbert spaces.

We first present a general convergence result for Galerkin approximations of abstract nonlinear differential equations in a Hilbert space X, endowed with a norm $\|\cdot\|_X$. The mathematical setting is somewhat classical but we recall it below for the reader's convenience and later use.

In that respect, we assume in this Section the linear operator $\mathcal{L}$ to be the infinitesimal generator of a C_0-semigroup of bounded linear operators $T(t)$ on X. Recall that in that case the domain $D(\mathcal{L})$ of $\mathcal{L}$ is dense in X and that $\mathcal{L}$ is a closed operator; see [Paz83, Cor. 2.5, p. 5].

Under these assumptions, recall that there exists $M \geq 1$ and $\omega \geq 0$ [Paz83, Thm. 2.2, p. 4] such that

$$
(4.1) \qquad \|T(t)\| \leq M e^{\omega t}, \qquad t \geq 0,
$$

where $\|\cdot\|$ denotes the operator norm subordinated to $\|\cdot\|_X$.

We are concerned with finite-dimensional approximations of the following initial-value problem:

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \mathcal{L}u + \mathcal{G}(u),$$

(4.2)

$$u(0) = u_0,$$

where $u_0 \in X$.

A *mild solution* of (4.2) over $[0, T]$, will be any function $u \in C([0, T], X)$ such that for $u_0 \in X$,

$$u(t) = T(t)u_0 + \int_0^t T(t - s)\mathcal{G}(u(s))\mathrm{d}s.$$

(4.3)

Let $\{X_N : N \in \mathbb{N}\}$ be a sequence of subspaces of X associated with *orthogonal projectors*

(4.4)
$$\Pi_N : X \to X_N,$$

such that

(4.5)
$$\|\Pi_N - \mathrm{Id}_X\| \underset{N \to \infty}{\longrightarrow} 0,$$

and

(4.6)
$$X_N \subset D(\mathcal{L}), \ \forall N.$$

The corresponding Galerkin approximation of (4.2) associated with X_N is then given by:

$$\frac{\mathrm{d}u_N}{\mathrm{d}t} = \mathcal{L}_N u_N + \Pi_N \mathcal{G}(u_N),$$

(4.7)

$$u_N(0) = \Pi_N u_0, \ u_0 \in X,$$

where

(4.8)
$$\mathcal{L}_N := \Pi_N \mathcal{L} \Pi_N : X \to X_N.$$

In particular, the domain $D(\mathcal{L}_N)$ of $\mathcal{L}_N$ is X, because of (4.6).

As we will see in Section 4.2, the choice of vectorized Koornwinder polynomials as a basis function will allow us to define subspaces X_N naturally associated with orthogonal projectors Π_N that satisfy the above properties in contrast to other polynomial functions used for (non-standard) Galerkin approximation or other Legendre-tau approximations of systems of DDEs used so far; see e.g. [IT86, Kap86]. See also Remark 2.1-iii).

These nice properties will allow us also to rely on the following general convergence result regarding *standard* Galerkin schemes, for the case of nonlinear systems of DDEs; see Section 4.2 below.

Theorem 4.1. Let $\mathcal{L}$ and $\{X_N\}_{N \geq 0}$ be as described above. Assume furthermore the following set of assumptions:

(A1) For each $N \in \mathbb{N}$, the linear flow $e^{\mathcal{L}_N t} : X_N \to X_N$ extends to a C_0-semigroup $T_N(t)$ on X. Furthermore the following uniform bound is satisfied by the family $\{T_N(t)\}_{N \geq 0, t \geq 0}$

(4.9)
$$\|T_N(t)\| \leq Me^{\omega t}, \quad N \geq 0, \quad t \geq 0,$$

where $\|T_N(t)\| = \sup\{\|T_N(t)x\|_X, \ \|x\|_X \leq 1, x \in X\}$.

(A2) The following convergence holds

(4.10)
$$\lim_{N \to \infty} \|\mathcal{L}_N \phi - \mathcal{L}\phi\|_X = 0, \quad \forall \phi \in D(\mathcal{L}).$$

(A3) $\mathcal{G}$ is globally Lipschitz.

Then for any $u_0 \in X$, there exists a unique mild solution of (4.2) and such a solution can be approximated uniformly on each bounded interval $[0, T]$ by the sequence $\{t \mapsto u_N(t; \Pi_N u_0)\}_{N \geq 0}$ of mild solutions obtained from (4.7), i.e.:

$$(4.11) \qquad \lim_{N \to \infty} \sup_{t \in [0,T]} \|u_N(t; \Pi_N u_0) - u(t; u_0)\|_X = 0, \qquad \forall\, T > 0.$$

Proof. Recall that the existence and uniqueness of solutions to Eq. (4.3) emanating from any initial data $u_0 \in X$ can be proved by a fixed point argument in $C([0, T], X)$ as in the proof of e.g. [CH98, Prop. 4.3.3], by relaxing the semigroup of contractions requirement therein to the C_0-semigroup setting adopted here; see also [Lun04, Thm. 6.1.1].

Given $u_0 \in X$, let u be thus the unique mild solution of Eq. (4.2). By the variation-of-constants formula applied to Eq. (4.7) we have on the other hand, for $0 \leq t \leq T$,

$$(4.12) \qquad u_N(t) = e^{\mathcal{L}_N t} \Pi_N u_0 + \int_0^t e^{\mathcal{L}_N(t-s)} \Pi_N \mathcal{G}(u_N(s)) \mathrm{d}s.$$

Then $v_N(t) = u(t) - u_N(t)$ satisfies

$$v_N(t) = T(t)u_0 - e^{\mathcal{L}_N t} \Pi_N u_0 + \int_0^t T(t-s)\mathcal{G}(u(s))\mathrm{d}s - \int_0^t e^{\mathcal{L}_N(t-s)} \Pi_N \mathcal{G}(u_N(s))\mathrm{d}s$$

$$(4.13) \qquad = T(t)u_0 - e^{\mathcal{L}_N t} \Pi_N u_0 + \int_0^t \big(T(t-s) - e^{\mathcal{L}_N(t-s)}\Pi_N\big)\mathcal{G}(u(s))\mathrm{d}s$$

$$+ \int_0^t e^{\mathcal{L}_N(t-s)}\Pi_N\big(\mathcal{G}(u(s)) - \mathcal{G}(u_N(s))\big)\mathrm{d}s.$$

Let us introduce

$$(4.14) \qquad \begin{aligned} r_N(s) &:= \|u(s) - u_N(s)\|_X, \\ \epsilon_N(u_0) &:= \sup_{t \in [0,T]} \|T(t)u_0 - e^{\mathcal{L}_N t}\Pi_N u_0\|_X, \\ d_N(s) &:= \sup_{t \in [s,T]} \|\big(T(t-s) - e^{\mathcal{L}_N(t-s)}\Pi_N\big)\mathcal{G}(u(s))\|_X. \end{aligned}$$

We obtain then from (4.13) that

$$r_N(t) \leq \epsilon_N(u_0) + \int_0^t d_N(s)\mathrm{d}s + \int_0^t \|e^{\mathcal{L}_N(t-s)}\Pi_N\big(\mathcal{G}(u(s)) - \mathcal{G}(u_N(s))\big)\|_X \mathrm{d}s$$

$$(4.15) \qquad \leq \epsilon_N(u_0) + \int_0^t d_N(s)\mathrm{d}s + M\mathrm{Lip}(\mathcal{G})\int_0^t e^{\omega(t-s)} r_N(s)\mathrm{d}s$$

$$\leq \epsilon_N(u_0) + \int_0^T d_N(s)\mathrm{d}s + M\mathrm{Lip}(\mathcal{G})e^{\omega T}\int_0^t r_N(s)\mathrm{d}s,$$

where we have used the global Lipschitz condition on $\mathcal{G}$ and (4.9) to derive the second inequality.

It follows then from Gronwall's inequality that

$$(4.16) \qquad r_N(t) \leq \Big(\epsilon_N(u_0) + \int_0^T d_N(s)\mathrm{d}s\Big)e^{M\mathrm{Lip}(\mathcal{G})Te^{\omega T}}, \qquad \forall\, t \in [0, T].$$

We are thus left with the estimation of $\epsilon_N(u_0)$ and $\int_0^T d_N(s)\mathrm{d}s$ as $N \to \infty$. The assumptions **(A1)**–**(A2)** allow us to use a version of the Trotter-Kato theorem [Paz83, Thm. 4.5, p.88][6] which implies

---

[6]Recall that because $\mathcal{L}$ is the generator of a C_0-semigroup $T(t)$ on X, it satisfies $\|T(t)\| \leq Me^{\omega t}$, and as a consequence the resolvent set of $\mathcal{L}$ contains the interval $]\omega, \infty[$; see [Paz83, Thm. 5.3 p. 20]. In particular, for any $f \in \mathcal{H}$ and any

together with (4.5) that

$$\lim_{N \to \infty} e^{\mathcal{L}_N t} \Pi_N \phi = T(t)\phi, \quad \forall \, \phi \in X, \tag{4.17}$$

uniformly for t in bounded intervals.

It follows that

$$\lim_{N \to \infty} \epsilon_N(u_0) = 0, \quad \forall \, u_0 \in X, \tag{4.18}$$

and that d_N converges point-wisely to zero on $[0, T]$, i.e.

$$\lim_{N \to \infty} d_N(s) = 0, \quad \forall \, s \in [0, T]. \tag{4.19}$$

On the other hand, from (4.1), (4.9), and **(A3)**, we get

$$\begin{aligned}
\|\big(T(t-s) - e^{\mathcal{L}_N(t-s)}\Pi_N\big)\mathcal{G}(u(s))\|_X &\leq 2Me^{\omega(t-s)}\|\mathcal{G}(u(s))\|_X \\
&\leq 2Me^{\omega(t-s)}\big(\mathrm{Lip}(\mathcal{G})\|u(s)\|_X + \|\mathcal{G}(0)\|_X\big),
\end{aligned} \tag{4.20}$$

which implies

$$d_N(s) \leq 2Me^{\omega T}\big(\mathrm{Lip}(\mathcal{G})\|u(s)\|_X + \|\mathcal{G}(0)\|_X\big), \quad \forall \, s \in [0, T]. \tag{4.21}$$

Since $u \in C([0, T], X)$, $s \mapsto \|u(s)\|_X$ is integrable on $[0, T]$, and the Lebesgue's dominated convergence theorem allows us then to conclude from (4.19) and (4.21) that

$$\lim_{N \to \infty} \int_0^T d_N(s)\mathrm{d}s = 0. \tag{4.22}$$

The desired uniform convergence estimate (4.11) is then trivially obtained from (4.16).

$\square$

4.2. **Galerkin approximations of nonlinear systems of DDEs.** In this section, given the Hilbert product space

$$\mathcal{H} := L^2([-\tau, 0); \mathbb{R}^d) \times \mathbb{R}^d, \ \ d \geq 1,$$

endowed with the inner product (2.2), we restrict our attention to the following abstract ODE:

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \mathcal{A}u + \mathcal{F}(u), \tag{4.23}$$

where $\mathcal{F}$ is a nonlinear operator that will be specified later on, and where — given L_D, a bounded linear operator from $H^1([-\tau, 0); \mathbb{R}^d)$ to $\mathbb{R}^d$ and, L_S, a bounded linear operator from $\mathbb{R}^d$ to $\mathbb{R}^d$ — the linear operator $\mathcal{A}$ is given by

$$[\mathcal{A}\Psi](\theta) := \begin{cases} \dfrac{\mathrm{d}^+\Psi^D}{\mathrm{d}\theta}, & \theta \in [-\tau, 0), \\ L_S\Psi^S + L_D\Psi^D, & \theta = 0, \end{cases} \tag{4.24}$$

for any $\Psi = (\Psi^D, \Psi^S)$ that lives in the domain, $D(\mathcal{A})$, defined as

$$D(\mathcal{A}) := \Big\{ \Psi \in \mathcal{H} : \Psi^D \in H^1([-\tau, 0); \mathbb{R}^d), \lim_{\theta \to 0^-} \Psi^D(\theta) = \Psi^S \Big\}. \tag{4.25}$$

Such an abstract setting arises naturally in the reformulation of a broad class of nonlinear systems of DDEs as an abstract ODE in $\mathcal{H}$; see e.g. [BHS83, CZ95]. Examples of operators L_D depending explicitly on the delay τ are given below; see (4.47).

$\lambda > \omega$, the equation $(\lambda I - \mathcal{L})x = f$ possesses a unique solution $x \in D(\mathcal{L})$, which implies in particular that $(\lambda I - \mathcal{L})D(\mathcal{L})$ is dense in X as required by the version of the Trotter-Kato theorem used here. This explains why this density requirement, consequence of our working assumptions, is omitted in the formulation of Theorem 4.1.

It is well-known that under these assumptions, the operator $\mathcal{A}$ defines a C_0-semigroup on $\mathcal{H}$ [BHS83, Thm. 2.3], and in particular $\mathcal{A}$ is dense in $\mathcal{H}$ and is a closed operator.

We turn now to the definition of the subspaces X_N and Π_N of the previous section. For each positive integer N, we define the Nd-dimensional subspace $\mathcal{H}_N \subset \mathcal{H}$ to be spanned by the first Nd vectorized Koornwinder polynomials introduced in (3.32), namely

$$(4.26) \qquad \mathcal{H}_N = \mathrm{span}\Big\{ \mathbb{K}_1^\tau, \cdots, \mathbb{K}_{Nd}^\tau \Big\}.$$

As noted in Section 3.3, these polynomials are orthogonal for the inner product with a point mass such as defined in (2.2).

The subspace $\mathcal{H}_N$ is thus naturally associated with an orthogonal projector Π_N, as required in the previous section. The approximation property (4.5) is satisfied due to the density arguments outlined in Appendix A.

Recall finally that by construction $\mathbb{K}_j^\tau \in D(\mathcal{A})$ for any $j \in \mathbb{N}^*$, and therefore

$$(4.27) \qquad \mathcal{H}_N \subset D(\mathcal{A}).$$

The corresponding N-dimensional Galerkin approximation of Eq. (4.23) reads then:

$$(4.28) \qquad \frac{\mathrm{d}u_N}{\mathrm{d}t} = \mathcal{A}_N u_N + \Pi_N \mathcal{F}(u_N),$$

with

$$(4.29) \qquad \mathcal{A}_N := \Pi_N \mathcal{A} \Pi_N,$$

which is therefore well defined on $\mathcal{H}$ because of (4.27).

We are now in position to check Conditions **(A1)** and **(A2)** of Theorem 4.1. To check Condition **(A1)**, we will make usage of the following extension of the linear flow $e^{\mathcal{A}_N t}$:

$$(4.30) \qquad T_N(t)u = e^{\mathcal{A}_N t}\Pi_N u + (I - \Pi_N)u, \ u \in \mathcal{H}.$$

Such an extension leads naturally to a C_0-semigroup on $\mathcal{H}$. The stability condition (4.9), will require however some specifications of the operator L_D that will be made clear later.

Condition **(A2)** can be however checked in the general setting by making an appropriate use of the properties of the vectorized Koornwinder polynomials summarized in Lemma 3.2. More precisely,

Lemma 4.1. Let $\mathcal{H}_N$ be the subspace defined in (4.26). Then for $\mathcal{A}$ defined in (4.24) and $\mathcal{A}_N$ defined in (4.29) associated with the orthogonal projector Π_N onto $\mathcal{H}_N$, we have

$$(4.31) \qquad \lim_{N \to \infty} \|\mathcal{A}_N \Psi - \mathcal{A}\Psi\|_{\mathcal{H}} = 0, \qquad \forall\, \Psi \in D(\mathcal{A}).$$

Proof. By construction $\mathbb{K}_j^\tau \in D(\mathcal{A})$ for each $j \in \mathbb{N}^*$. Since $\big\{ \frac{\mathbb{K}_j^\tau}{\|\mathbb{K}_j^\tau\|_{\mathcal{H}}} : j \in \mathbb{N}^* \big\}$ forms a Hilbert basis of $\mathcal{H}$, it suffices to show that

$$(4.32) \qquad \lim_{N \to \infty} \|\mathcal{A}_N \Psi - \mathcal{A}\Psi\|_{\mathcal{H}} = 0, \ \ \Psi \in \bigcup_{k \geq 1} \mathcal{H}_k.$$

We recall from (3.34) that $\|\mathbb{K}_j^\tau\|_{\mathcal{H}} = \|\mathcal{K}_{j_q}^\tau\|_{\mathcal{H}_1}$ for all $j \in \mathbb{N}$. It follows then that the orthogonal projector Π_N associated with the subspace $\mathcal{H}_N$ takes the following explicit form:

$$
\begin{aligned}
(4.33) \quad \Pi_N \Psi &= \sum_{j=1}^{Nd} \frac{\langle \Psi, \mathbb{K}_j^\tau \rangle_{\mathcal{H}}}{\|\mathcal{K}_{j_q}^\tau\|_{\mathcal{H}_1}^2} \mathbb{K}_j^\tau \\
&= \sum_{j=1}^{Nd} \left(\frac{1}{\tau} \langle \Psi^D, \mathbf{K}_j^\tau \rangle_{L^2} + \langle \Psi^S, \mathbf{K}_j^\tau(0) \rangle \right) \frac{\mathbb{K}_j^\tau}{\|\mathcal{K}_{j_q}^\tau\|_{\mathcal{H}_1}^2} \\
&= \begin{pmatrix} p_N & q_N \\ p_N' & q_N' \end{pmatrix} \begin{pmatrix} \Psi^D \\ \Psi^S \end{pmatrix},
\end{aligned}
$$

where the operators p_N, p_N', q_N, q_N' are defined as following:

$$
(4.34a) \qquad p_N : L^2([-\tau, 0]; \mathbb{R}^d) \to L^2([-\tau, 0]; \mathbb{R}^d), \qquad \Psi^D \mapsto \sum_{j=1}^{Nd} \frac{\langle \Psi^D, \mathbf{K}_j^\tau \rangle_{L^2}}{\tau \|\mathcal{K}_{j_q}^\tau\|_{\mathcal{H}_1}^2} \mathbf{K}_j^\tau;
$$

$$
(4.34b) \qquad p_N' : L^2([-\tau, 0]; \mathbb{R}^d) \to \mathbb{R}^d, \qquad \Psi^D \mapsto \sum_{j=1}^{Nd} \frac{\langle \Psi^D, \mathbf{K}_j^\tau \rangle_{L^2}}{\tau \|\mathcal{K}_{j_q}^\tau\|_{\mathcal{H}_1}^2} \mathbf{K}_j^\tau(0);
$$

$$
(4.34c) \qquad q_N : \mathbb{R}^d \to L^2([-\tau, 0]; \mathbb{R}^d), \qquad \Psi^S \mapsto \sum_{j=1}^{Nd} \frac{\langle \Psi^S, \mathbf{K}_j^\tau(0) \rangle}{\|\mathcal{K}_{j_q}^\tau\|_{\mathcal{H}_1}^2} \mathbf{K}_j^\tau;
$$

$$
(4.34d) \qquad q_N' : \mathbb{R}^d \to \mathbb{R}^d, \qquad \Psi^S \mapsto \sum_{j=1}^{Nd} \frac{\langle \Psi^S, \mathbf{K}_j^\tau(0) \rangle}{\|\mathcal{K}_{j_q}^\tau\|_{\mathcal{H}_1}^2} \mathbf{K}_j^\tau(0).
$$

In the following, we arbitrarily fix $\Phi \in \mathcal{H}_k$ for some integer $k > 0$. Now let us choose N such that $Nd \geq k$, then $\Pi_{Nd}\Phi = \Phi$, and we get for each such N

$$
(4.35) \qquad \mathcal{A}_N \Phi = \Pi_N \mathcal{A} \Pi_N \Phi = \Pi_N \mathcal{A} \Phi = \begin{pmatrix} p_N \frac{\mathrm{d}^+}{\mathrm{d}\theta} \Phi^D + q_N(L_S \Phi^S + L_D \Phi^D) \\ p_N' \frac{\mathrm{d}^+}{\mathrm{d}\theta} \Phi^D + q_N'(L_S \Phi^S + L_D \Phi^D) \end{pmatrix}.
$$

We obtain then

$$
(4.36) \qquad (\mathcal{A} - \mathcal{A}_N)\Phi = \begin{pmatrix} (I^D - p_N)\frac{\mathrm{d}^+}{\mathrm{d}\theta} \Phi^D - q_N\big(L_S \Phi^S + L_D \Phi^D\big) \\ -p_N' \frac{\mathrm{d}^+}{\mathrm{d}\theta} \Phi^D + (I^S - q_N')\big(L_S \Phi^S + L_D \Phi^D\big) \end{pmatrix}, \quad \text{if } Nd \geq k,
$$

where I^D and I^S denote the identity maps on $L^2([-\tau, 0]; \mathbb{R}^d)$ and $\mathbb{R}^d$, respectively.

We show below that the RHS of (4.36) converges to zero. Let us begin with the term $(I^D - p_N)\frac{\mathrm{d}^+}{\mathrm{d}\theta} \Phi^D$. Note that by comparing the definition of p_N give by (4.34a) and the decomposition of $L^2([-\tau, 0]; \mathbb{R}^d)$ functions given by (3.37), we see that for each $f \in L^2([-\tau, 0]; \mathbb{R}^d)$, the term $p_N f$ is the partial sum of the first N terms in the corresponding decomposition. It follows then that

$$
(4.37) \qquad \lim_{N \to \infty} \|(I^D - p_N)f\|_{L^2} = 0, \quad \forall f \in L^2([-\tau, 0]; \mathbb{R}^d).
$$

Since $\Phi \in \mathcal{H}_k \subset D(\mathcal{A})$, it holds that $\frac{\mathrm{d}^+}{\mathrm{d}\theta} \Phi^D \in L^2([-\tau, 0]; \mathbb{R}^d)$. We obtain then from (4.37) that

$$
(4.38) \qquad \lim_{N \to \infty} \left\| (I^D - p_N)\Big(\frac{\mathrm{d}^+}{\mathrm{d}\theta} \Phi^D\Big) \right\|_{L^2} = 0.
$$

We turn now to the estimates for $q_N(L_S\Phi^S + L_D\Phi^D)$. By the definition of q_N in (4.34c), we get

$$(4.39) \qquad q_N\big(L_S\Phi^S + L_D\Phi^D\big) = \sum_{j=1}^{Nd} \frac{\langle L_S\Phi^S + L_D\Phi^D, \mathbf{K}_j^\tau(0)\rangle}{\|\mathcal{K}_{j_q}^\tau\|_{\mathcal{H}_1}^2} \mathbf{K}_j^\tau.$$

It follows then from the identity (3.35) that

$$(4.40) \qquad \lim_{N\to\infty} \Big\| q_N\big(L_S\Phi^S + L_D\Phi^D\big) \Big\|_{L^2} = 0.$$

For the term $p_N' \frac{\mathrm{d}^+}{\mathrm{d}\theta}\Phi^D$, since $\frac{\mathrm{d}^+}{\mathrm{d}\theta}\Phi^D \in L^2([-\tau,0);\mathbb{R}^d)$, it follows from the definition of p_N' given in (4.34b) and the identity (3.38) that

$$(4.41) \qquad \lim_{N\to\infty} \left| p_N' \frac{\mathrm{d}^+}{\mathrm{d}\theta}\Phi^D \right| = 0,$$

where $|\cdot|$ denotes the Euclidean norm of $\mathbb{R}^d$.

By using the identity (3.36), we also get

$$(4.42) \qquad \lim_{N\to\infty} \left| (I^S - q_N')\big(L_S\Phi^S + L_D\Phi^D\big) \right| = 0.$$

Now, by using the estimates (4.38), (4.40), (4.41), and (4.42), we get from (4.36) that

$$(4.43) \qquad \lim_{N\to\infty} \|(\mathcal{A} - \mathcal{A}_N)\Phi\|_{\mathcal{H}} = 0,$$

and (4.32) follows.

$\square$

Remark 4.1. We explain here how the truncated linear operator $\mathcal{A}_N$ defined in (4.35) is related to an interesting class of nonlocal linear PDEs. For the sake of clarity, we discuss the case $d = 1$. For convenience, let us write $v(\theta) = \Phi^D(\theta)$ and, recalling e.g. Eq. (1.2) in the Introduction, replace $\mathrm{d}^+v/\mathrm{d}\theta$ by v_θ in (4.35).

One then obtains that, when $v \in \mathcal{H}_N$,

$$(4.44) \qquad \begin{pmatrix} p_N v_\theta \\ p_N' v_\theta \end{pmatrix} = \begin{pmatrix} v_\theta \\ v_\theta(0) \end{pmatrix} - \sum_{n=0}^{N-1} \frac{v_\theta(0)}{\|\mathcal{K}_n^\tau\|_{\mathcal{H}}^2} \mathcal{K}_n^\tau.$$

Next, we use the expressions of $\mathcal{A}_N$, p_N and q_N — given, respectively, in (4.35), (4.34a) and (4.34c) — to note that the D-component v of any solution u of

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \mathcal{A}_N u,$$

which emanates from initial data taken[7] in $\mathcal{H}_N$, satisfies the following *nonlocal linear PDE*:

$$(4.45) \qquad \partial_t v = \partial_\theta v + b_N(\theta)\Big(L_S v(t,0) + L_D v - \partial_\theta v\big|_{\theta=0}\Big);$$

here

$$(4.46) \qquad b_N(\theta) = \sum_{n=0}^{N-1} \frac{K_n^\tau(\theta)}{\|\mathcal{K}_n^\tau\|_{\mathcal{H}}^2},$$

and the rescaled Koornwinder polynomials K_n^τ are given by (3.18), since $d = 1$. We see therewith that the Galerkin scheme used herein introduces a nonlocal perturbation term with respect to the

---

[7]For such initial data, the solution stays in $\mathcal{H}_N$, by the definition of $\mathcal{A}_N$.

D-component of $\mathcal{A}$ given in (4.24). This perturbation term results from the difference between the S-component of $\mathcal{A}$ and the derivative at 0, when applied to functions in $\mathcal{H}_N$.

From Lemma 4.1 above and Lemmas 4.2 and 4.3 below, one can infer by the Trotter-Kato theorem that the effects on the solutions of Eq. (4.45) of this nonlocal perturbation — which vanishes as $N \to \infty$, due to (3.22) — do not lead to a degenerate situation and that actually these solutions converge over finite intervals to those of the local PDE, $\partial_t v = \partial_\theta v$. This nice convergence is due to the key properties of the Koornwinder polynomials, as summarized in Lemma 3.1 for $d = 1$, and in Lemma 3.2 for the multidimensional case; see also Fig. 1 for the nature of the approximation at $\theta = 0$.

To conclude this remark, we return now to the discussion in the Introduction concerning the approximation of discontinuities that arise, for instance, in the first-order derivative of a DDE's solution, cf. [GZT08, Appendix A and references therein].

In Table 1, we report the corresponding differences over the interval $[0, 2]$ between the analytic solution of Eq. (1.1) with $\tau = 1$ and history $x(t) \equiv 1$, on the one hand, and low-dimensional Galerkin approximations obtained by application of the formulas derived hereafter in Section 5, on the other.

TABLE 1. Errors in Galerkin approximation of Eq. (1.1)

N	Max. error in Galerkin solution	Max. error in 1st-order derivative of Galerkin solution
4	6.9×10^{-3}	5.9×10^{-2}
8	9.3×10^{-4}	2.2×10^{-2}
16	2.3×10^{-4}	1.1×10^{-2}
32	9.0×10^{-5}	5.4×10^{-3}

The second column of this table is obviously consistent with the rigorous convergence result of Corollary 4.1 proved below. The third column shows that, furthermore, the aforementioned discontinuities present in the derivative of the DDE's solutions are well captured by the proposed methodology as well. □

In the following, we restrict the linear operator $\mathcal{A}$ defined in (4.24) to be such that $L_S : \mathbb{R}^d \to \mathbb{R}^d$ is a bounded linear operator, and L_D is defined by

(4.47)
$$L_D : H^1([-\tau, 0); \mathbb{R}^d) \to \mathbb{R}^d,$$
$$\Psi^D \mapsto B\Psi^D(-\tau) + \int_{-\tau}^0 C(s)\Psi^D(s)\mathrm{d}s,$$

with $B : \mathbb{R}^d \to \mathbb{R}^d$ being a bounded linear operator[8], and $C(\cdot) \in L^2([-\tau, 0); \mathbb{R}^{d \times d})$.

In the following preparatory lemma, Lemma 4.2, we recall by means of basic estimates, that the existence of $\omega > 0$ such that $\mathcal{A} - \omega I$ is dissipative in $\mathcal{H}$, can be guaranteed. This fact will be used to establish a stability condition of type (4.9) (with $M = 1$) for the semigroups $T(t)$ and $T_N(t)$, generated respectively by $\mathcal{A}$ and its finite-dimensional approximation $\mathcal{A}_N$; see Lemma 4.3. The proofs of these Lemmas are quite straightforward but are reproduced below for the sake of completeness.

Lemma 4.2. Let $\mathcal{A}$ be defined such as in (4.24) with L_D such as specified in (4.47) and $L_S : \mathbb{R}^d \to \mathbb{R}^d$ to be a bounded linear operator. Then

(4.48) $$\langle \mathcal{A}\Psi, \Psi \rangle_{\mathcal{H}} \leq \omega \|\Psi\|_{\mathcal{H}}^2, \quad \forall \, \Psi \in D(\mathcal{A}),$$

---

[8]Note that $\Psi^D(-\tau)$ is well-defined for $\Psi^D \in H^1([-\tau, 0); \mathbb{R}^d)$, since the latter Sobolev space is continuously embedded in the space of continuous functions $C([-\tau, 0]; \mathbb{R}^d)$; see [Bré10, Thm. 8.8].

with[9]

$$\omega = \left(1 + \frac{1}{2\tau} + |L_S| + \frac{\tau}{2}|B|^2 + \frac{\tau}{4}\|C\|_{L^2}^2\right). \tag{4.49}$$

Proof. Let $\Psi \in D(\mathcal{A})$, then by the definition of $\mathcal{A}$ given in (4.24), we have

$$\langle \mathcal{A}\Psi, \Psi \rangle_{\mathcal{H}} = \frac{1}{\tau}\int_{-\tau}^{0}\left\langle\frac{d^+\Psi^D}{d\theta}(\theta), \Psi^D(\theta)\right\rangle d\theta + \langle L_S\Psi^S, \Psi^S\rangle + \langle L_D\Psi^D, \Psi^S\rangle. \tag{4.50}$$

Note that

$$\frac{1}{\tau}\int_{-\tau}^{0}\left\langle\frac{d^+\Psi^D}{d\theta}(\theta), \Psi^D(\theta)\right\rangle d\theta = \frac{1}{2\tau}\int_{-\tau}^{0} d|\Psi^D(\theta)|^2$$

$$= \frac{1}{2\tau}\left(|\Psi^S|^2 - |\Psi^D(-\tau)|^2\right). \tag{4.51}$$

Using the definition of L_D given in (4.47), we obtain

$$\langle L_D\Psi^D, \Psi^S\rangle = \left\langle B\Psi^D(-\tau) + \int_{-\tau}^{0}C(\theta)\Psi^D(\theta)d\theta, \Psi^S\right\rangle$$

$$\leq |B||\Psi^D(-\tau)||\Psi^S| + \|C\|_{L^2}\|\Psi^D\|_{L^2}|\Psi^S|. \tag{4.52}$$

It follows then from Young's inequality that

$$\langle L_D\Psi^D, \Psi^S\rangle \leq \left(\frac{1}{2\tau}|\Psi^D(-\tau)|^2 + \frac{\tau}{2}|B|^2|\Psi^S|^2\right)$$

$$+ \left(\frac{1}{\tau}\|\Psi^D\|_{L^2}^2 + \frac{\tau}{4}\|C\|_{L^2}^2|\Psi^S|^2\right). \tag{4.53}$$

Note also that

$$\langle L_S\Psi^S, \Psi^S\rangle \leq |L_s||\Psi^S|^2. \tag{4.54}$$

Now, by using (4.51), (4.53), and (4.54) in (4.50), we get

$$\langle \mathcal{A}\Psi, \Psi \rangle_{\mathcal{H}} \leq \left(\frac{1}{2\tau} + |L_S| + \frac{\tau}{2}|B|^2 + \frac{\tau}{4}\|C\|_{L^2}^2\right)|\Psi^S|^2 + \frac{1}{\tau}\|\Psi^D\|_{L^2}^2$$

$$\leq \left(1 + \frac{1}{2\tau} + |L_S| + \frac{\tau}{2}|B|^2 + \frac{\tau}{4}\|C\|_{L^2}^2\right)\left(|\Psi^S|^2 + \frac{1}{\tau}\|\Psi^D\|_{L^2}^2\right)$$

$$= \left(1 + \frac{1}{2\tau} + |L_S| + \frac{\tau}{2}|B|^2 + \frac{\tau}{4}\|C\|_{L^2}^2\right)\|\Psi\|_{\mathcal{H}}^2, \tag{4.55}$$

leading thus to the desired estimate.

$\square$

Lemma 4.3. Let $\mathcal{A}$ be defined such as in (4.24) with L_D such as specified in (4.47) and $L_S : \mathbb{R}^d \to \mathbb{R}^d$ to be a bounded linear operator.

Then, the linear semigroups $T(t)$ and $T_N(t)$ generated respectively by $\mathcal{A}$ and $\mathcal{A}_N$ defined in (4.29), satisfy

$$\|T(t)\| \leq e^{\omega t} \quad \text{and} \quad \|T_N(t)\| \leq e^{\omega t}, \qquad t \geq 0, \tag{4.56}$$

with ω given by (4.49).

---

[9]Throughout this article, we will denote indistinguishably by $|\cdot|$, either the Euclidean norm of a vector in $\mathbb{R}^d$, or its subordinated (operator) norm, in the case of a $d \times d$ matrix. It should be clear from the context which norm is used.

Proof. Since $T(t)$ is a C_0-semigroup with infinitesimal generator $\mathcal{A}$, we have that $T(t)u_0 \in D(\mathcal{A})$ for all $u_0 \in D(\mathcal{A})$, and that

$$(4.57) \qquad \frac{\mathrm{d}}{\mathrm{d}t}T(t)u_0 = \mathcal{A}T(t)u_0, \qquad \forall\, u_0 \in \mathcal{A},\ t \geq 0;$$

cf. [Paz83, Thm. 2.4 c) p.5].

We obtain thus

$$(4.58) \qquad \frac{\mathrm{d}}{\mathrm{d}t}\|T(t)u_0\|_{\mathcal{H}}^2 = 2\langle \mathcal{A}T(t)u_0, T(t)u_0\rangle_{\mathcal{H}} \leq 2\omega\|T(t)u_0\|_{\mathcal{H}}^2, \qquad \forall\, u_0 \in D(\mathcal{A}),$$

where we have used Lemma 4.2 to obtain the last inequality above with ω given by (4.49).

It follows then from Gronwall's inequality that

$$(4.59) \qquad \|T(t)u_0\|_{\mathcal{H}}^2 \leq e^{2\omega t}\|u_0\|_{\mathcal{H}}^2, \qquad \forall\, u_0 \in D(\mathcal{A}).$$

Since $D(\mathcal{A})$ is dense in $\mathcal{H}$ and $T(t)$ are bounded operators on $\mathcal{H}$, the estimate (4.59) still holds for general initial data in $\mathcal{H}$, leading in turn to

$$(4.60) \qquad \|T(t)\| \leq e^{\omega t}, \qquad t \geq 0.$$

The estimate for T_N is also trivial and proceeds as follows. First note that by the definition of T_N given by (4.30), we have

$$(4.61) \qquad \begin{aligned} \|T_N(t)u_0\|_{\mathcal{H}}^2 &= \langle e^{\mathcal{A}_N t}\Pi_N u_0 + (I - \Pi_N)u_0, e^{\mathcal{A}_N t}\Pi_N u_0 + (I - \Pi_N)u_0\rangle_{\mathcal{H}} \\ &= \langle e^{\mathcal{A}_N t}\Pi_N u_0, e^{\mathcal{A}_N t}\Pi_N u_0\rangle_{\mathcal{H}} + \langle (I - \Pi_N)u_0, (I - \Pi_N)u_0\rangle_{\mathcal{H}} \\ &= \|e^{\mathcal{A}_N t}\Pi_N u_0\|_{\mathcal{H}}^2 + \|(I - \Pi_N)u_0\|_{\mathcal{H}}^2. \end{aligned}$$

It follows that

$$(4.62) \qquad \begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\|T_N(t)u_0\|_{\mathcal{H}}^2 &= \frac{\mathrm{d}}{\mathrm{d}t}\|e^{\mathcal{A}_N t}\Pi_N u_0\|_{\mathcal{H}}^2 \\ &= 2\langle \mathcal{A}_N e^{\mathcal{A}_N t}\Pi_N u_0, e^{\mathcal{A}_N t}\Pi_N u_0\rangle_{\mathcal{H}}, \qquad \forall\, u_0 \in \mathcal{H}. \end{aligned}$$

Note also that

$$(4.63) \qquad \begin{aligned} \langle \mathcal{A}_N\Psi, \Psi\rangle_{\mathcal{H}} &= \langle \Pi_N\mathcal{A}\Pi_N\Psi, \Psi\rangle_{\mathcal{H}} \\ &= \langle \Pi_N\mathcal{A}\Pi_N\Psi, \Pi_N\Psi\rangle_{\mathcal{H}} + \langle \Pi_N\mathcal{A}\Pi_N\Psi, (I - \Pi_N)\Psi\rangle_{\mathcal{H}} \\ &= \langle \Pi_N\mathcal{A}\Pi_N\Psi, \Pi_N\Psi\rangle_{\mathcal{H}} \\ &= \langle \mathcal{A}\Pi_N\Psi, \Pi_N\Psi\rangle_{\mathcal{H}} \\ &\leq \omega\|\Pi_N\Psi\|_{\mathcal{H}} \\ &\leq \omega\|\Psi\|_{\mathcal{H}}. \end{aligned}$$

We obtain then from (4.62) that

$$(4.64) \qquad \begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\|T_N(t)u_0\|_{\mathcal{H}}^2 &\leq 2\omega\|e^{\mathcal{A}_N t}\Pi_N u_0\|_{\mathcal{H}}^2 \\ &\leq 2\omega\big(\|e^{\mathcal{A}_N t}\Pi_N u_0\|_{\mathcal{H}}^2 + \|(I - \Pi_N)u_0\|_{\mathcal{H}}^2\big) \\ &= 2\omega\|e^{\mathcal{A}_N t}\Pi_N u_0 + (I - \Pi_N)u_0\|_{\mathcal{H}}^2 \\ &= 2\omega\|T_N(t)u_0\|_{\mathcal{H}}^2, \qquad \forall\, u_0 \in \mathcal{H}. \end{aligned}$$

The desired estimate for $\|T_N(t)\|$ can be derived now from (4.64) by using Gronwall's inequality. $\qquad\square$

Remark 4.2. Note that the estimate about T_N in (4.56) shows in particular that solutions of (4.45) grow at most exponentially with a rate independent of N, and stay uniformly bounded over finite intervals. □

With these preparatory lemmas, we are now in position to obtain as corollaries of Theorem 4.1, the convergence results for the Galerkin approximation (4.28) of d-dimensional nonlinear systems of DDEs of the form

$$(4.65) \qquad \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = L_S\mathbf{x}(t) + B\mathbf{x}(t-\tau) + \int_{t-\tau}^t C(s-t)\mathbf{x}(s)\mathrm{d}s + \mathbf{F}\Big(\mathbf{x}(t), \int_{t-\tau}^t \mathbf{x}(s)\mathrm{d}s\Big),$$

where $\mathbf{F} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, and L_S, B, and C are as given in (4.47).

We first sate the result for the case of global Lipschitz nonlinearity, keeping in mind that already for the case of scalar DDEs ($d = 1$), chaotic dynamics can take place under such a simple nonlinear setting; see Section 6.1 for a numerical illustration.

Corollary 4.1. Let $\mathcal{A}$ be defined such as in (4.24) with L_D such as specified in (4.47) and $L_S : \mathbb{R}^d \to \mathbb{R}^d$ to be a bounded linear operator. Assume that the nonlinearity $\mathcal{F} : \mathcal{H} \to \mathcal{H}$ defined by

$$(4.66) \qquad [\mathcal{F}(\Psi)](\theta) := \begin{cases} 0, & \theta \in [-\tau, 0), \\ \mathbf{F}\Big(\Psi^S, \int_{-\tau}^0 \Psi^D(s)\mathrm{d}s\Big), & \theta = 0, \end{cases} \qquad \forall\, \Psi = (\Psi^D, \Psi^S) \in \mathcal{H},$$

is globally Lipschitz.

Then, for each $u_0 \in \mathcal{H}$, the mild solution of (4.28) emanating from $\Pi_N u_0$ converges uniformly to the mild solution of (4.23) emanating from u_0 on each bounded interval $[0, T]$, i.e.:

$$(4.67) \qquad \lim_{N\to\infty} \sup_{t\in[0,T]} \|u_N(t; \Pi_N u_0) - u(t; u_0)\|_{\mathcal{H}} = 0, \qquad \forall\, T > 0,\ u_0 \in \mathcal{H}.$$

Proof. This corollary is a direct consequence of Lemma 4.3 and Lemma 4.1, ensuring respectively, Conditions **(A1)** and **(A2)** of Theorem 4.1. □

In the next two corollaries, we relax the global Lipschitz condition assumed in Corollary 4.1 to a local Lipschitz condition in addition to either a sublinear growth for $\mathcal{F}$ (see Corollary 4.2) or an energy inequality satisfied by $\mathcal{F}$; see Corollary 4.3.

Corollary 4.2. Let $\mathcal{A}$ be defined in (4.24) with L_D specified in (4.47) and $L_S : \mathbb{R}^d \to \mathbb{R}^d$ to be a bounded linear operator.

Assume that the nonlinearity $\mathcal{F}$ given by (4.66) is locally Lipschitz in the sense that for all $r > 0$ there exists $L(r) > 0$ such that for any Ψ_1 and Ψ_2 in $\mathcal{H}$, we have

$$(4.68) \qquad \|\Psi_1\|_{\mathcal{H}} < r \text{ and } \|\Psi_2\|_{\mathcal{H}} < r \implies \|\mathcal{F}(\Psi_1) - \mathcal{F}(\Psi_2)\|_{\mathcal{H}} \le L(r)\|\Psi_1 - \Psi_2\|_{\mathcal{H}}.$$

Assume also that $\mathcal{F}$ satisfies the following sublinear growth:

$$(4.69) \qquad \|\mathcal{F}(\Psi)\|_{\mathcal{H}} \le \gamma_1\|\Psi\|_{\mathcal{H}} + \gamma_2, \qquad \forall\, \Psi \in \mathcal{H},$$

where $\gamma_1 > 0$ and $\gamma_2 \ge 0$.

Then, for each $u_0 \in \mathcal{H}$, the mild solution $u_N(t; \Pi_N u_0)$ of (4.28) emanating from $\Pi_N u_0$ and, the mild solution $u(t; u_0)$ of (4.23) emanating from u_0, do not blow up in any finite time. Moreover, $u_N(t; \Pi_N u_0)$ converges uniformly to $u(t; u_0)$ on each bounded interval $[0, T]$, i.e.:

$$(4.70) \qquad \lim_{N\to\infty} \sup_{t\in[0,T]} \|u_N(t; \Pi_N u_0) - u(t; u_0)\|_{\mathcal{H}} = 0, \qquad \forall\, T > 0,\ u_0 \in \mathcal{H}.$$

Proof. Recall that the local Lipschitz condition (4.68) on $\mathcal{F}$ ensures the existence and uniqueness of a local mild solution $u(t; u_0)$ to (4.23) emanating from any $u_0 \in \mathcal{H}$; see e.g. [CH98, Prop. 4.3.3].[10]

By recalling that $\|T(t)\|_{\mathcal{H}} \leq e^{\omega t}$ (see Lemma 4.3) and by using the sublinear growth assumption (4.69) on $\mathcal{F}$, we obtain for mild solutions

$$
\begin{aligned}
\|u(t)\|_{\mathcal{H}} &\leq \|T(t)u_0\|_{\mathcal{H}} + \int_0^t \|T(t-s)\mathcal{F}(u(s))\|_{\mathcal{H}} \mathrm{d}s \\
&\leq e^{\omega t}\|u_0\|_{\mathcal{H}} + \int_0^t e^{\omega(t-s)}\Big(\gamma_1\|u(s)\|_{\mathcal{H}} + \gamma_2\Big)\mathrm{d}s \\
&\leq e^{\omega t}\|u_0\|_{\mathcal{H}} + \frac{\gamma_2(e^{\omega t}-1)}{\omega} + \gamma_1 \int_0^t e^{\omega(t-s)}\|u(s)\|_{\mathcal{H}}\mathrm{d}s,
\end{aligned}
\tag{4.71}
$$

where the positive constant ω is given by (4.49).

A simple multiplication by $e^{-\omega t}$ to both sides of (4.71), leads then trivially to

$$
\begin{aligned}
e^{-\omega t}\|u(t)\|_{\mathcal{H}} &\leq \|u_0\|_{\mathcal{H}} + \frac{\gamma_2(1-e^{-\omega t})}{\omega} + \gamma_1 \int_0^t e^{-\omega s}\|u(s)\|_{\mathcal{H}}\mathrm{d}s \\
&\leq \|u_0\|_{\mathcal{H}} + \frac{\gamma_2}{\omega} + \gamma_1 \int_0^t e^{-\omega s}\|u(s)\|_{\mathcal{H}}\mathrm{d}s.
\end{aligned}
\tag{4.72}
$$

An application of the Gronwall's inequality to $v(t) := e^{-\omega t}\|u(t)\|_{\mathcal{H}}$, gives then

$$
\|u(t)\|_{\mathcal{H}} \leq \Big(\|u_0\|_{\mathcal{H}} + \frac{\gamma_2}{\omega}\Big)e^{(\omega+\gamma_1)t},
\tag{4.73}
$$

preventing thus the blow up of a mild solution in finite time.

Similarly, for mild solutions u_N of (4.28), we have

$$
\|u_N(t)\|_{\mathcal{H}} \leq \Big(\|u_0\|_{\mathcal{H}} + \frac{\gamma_2}{\omega}\Big)e^{(\omega+\gamma_1)t}.
\tag{4.74}
$$

by noting that $\|\Pi_N\| < 1$ for all $N \geq 1$, $\|T_N(t)\|_{\mathcal{H}} \leq e^{\omega t}$ (see Lemma 4.3), and by using the sublinear growth assumption (4.69) on $\mathcal{F}$, preventing also the blow up of any mild solution u_N of (4.28), in finite time.

Finally, (4.73) and (4.74) lead to

$$
\begin{aligned}
\|u(t; u_0)\|_{\mathcal{H}} &\leq C(T, \|u_0\|_{\mathcal{H}}), && \forall\, t \in [0, T], \\
\|u_N(t; \Pi_N u_0)\|_{\mathcal{H}} &\leq C(T, \|u_0\|_{\mathcal{H}}), && \forall\, t \in [0, T] \text{ and } N \in \mathbb{N}^*,
\end{aligned}
\tag{4.75}
$$

where

$$
C(T, \|u_0\|_{\mathcal{H}}) := \Big(\|u_0\|_{\mathcal{H}} + \frac{\gamma_2}{\omega}\Big)e^{(\omega+\gamma_1)T}.
$$

Now, we can follow the proof of Theorem 4.1 to obtain the desired convergence result (4.70), with the only difference consisting of the global Lipschitz estimates used therein, by the local Lipschitz condition (4.68) applied to Ψ such that

$$
\|\Psi\|_{\mathcal{H}} \leq 2C(T, \|u_0\|_{\mathcal{H}}).
$$

$\square$

Corollary 4.3. Let $\mathcal{A}$ be defined in (4.24) with L_D specified in (4.47) and L_S to be a bounded linear operator from $\mathbb{R}^d$ to $\mathbb{R}^d$.

---

[10] [CH98, Prop. 4.3.3] is derived for the case of contraction semigroups. However, the proof can be easily adapted to the case of more general C_0-semigroups $T(t)$ for which $\|T(t)\| \leq Me^{\omega t}$.

Assume that the nonlinearity $\mathcal{F}$ given by (4.66) is locally Lipschitz in the sense of (4.68). Assume also that the following energy inequality holds for $\mathcal{F}$

$$(4.76) \qquad \langle \mathcal{F}(\Psi), \Psi \rangle_{\mathcal{H}} \leq \gamma_1 \|\Psi\|_{\mathcal{H}}^2 + \gamma_2, \qquad \forall\, \Psi \in \mathcal{H},$$

where $\gamma_1 \in \mathbb{R}$ and $\gamma_2 \geq 0$.

Then, for each $u_0 \in D(\mathcal{A})$, the strong solution $u_N(t; \Pi_N u_0)$ of (4.28) emanating from $\Pi_N u_0$, and the strong solution[11] $u(t; u_0)$ of (4.23) emanating from u_0, do not blow up in any finite time. Moreover, $u_N(t; \Pi_N u_0)$ converges uniformly to $u(t; u_0)$ on each bounded interval $[0, T]$, i.e.:

$$(4.77) \qquad \lim_{N \to \infty} \sup_{t \in [0,T]} \|u_N(t; \Pi_N u_0) - u(t; u_0)\|_{\mathcal{H}} = 0, \qquad \forall\, T > 0,\ u_0 \in D(\mathcal{A}).$$

Furthermore, any strong solutions $v = u$ of (4.23) or $v = u_N$ of (4.28), emanating respectively from $v(0) = u_0 \in D(\mathcal{A})$ or $v(0) = \Pi_N u_0$, have their $\mathcal{H}$-norm controlled as follows:

$$(4.78) \qquad \|v(t)\|_{\mathcal{H}}^2 \leq e^{\kappa t} \|v(0)\|_{\mathcal{H}}^2 + \frac{2\gamma_2}{\kappa}(e^{\kappa t} - 1),\ t > 0,$$

where $\kappa = 2(\omega + \gamma_1)$, with ω given in (4.49).

Proof. Let $u_0 \in D(\mathcal{A})$, and let u be the mild solution of (4.23) emanating from u_0, such as ensured by the local Lipschitz condition on $\mathcal{F}$. Then by adapting the proof of e.g. [CH98, Prop. 4.3.9],[12] we have that there exists a map $T : D(\mathcal{A}) \to (0, \infty]$, for which $u \in C([0, T(u_0)), D(\mathcal{A})) \cap C^1([0, T(u_0)), \mathcal{H})$ and u solves the initial-value problem

$$(4.79a) \qquad \frac{\mathrm{d}u}{\mathrm{d}t} = \mathcal{A}u + \mathcal{F}(u),$$

$$(4.79b) \qquad u(0) = u_0.$$

By taking the $\mathcal{H}$-inner product on both sides of (4.79a) with the solution $u \in \mathcal{D}(\mathcal{A})$, and using the energy inequality (4.76) and the stability property $\langle \mathcal{A}\Psi, \Psi \rangle_{\mathcal{H}} \leq \omega \|\Psi\|_{\mathcal{H}}^2$ from Lemma 4.2, we obtain

$$(4.80) \qquad \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u\|_{\mathcal{H}}^2 = \langle \mathcal{A}u, u \rangle_{\mathcal{H}} + \langle \mathcal{F}(u), u \rangle_{\mathcal{H}} \leq \omega\|u\|_{\mathcal{H}}^2 + \gamma_1\|u\|_{\mathcal{H}}^2 + \gamma_2,$$

where the positive constant ω is given by (4.49).

It follows then from Gronwall's inequality that

$$(4.81) \qquad \|u(t; u_0)\|_{\mathcal{H}}^2 \leq e^{\kappa t}\|u_0\|_{\mathcal{H}}^2 + \frac{2\gamma_2}{\kappa}(e^{\kappa t} - 1), \quad t \in [0, T(u_0)),\ u_0 \in D(\mathcal{A}),$$

where $\kappa = \omega + \gamma_1$.

Similarly, by noting that

$$\langle \Pi_N \mathcal{F}(\Psi), \Psi \rangle_{\mathcal{H}} = \langle \mathcal{F}(\Psi), \Psi \rangle_{\mathcal{H}}, \qquad \forall\, \Psi \in \mathcal{H}_N,$$

we have

$$(4.82) \qquad \|u_N(t; \Pi_N u_0)\|_{\mathcal{H}}^2 \leq e^{\kappa t}\|u_0\|_{\mathcal{H}}^2 + \frac{2\gamma_2}{\kappa}(e^{\kappa t} - 1), \quad t \in [0, T(u_0)),\ u_0 \in \mathcal{H}.$$

We have thus shown that for any initial data $u_0 \in D(\mathcal{A})$, the strong solutions $u(t; u_0)$ and $u_N(t; \Pi_N u_0)$ do not blow up in any finite time. The convergence result (4.77) can then be deduced as in the proof of Corollary 4.2.

$\square$

---

[11] By strong solutions of (4.23), we mean a solution in $C([0, T], D(\mathcal{A})) \cap C^1([0, T], \mathcal{H})$ of (4.23).

[12] The regularity result [CH98, Prop. 4.3.9] is stated for the case of contraction semigroups. However, the proof can be adapted to the case of C_0-semigroups for which $\|T(t)\| \leq e^{\omega t}$ (i.e. with $M = 1$) such as encountered here when L_D is as specified in (4.47).

Remark 4.3. It is worth mentioning that the conclusions of Corollaries 4.1, 4.2 and 4.3 still hold when the underlying system of DDEs (4.65) is perturbed by a suitable time-dependent forcing, $\mathbf{g}(t)$. For instance, it suffices to assume that $\mathbf{g}(t) \in L^2_{\text{loc}}([0, \infty); \mathbb{R}^d)$ to still get the convergence results.   □

4.3. **Examples.** In this section we provide some class of nonlinear scalar DDEs of the form (2.4) that fit with the assumptions of Corollary 4.3. In that respect we restrict our attention to the case of $\mathcal{A}$ such as defined in (2.8) for $d = 1$. We discuss below some classes of nonlinearities that verify the local Lipschitz condition (4.68) and the energy inequality (4.76). Extension to systems can be easily built out of these examples and are thus left to the reader.

4.3.1. *Delay equations with a global Lipschitz nonlinearity.* Let $\mathcal{F}$ be given such as in (2.10), and for which F is assumed to be of the form

(4.83)
$$F\Big(\Psi^S, \int_{-\tau}^0 \Psi^D(\theta)\mathrm{d}\theta\Big) = g\Big(\int_{-\tau}^0 \Psi^D(\theta)\mathrm{d}\theta\Big) + h(\Psi^S),$$

with g and h to be global Lipschitz maps from $\mathbb{R}$ to $\mathbb{R}$, of constants L_1 and L_2, respectively. Then, we have

$$\langle \mathcal{F}(\Psi), \Psi \rangle_{\mathcal{H}} = g\Big(\int_{-\tau}^0 \Psi^D(\theta)\mathrm{d}\theta\Big)\Psi^S + h(\Psi^S)\Psi^S,$$

which gives

$$\langle \mathcal{F}(\Psi), \Psi \rangle_{\mathcal{H}} \leq L_1 \Big| \int_{-\tau}^0 \Psi^D(\theta)\mathrm{d}\theta \Big| |\Psi^S| + (|g(0)| + |h(0)|)|\Psi^S| + L_2 |\Psi^S|^2$$
$$\leq \gamma_1 \|\Psi\|^2_{\mathcal{H}} + \gamma_2,$$

with $\gamma_1 > 0$ and $\gamma_2 \geq 0$, and thus F satisfies the energy inequality (4.76). The local Lipschitz condition (4.68) for $\mathcal{F}$ is trivially satisfied under the assumptions on F.

Note that such nonlinear equations arise in many applications where a delayed monotone feedback mechanism is naturally involved in the description of the system's evolution; see [GZT08, Kri08, MPS96]. It is also interesting to mention that such seemingly simple scalar DDEs with global Lipschitz nonlinearity can also support chaotic dynamics as illustrated in Section 6.1 below.

4.3.2. *Delay equations with locally Lipschitz nonlinearity.* We relax now the global Lipschitz requirement. In that respect, we consider

(4.84)
$$F\Big(\Psi^S, \int_{-\tau}^0 \Psi^D(\theta)\mathrm{d}\theta\Big) = -g_1\Big(\int_{-\tau}^0 \Psi^D(\theta)\mathrm{d}\theta\Big)g_2\big(\Psi^S\big),$$

and assume that

    (i) $g_1 : \mathbb{R} \to \mathbb{R}^+$ is locally Lipschitz;
    (ii) $g_2 : \mathbb{R} \to \mathbb{R}$ is locally Lipschitz and verifies the condition

(4.85)
$$g_2(x)x \geq 0, \qquad x \in \mathbb{R}.$$

These assumptions allow us to consider a broad class of nonlinear effects that are not necessarily bounded or polynomial.

We check below that the abstract nonlinear map $\mathcal{F}$ defined in (2.10) with F given by (4.84) and that satisfy (i) and (ii), satisfies also the conditions of Corollary 4.3. We first check the local Lipschitz condition.

Trivially, let us first remark that

(4.86)
$$\|\mathcal{F}(\Psi_1) - \mathcal{F}(\Psi_2)\|_{\mathcal{H}} = \Big| F\Big(\Psi_1^S, \int_{-\tau}^0 \Psi_1^D(\theta)\mathrm{d}\theta\Big) - F\Big(\Psi_2^S, \int_{-\tau}^0 \Psi_2^D(\theta)\mathrm{d}\theta\Big) \Big|.$$

Let us introduce the notations $\alpha_i := \Psi_i^S$ and $\beta_i := \int_{-\tau}^0 \Psi_i^D(\theta)\mathrm{d}\theta$, $i = 1, 2$. Let Ψ_i be chosen such that for $i = 1, 2$, $\|\Psi_i\|_{\mathcal{H}} \leq R$ for some $R > 0$. It follows that

$$\text{(4.87)} \qquad |\alpha_i| \leq R, \quad |\beta_i| = \left| \int_{-\tau}^0 \Psi_i^D(\theta)\mathrm{d}\theta \right| \leq \sqrt{\tau} \|\Psi_i^D\|_{L^2},$$

and thus by definition of the $\mathcal{H}$-inner product (2.2)

$$\text{(4.88)} \qquad |\beta_i| \leq \tau \|\Psi_i\|_{\mathcal{H}} \leq \tau R,$$

which leads to

$$\text{(4.89)} \qquad \begin{aligned} &|F(\alpha_1, \beta_1) - F(\alpha_2, \beta_2)| \\ &\leq |g_1(\beta_1)g_2(\alpha_1) - g_1(\beta_2)g_2(\alpha_2)| \\ &\leq \left| g_1(\beta_1)\big(g_2(\alpha_1) - g_2(\alpha_2)\big) \right| + \left| \big(g_1(\beta_1) - g_1(\beta_2)\big)g_2(\alpha_2) \right| \\ &\leq \mathrm{L}_2(R)|g_1(\beta_1)||\alpha_1 - \alpha_2| + \mathrm{L}_1(\tau R)|g_2(\alpha_2)||\beta_1 - \beta_2|, \end{aligned}$$

where $\mathrm{L}_1(r)$ (resp. $\mathrm{L}_2(r)$) denotes the local Lipschitz constant associated with $g_1(x)$ (resp. $g_2(x)$) for $|x| < r$.

On the other hand,

$$\begin{aligned} |g_1(\beta_1)| &\leq |g_1(\beta_1) - g_1(0)| + |g_1(0)| \\ &\leq \mathrm{L}_1(\tau R)|\beta_1| + |g_1(0)| \\ &\leq \tau R \mathrm{L}_1(\tau R) + |g_1(0)|, \end{aligned}$$

and

$$|g_2(\alpha_2)| \leq R\mathrm{L}_2(R) + |g_2(0)|.$$

Note also that

$$|\alpha_1 - \alpha_2| \leq \|\Psi_1 - \Psi_2\|_{\mathcal{H}},$$

and that

$$|\beta_1 - \beta_2| \leq \sqrt{\tau}\|\Psi_1^D - \Psi_2^D\|_{L^2} \leq \tau\|\Psi_1 - \Psi_2\|_{\mathcal{H}}.$$

We obtain then from (4.89) that

$$\text{(4.90)} \qquad |F(\alpha_1, \beta_1) - F(\alpha_2, \beta_2)| \leq L(R)\|\Psi_1 - \Psi_2\|_{\mathcal{H}},$$

where

$$L(R) := \mathrm{L}_2(R)\Big(\sqrt{\tau}R\mathrm{L}_1(\tau R) + |g_1(0)|\Big) + \tau\mathrm{L}_1(\tau R)\Big(R\mathrm{L}_2(R) + |g_2(0)|\Big),$$

which gives the local Lipschitz property of $\mathcal{F}$ as a map from $\mathcal{H}$ to $\mathcal{H}$.

The energy inequality (4.76) is here readily satisfied since

$$\text{(4.91)} \qquad \langle \mathcal{F}(\Psi), \Psi \rangle_{\mathcal{H}} = F\Big(\Psi^S, \int_{-\tau}^0 \Psi^D(\theta)\mathrm{d}\theta\Big)\Psi^S \leq 0,$$

because of (4.85).

Remark 4.4.
   (i) Note that famous delayed models from population dynamics are covered by Corollary 4.3, although not satisfying (4.84). For instance, delayed logistic equations of the form

$$\text{(4.92)} \qquad \frac{\mathrm{d}x}{\mathrm{d}t} = rx(t)\Big(1 - K^{-1}\int_{t-\tau}^t \omega(s)f(x(s))\mathrm{d}s\Big), \ r, K, \tau > 0,$$

   with $\omega \in L^\infty(\mathbb{R}, \mathbb{R}^+)$ and $f \in L^1(\mathbb{R}, \mathbb{R}^+)$ satisfying the inequality

$$|f(x) - f(y)| \leq \gamma|x - y|,$$

for some $\gamma > 0$ and for almost every $x, y \in \mathbb{R}$, are still covered by Corollary 4.3.

(ii) Note also that many other nonlinear effects could have been considered in (2.4) for which the convergence result of Corollary 4.3 would hold. For instance we could have considered

$$(4.93) \qquad F\Big(\Psi^S, \int_{-\tau}^0 \Psi^D(\theta)\mathrm{d}\theta\Big) = \sum_{j=1}^{2p-1} b_j\Big(\int_{-\tau}^0 \Psi^D(\theta)\mathrm{d}\theta\Big)(\Psi^S)^j, p \geq 2,$$

where each b_j is a local Lipschitz function in $L^\infty(\mathbb{R})$ and $b_{2p-1}(\cdot) \leq \beta < 0$ for some constant β.

Under these assumptions, the nonlinearity (4.93) satisfies the energy inequality (4.76) with $\gamma_1 = 0$ and with γ_2 sufficiently large, depending on the L^∞ norm of the functions b_j's. This is because the term with the largest power $(\Psi^S)^{2p-1}$ is strictly negative by assumption on b_{2p-1}, and the other terms with a lower degree can be controlled by using Young's inequality. $\qquad \square$

5. Galerkin approximation: Analytic formulas for scalar DDEs

This section is devoted to the derivation of explicit expressions of the Galerkin approximation (4.28) associated with nonlinear DDEs. For simplicity, we focus on the case of a scalar DDE taking the form given by (2.4). The more general case of nonlinear systems of DDEs is dealt with in Appendix C; see also Appendix C.2 for the case where the linear part of (2.4) involves a distributed-delay term as in (4.65).

As a preparation for the forthcoming analytic derivations, we need to express the derivative of the Koornwinder polynomials in terms of the polynomials themselves. This is the content of the following proposition.

Proposition 5.1. The Koornwinder polynomial K_n of degree $n \in \mathbb{N}$ defined in (3.3) satisfies the differential relation

$$(5.1) \qquad \frac{\mathrm{d}K_n}{\mathrm{d}s}(s) = \sum_{k=0}^{n-1} a_{n,k} K_k(s), \quad s \in (-1, 1),$$

where the coefficients $a_n := (a_{n,0}, \cdots, a_{n,n-1})^{\mathrm{tr}}$, satisfy the upper triangular system of linear equations

$$(5.2) \qquad \mathbf{T}a_n = b_n,$$

with $\mathbf{T} := (\mathbf{T}_{i,j})_{n \times n}$ and $b_n := (b_{n,0}, \cdots, b_{n,n-1})^{\mathrm{tr}}$ given by

$$(5.3) \qquad
\begin{aligned}
\mathbf{T}_{i,j} &= \begin{cases} 0, & \text{if } j < i, \\ i^2 + 1, & \text{if } j = i, \\ -(2i+1), & \text{if } j > i, \end{cases} \quad \text{where} \quad 0 \leq i, j \leq n-1, \\[2mm]
b_{n,i} &= \begin{cases} \frac{1}{2}(2i+1)(n+i+1)(n-i), & \text{if } n+i \text{ is even,} \\[2mm] (n^2+n)(2i+1) - \frac{i}{2}(n+i)(n-i+1) \\ \qquad - \frac{i}{2}(i+1)(n-i-1)(n+i+2), & \text{if } n+i \text{ is odd.} \end{cases}
\end{aligned}$$

For the rescaled version K_n^τ defined by (3.18), it holds that

$$(5.4) \qquad \frac{\mathrm{d}K_n^\tau}{\mathrm{d}\theta}(\theta) = \frac{2}{\tau} \sum_{k=0}^{n-1} a_{n,k} K_k^\tau(\theta), \qquad n \in \mathbb{N}, \quad \theta \in (-\tau, 0).$$

Proof. See Appendix B.

$\qquad \square$

Let us now rewrite the unknown u_N in the Galerkin system (4.28) in terms of the first N rescaled Koornwinder polynomials, i.e.:

$$(5.5) \qquad u_N(t) = \sum_{n=0}^{N-1} y_n(t) \mathcal{K}_n^\tau, \qquad t \ge 0,$$

where

$$(5.6) \qquad y_n(t) = \frac{\langle u_N(t), \mathcal{K}_n^\tau \rangle_{\mathcal{H}}}{\|\mathcal{K}_n^\tau\|_{\mathcal{H}}^2}.$$

We then replace u_N in Eq. (4.28) by the expansion given in (5.5), and take the $\mathcal{H}$-inner product on both sides with $\mathcal{K}_j^\tau$ for each $j \in \{0, \cdots, N-1\}$ to obtain:

$$(5.7) \qquad \|\mathcal{K}_j^\tau\|_{\mathcal{H}}^2 \frac{\mathrm{d}y_j}{\mathrm{d}t} = \sum_{n=0}^{N-1} y_n(t) \left\langle \mathcal{A}_N \mathcal{K}_n^\tau, \mathcal{K}_j^\tau \right\rangle_{\mathcal{H}} + \left\langle \Pi_N \mathcal{F}\left(\sum_{n=0}^{N-1} y_n(t) \mathcal{K}_n^\tau \right), \mathcal{K}_j^\tau \right\rangle_{\mathcal{H}}.$$

Recall that the linear operator $\mathcal{A}$ here is defined by (2.8). Then, for each $n \in \{0, \cdots, N-1\}$, it holds that

$$(5.8) \qquad \begin{aligned} \mathcal{A}_N \mathcal{K}_n^\tau = \Pi_N \mathcal{A} \mathcal{K}_n^\tau = \sum_{l=0}^{N-1} \Big(\frac{1}{\tau} \Big\langle \frac{\mathrm{d}^+}{\mathrm{d}\theta} K_n^\tau, K_l^\tau \Big\rangle_{L^2} \\ + \Big(a K_n^\tau(0) + b K_n^\tau(-\tau) + c \int_{-\tau}^0 K_n^\tau(\theta) \mathrm{d}\theta \Big) \Big) \frac{\mathcal{K}_l^\tau}{\|\mathcal{K}_l^\tau\|_{\mathcal{H}}^2}. \end{aligned}$$

We obtain then that

$$(5.9) \qquad \begin{aligned} \left\langle \mathcal{A}_N \mathcal{K}_n^\tau, \mathcal{K}_j^\tau \right\rangle_{\mathcal{H}} = \frac{1}{\tau} \Big\langle \frac{\mathrm{d}^+}{\mathrm{d}\theta} K_n^\tau, K_j^\tau \Big\rangle_{L^2} \\ + \Big(a K_n^\tau(0) + b K_n^\tau(-\tau) + c \int_{-\tau}^0 K_n^\tau(\theta) \mathrm{d}\theta \Big) K_j^\tau(0). \end{aligned}$$

It follows from the expression of $\frac{\mathrm{d}K_n^\tau}{\mathrm{d}\theta}$ in (5.4) that

$$(5.10) \qquad \begin{aligned} \frac{1}{\tau} \int_{-\tau}^0 \frac{\mathrm{d}^+ K_n^\tau}{\mathrm{d}\theta}(\theta) K_j^\tau(\theta) \mathrm{d}\theta &= \frac{2}{\tau} \sum_{k=0}^{n-1} a_{n,k} \left(\frac{1}{\tau} \int_{-\tau}^0 K_k^\tau(\theta) K_j^\tau(\theta) \mathrm{d}\theta \right) \\ &= \frac{2}{\tau} \sum_{k=0}^{n-1} a_{n,k} \left(\langle \mathcal{K}_k^\tau, \mathcal{K}_j^\tau \rangle_{\mathcal{H}} - K_k^\tau(0) K_j^\tau(0) \right) \\ &= \frac{2}{\tau} \sum_{k=0}^{n-1} a_{n,k} \left(\delta_{j,k} \|\mathcal{K}_j^\tau\|_{\mathcal{H}}^2 - 1 \right), \end{aligned}$$

where $\delta_{j,k}$ denotes the Kronecker delta, and the last equality above follows from the orthogonal property of the Koornwinder polynomials as well as the normalization property $K_n^\tau(0) = 1$; cf. (3.20).

Note that $K_0^\tau \equiv 1$, which follows from the definition of the Koornwinder polynomials K_n given by (3.3) and the fact that the first Legendre polynomial L_0 is identically 1. We get then $\|\mathcal{K}_0^\tau\|_{\mathcal{H}}^2 = 2$. Note also that

$$(5.11) \qquad \langle \mathcal{K}_n^\tau, \mathcal{K}_0^\tau \rangle_{\mathcal{H}} = \frac{1}{\tau} \int_{-\tau}^0 K_n^\tau(\theta) \mathrm{d}\theta + 1 = \delta_{n,0} \|\mathcal{K}_0^\tau\|_{\mathcal{H}}^2 = 2\delta_{n,0},$$

leading thus to

$$(5.12) \qquad \int_{-\tau}^{0} K_n^\tau(\theta) \mathrm{d}\theta = \tau(2\delta_{n,0} - 1), \quad n \in \mathbb{N}.$$

By using (5.12), the normalization property $K_j^\tau(0) = 1$ and the identity $K_j^\tau(-\tau) = K_j(-1)$ (valid for any $j \geq 0$), we obtain

$$(5.13) \qquad \Big(aK_n^\tau(0) + bK_n^\tau(-\tau) + c\int_{-\tau}^{0} K_n^\tau(\theta)\mathrm{d}\theta\Big) K_j^\tau(0) = a + bK_n(-1) + c\tau(2\delta_{n,0} - 1).$$

Now, by using (5.10) and (5.13) in (5.9), we obtain

$$(5.14) \qquad \sum_{n=0}^{N-1} y_n(t) \left\langle \mathcal{A}_N \mathcal{K}_n^\tau, \mathcal{K}_j^\tau \right\rangle_{\mathcal{H}} = \sum_{n=0}^{N-1} \Big(a + bK_n(-1) + c\tau(2\delta_{n,0} - 1) \\ + \frac{2}{\tau} \sum_{k=0}^{n-1} a_{n,k} \left(\delta_{j,k} \|\mathcal{K}_j\|_{\mathcal{H}}^2 - 1 \right) \Big) y_n(t).$$

For the nonlinear part, since $\langle \Pi_N \Phi, \mathcal{K}_n^\tau \rangle_{\mathcal{H}} = \langle \Phi, \mathcal{K}_n^\tau \rangle_{\mathcal{H}}$ for all $\Phi \in \mathcal{H}$ and all $n \in \{0, \cdots, N-1\}$, together with the definition of $\mathcal{F}$ given in (2.10), we obtain

$$(5.15) \qquad \left\langle \Pi_N \mathcal{F}\Big(\sum_{n=0}^{N-1} y_n(t)\mathcal{K}_n^\tau \Big), \mathcal{K}_j^\tau \right\rangle_{\mathcal{H}} = F\left(\sum_{n=0}^{N-1} y_n(t), \int_{-\tau}^{0} \sum_{n=0}^{N-1} y_n(t) K_n^\tau(\theta)\mathrm{d}\theta \right).$$

From (5.12), it also follows that

$$(5.16) \qquad \int_{-\tau}^{0} \sum_{n=0}^{N-1} y_n(t) K_n^\tau(\theta)\mathrm{d}\theta = \tau y_0(t) - \tau \sum_{n=1}^{N-1} y_n(t).$$

By using (5.16) in (5.15), we get

$$(5.17) \qquad \left\langle \Pi_N \mathcal{F}\Big(\sum_{n=0}^{N-1} y_n(t)\mathcal{K}_n^\tau \Big), \mathcal{K}_j^\tau \right\rangle_{\mathcal{H}} = F\left(\sum_{n=0}^{N-1} y_n(t), \tau y_0(t) - \tau \sum_{n=1}^{N-1} y_n(t) \right).$$

Now, by using (5.14) and (5.17) and recalling that $\|\mathcal{K}_n^\tau\|_{\mathcal{H}} = \|\mathcal{K}_n\|_{\mathcal{E}}$, we obtain from Eq. (5.7) the following explicit form of the N-dimensional Galerkin system (4.28):

$$(5.18) \qquad \begin{aligned} \frac{\mathrm{d}y_j}{\mathrm{d}t} = {} & \frac{1}{\|\mathcal{K}_j\|_{\mathcal{E}}^2} \sum_{n=0}^{N-1} \Big(a + bK_n(-1) + c\tau(2\delta_{n,0} - 1) \\ & + \frac{2}{\tau} \sum_{k=0}^{n-1} a_{n,k} \left(\delta_{j,k} \|\mathcal{K}_j\|_{\mathcal{E}}^2 - 1 \right) \Big) y_n(t) \\ & + \frac{1}{\|\mathcal{K}_j\|_{\mathcal{E}}^2} F\left(\sum_{n=0}^{N-1} y_n(t), \tau y_0(t) - \tau \sum_{n=1}^{N-1} y_n(t) \right), \quad 0 \leq j \leq N-1. \end{aligned}$$

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

$$(5.19) \qquad \boxed{\frac{\mathrm{d}y}{\mathrm{d}t} = Ay + G(y),}$$

where Ay denotes the linear part of Eq. (5.18), and $G(y)$ the nonlinear part. Namely, A is the $N \times N$ matrix whose elements are given by

(5.20)
$$
\begin{aligned}
(A)_{j,n} = \frac{1}{\|\mathcal{K}_j\|_{\mathcal{E}}^2} \Big(a + b K_n(-1) + c\tau(2\delta_{n,0} - 1) \\
+ \frac{2}{\tau} \sum_{k=0}^{n-1} a_{n,k} \left(\delta_{j,k}\|\mathcal{K}_j\|_{\mathcal{E}}^2 - 1 \right) \Big),
\end{aligned}
$$

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

(5.21)
$$
G_j(y) = \frac{1}{\|\mathcal{K}_j\|_{\mathcal{E}}^2} F\left(\sum_{n=0}^{N-1} y_n(t), \tau y_0(t) - \tau \sum_{n=1}^{N-1} y_n(t) \right), \ 0 \le j \le N-1.
$$

Remark 5.1. When a time-dependent forcing $g(t)$ is added to the RHS of the DDE (2.4), the only change in the corresponding Galerkin system is to add a term $\|\mathcal{K}_j\|_{\mathcal{E}}^{-2} g(t)$ to each G_j in (5.21). Recall also from Remark 4.3 that it is sufficient to require that $g \in L^2_{\text{loc}}(\mathbb{R}; \mathbb{R})$ in order to ensure the convergence result of the Galerkin system over any finite interval. $\square$

6. Approximation of chaotic dynamics: Numerical results

In this section, we report on the performance of our Galerkin schemein approximating quasi-periodic and chaotic dynamics. In the case of the latter, it is well known that any finite-time uniform convergence result — such as the one given by (4.77) and obtained in Corollary 4.3 — becomes less useful, due to sensitivity to initial data.

In this case, when individual solutions diverge exponentially, it is natural to consider instead the approximation of the strange attractor or of the statistics of the dynamics. More generally, the approximation of meaningful invariant probability measures supported by the strange attractor is of primary interest in chaotic dynamics. Nonlinear DDEs, like those considered here, are known to support such probability measures once a global attractor is known to exist [CGH12].

If, in addition to the finite-time uniform convergence (4.77), a uniform dissipativity assumption is satisfied, then any sequence of invariant measures associated with the Galerkin approximation converges weakly to an invariant measure of the full system; see [Wan09, Thm. 2.2]. The uniform dissipativity assumption referred to in [Wan09] is satisfied if one can establish, uniformly in N, the existence of an absorbing ball for the Galerkin reduced systems in another separable Hilbert space $\mathcal{V}$, which is compactly imbedded in $\mathcal{H}$, in the case of strongly dissipative systems.

We show in Sections 6.1 and 6.2 below that, for two simple nonlinear scalar DDEs that — even in instances in which the uniform dissipativity assumption of [Wan09, Thm. 2.2] is not guaranteed — our Galerkin scheme is still able to approximate significant statistical properties of the chaotic dynamics, or the strange attractor itself. Finally, we illustrate in Section 6.3, for a highly idealized ENSO model from climate dynamics that our approach also works in the case of periodically forced DDEs with multiple delays.

6.1. **"Nearly-Brownian" chaotic dynamics.** We consider here a modified version of the DDE analyzed in [Spr07] with a sinusoidal nonlinearity. The modification consists of replacing the discrete delay by distributed delays. As we will see, this modification does not affect the main dynamical properties identified in [Spr07] in the case of a discrete delay, once the proper parameter values are chosen.

More precisely, we consider the following DDE

$$(6.1) \qquad \dot{x} = a \sin \left(\int_{t-\tau}^{t} x(s) \mathrm{d}s \right),$$

with $a = 0.5$ and $\tau = 5.5$. This example fits within the general class discussed in Section 4.3.1, to which the rigorous convergence results described in Sections 4.2 in an abstract setting, and in Section 5 more concretely, do apply.

As pointed out in the introduction of this section, such finite-time uniform convergence results are essential in general, but they are not the ones we are necessarily looking for in approximating chaotic dynamics. We rely, therefore, on careful numerical simulations to explore the performance of our Koornwinder-polynomial–based Galerkin systems to approximate the statistical features of the chaotic dynamics in these examples.

A sample trajectory of Eq. (6.1) is shown in black in Fig. 2. While the governing DDE is perfectly deterministic, the visual resemblance of this trajectory to a sample path of Brownian motion is obvious. A trajectory with the same constant initial history over the interval $[-\tau, 0)$, but obtained by solving a 10D-approximation by our Galerkin scheme of the DDE, is shown as the red curve in the figure. It is clear that individual trajectories do have the same overall behavior, which is nearly Brownian, but the pointwise approximation of the exact trajectory by the approximate one is not good.

To study instead the statistics of the solutions, we have estimated — for a collection of $n = 10^4$ initial histories of constant value drawn uniformly in $[-1, 1]$ — the empirical probability density function (PDF) of the corresponding $x(t)$-values at a given time t. Figure 3 reports on the numerical results at $t = 2680$ when $x(t)$ is simulated from Eq. (6.1) by forward Euler integration (black curve) with a time step of $\Delta t = \tau/2^{10}$, and when $x(t)$ is approximated by a 10D-Galerkin approximation (red curve) obtained from the analytic formulas (5.19)–(5.21) applied to Eq. (6.1).

The Galerkin system of ODEs is integrated using a semi-implicit Euler method that still uses $\Delta t = \tau/2^{10}$, but in which the linear part Ay is treated implicitly, while the nonlinear term $G(y)$ is treated explicitly. The approximate solution $x_N(t)$, provided by an N-dimensional Galerkin system of the form (5.19), is obtained by using the expansion (5.5) into Koornwinder polynomials. More precisely, $x_N(t)$ is obtained as the state part of u_N given by (5.5) which, thanks to the normalization property $K_n^\tau(0) = 1$ given in (3.20), reduces to

$$(6.2) \qquad x_N(t) = \sum_{j=1}^{N} y_j(t),$$

where $y := (y_0, \cdots, y_{N-1})$ is the vector solution of (5.19).

Also shown in blue in panels (a, c) of Fig. 3, is a Gaussian distribution with the same mean and standard deviation, both of which are estimated from the $x(t)$-values at $t = 2680$ of the simulated DDE solutions. In both cases, the empirical distributions, as obtained from the simulated DDE solution or as obtained by using a Galerkin approximation with $N = 10$, closely follow a Gaussian law.

For both $x(t)$, as simulated by numerically solving Eq. (6.1), and $x_N(t)$, as obtained by using a Galerkin approximation with $N = 10$, the standard deviations of the corresponding collection of trajectories are shown versus time in panels (b) and (d) of Fig. 3. The best linear regression fit of $\log(\sigma)$ versus $\log(t)$ gives $\sigma = 0.195 t^{0.479}$, in the case of Eq. (6.1), and $\sigma = 0.190 t^{0.480}$, in the case of the 10D-Galerkin approximation.

In both cases, the slopes of the fitted lines indicate that the deterministic dynamics of Eq. (6.1) mimics that of Brownian motion, for which the slope would be 0.5, with a diffusion coefficient $D = \sigma^2/t$. In particular, the solutions do not stay within any bounded subset; the dynamics thus violates any dissipation criterion, while still possessing an attractor with strange behavior (not shown).
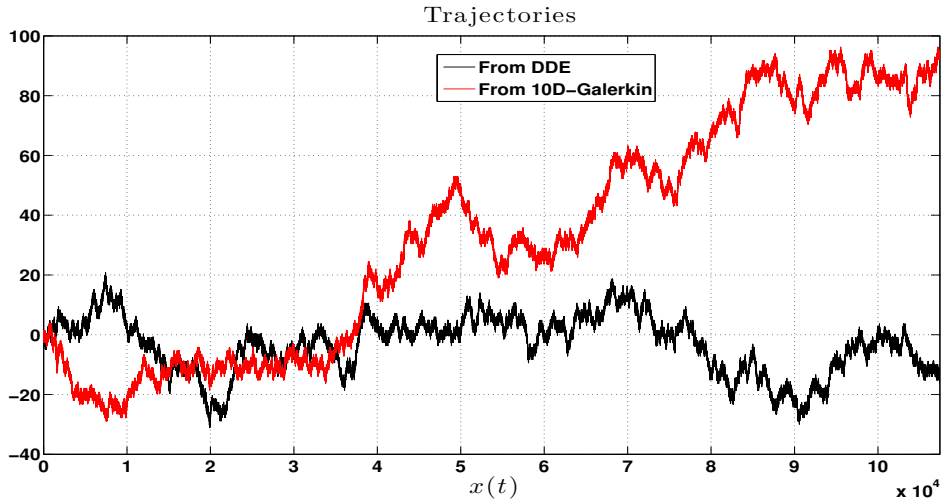
FIGURE 2. Trajectory $x(t)$ simulated by the DDE (6.1) (black curve), and its 10D-Galerkin approximation $x_N(t)$, with $N = 10$ (red curve), obtained by the method described in Section 5.

To the best of our knowledge, the 10D-Galerkin approximation computed here thus provides the first example of a chaotic system of ODEs whose solutions exhibit a statistical behavior close to that of Brownian motion.

6.2. **Bimodal chaotic dynamics with low-frequency variability.** The model studied in this subsection,

$$(6.3) \qquad \dot{x} = ax(t - \tau) - bx(t - \tau)^3,$$

is also based on [Spr07]; the parameter values used here are $a = 0.5$, $b = 20$ and $\tau = 3.35$. Remark 6.1 at the end of this subsection discusses similarities between the dynamics of the model above and important aspects of ENSO dynamics.

In contrast to the DDE considered in the previous subsection, Eq. (6.3) does not fit directly within the general framework of Section 4.2, for which rigorous convergence results are available. The discrete lag present in the cubic term leads to complications for a rigorous analysis. Replacing this lag effect by a distributed one, as in Eq. (6.1), would place the DDE of Eq. (6.3) into the class considered in Section 4.3.2, for which finite-time uniform convergence results are guaranteed. But even then, we cannot be assured *a priori* of an effective approximation of statistical features of the dynamics, as discussed above.

The purpose of this subsection is to show that, even in such a borderline case with respect to the theory presented in this article, statistical and even topological features can still be remarkably well approximated by the Galerkin systems of Section 5, when appropriately modified to handle the case of discrete delay in the nonlinear terms.[13]

---

[13]This modification consists just of noting that, by replacing the nonlinear term with distributed delays in (2.4) by $F(x(t), x(t - \tau))$, the identity (5.15) becomes

$$(6.4) \qquad \Big\langle \Pi_N \mathcal{F}\Big(\sum_{n=0}^{N-1} y_n(t)\mathcal{K}_n^\tau\Big), \mathcal{K}_j^\tau \Big\rangle_{\mathcal{H}} = F\Big(\sum_{n=0}^{N-1} y_n(t), \sum_{n=0}^{N-1} y_n(t)K_n(-1)\Big);$$

the latter, in turn, leads to the corresponding change in the formula (5.21) for the nonlinear vector field in Eq. (5.19).
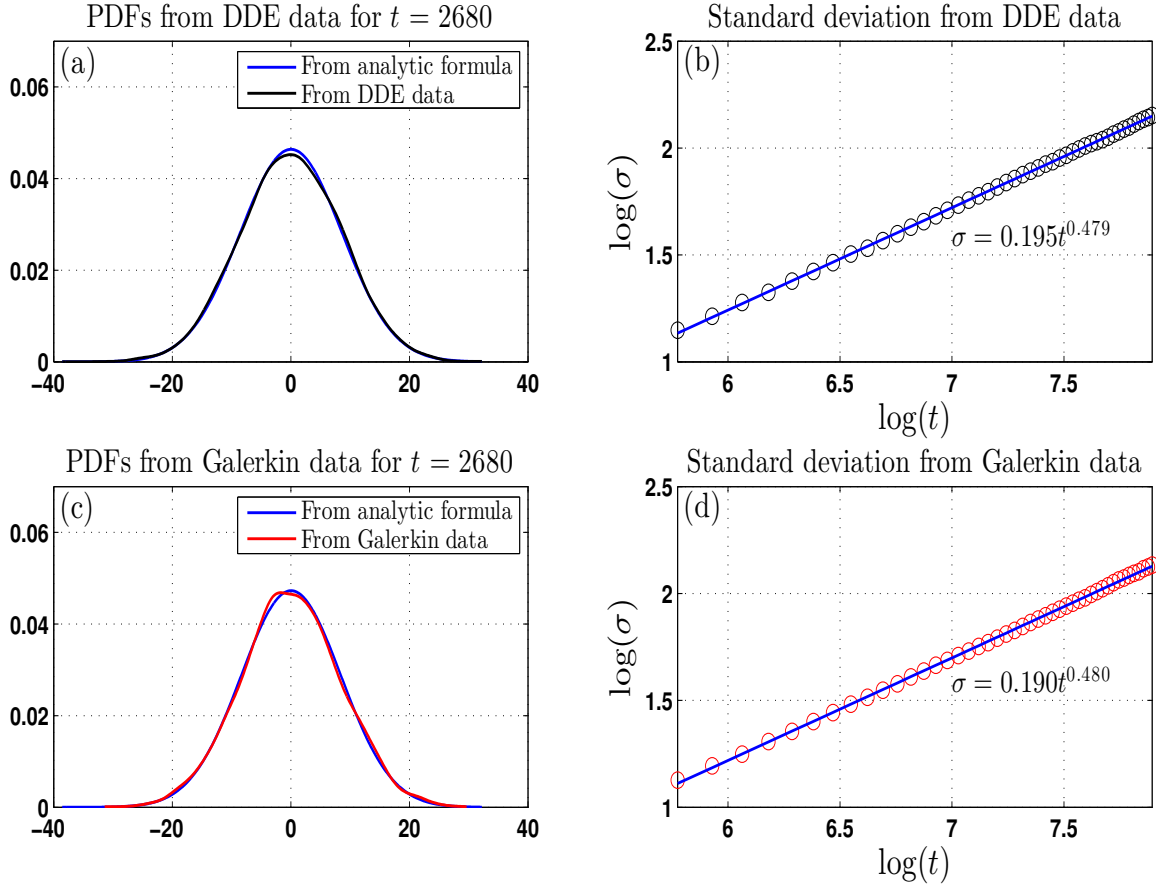
FIGURE 3. (a, c): Probability density functions (PDFs); and (b, d): standard deviations σ, as estimated from the simulations of the DDE (6.1) (black curve and black circles, in the left panels and the right panels, respectively), and its 10D-Galerkin approximation (red curve and red circles) by the method described in Section 5, respectively. In each of the panels (a) and (b), the results are compared with a Gaussian distribution estimated by standard analytic formulas (blue curves), and in each of the panels (c) and (d), the results are compared with the best linear regression fit (blue curve) of $\log(\sigma)$ versus $\log(t)$, providing the corresponding slope and the exponent reported therein.

As shown in the left panel of Figure 4 in natural delayed coordinates, the strange attractor associated with Eq. (6.3) (black) has a nearly symmetric topological shape and is constituted by two fairly high-density "islands" connected by a foliation of heteroclinic-like orbits. These properties of the attractor are very well captured by a 6D-Galerkin approximation (right panel, red).

These two high-density islands, along with the lower-density areas of heteroclinic-like connections between the two, give rise to a bimodal chaotic behavior, which gives rise, in turn, to an interesting time variability. Figure 5 plots the results of a standard numerical estimation of the spectral density — also known in the engineering literature as the power spectrum — associated with $x(t)$, as directly simulated using Eq. (6.3) (black curve) and as approximated by a 6D-Galerkin approximation (red curve). In both cases, the power spectra are estimated from the corresponding autocorrelation functions [ER85, GAD+02].

The numerical results show that the spectrum of $x(t)$ contains two broadband peaks at low frequencies that stand out above an exponentially decaying background. As shown in Fig. 5, these
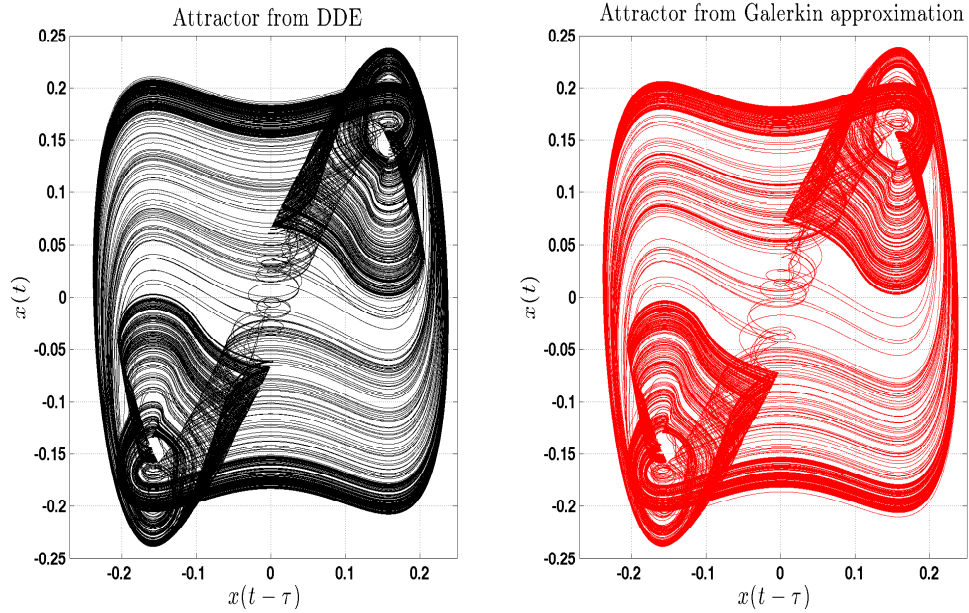
FIGURE 4. The attractor associated with Eq. (6.3) (left panel) and its approximation obtained from a 6D-Galerkin approximation (right panel).
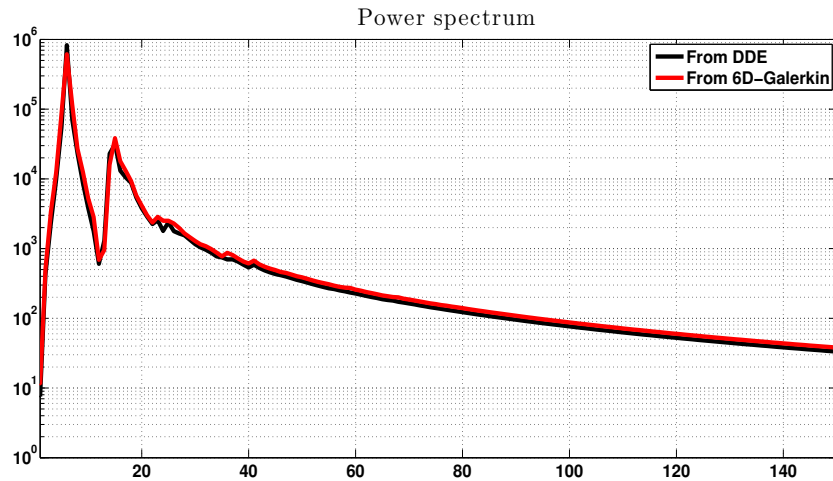

FIGURE 5. Power spectral density of $x(t)$ as simulated from Eq. (6.3), and from a 6D-Galerkin approximation derived from the analytic formulas described in Section 5 and applied to Eq. (6.3).

two peaks, as well as the exponential background, are strikingly well approximated by a 6D-Galerkin approximation, in both amplitude and frequency, as well as in the rate of decay for high frequencies.

The approximation by a truly low-dimensional, 6D-Galerkin model of these key features of the power spectrum of the solutions to the DDE (6.3) has deep dynamical implications in terms of the Ruelle-Pollicott resonances and mixing properties of the dynamics on the attractor [CNK+14]. These implications are beyond the scope of this article but we intend to discuss them elsewhere.

Remark 6.1. Equation. (6.3) can actually be seen as a highly idealized ENSO model with memory effects; see [BH89, CMZ90, Dij05, GCS15, GT00, GZT08, KS14, MCZ91, TSCJ94, ZG10, vR13] and references therein. Indeed, for the given parameter values, the solution of Eq. (6.3) admits two metastable states, as can be seen from the two islands in the attractor given in Figure 4. These two metastable states are analogous to the warm, El Niño and the cold, La Niña states in ENSO dynamics. Moreover, the two broadband peaks at low frequencies in the spectral density of the solution, as shown in Figure 5, are also reminiscent of the quasi-quadrennial and the quasi-biennial mode in ENSO [GR00, GAD$^+$02, JNG95] dynamics. The important role of such low-frequency variability in the understanding and prediction of climate dynamics on various time scales was emphasized in [CKG11, DG05, Ghi01, GR00, GR02] and references therein. □

6.3. **A highly idealized ENSO model with memory effects.** In this section, we consider the following periodically forced DDE with two discrete delays [GZT08]:

$$(6.5) \qquad \dot{x} = -\alpha \tanh(\kappa x(t-\tau_1)) + \beta \tanh(\kappa x(t-\tau_2)) + \gamma \cos(2\pi t).$$

This equation is a slightly modified version of the model used in [TSCJ94] for the study of the interaction between the seasonal forcing and the intrinsic ENSO variability.[14] We also refer to [BH89, CMZ90, Dij05, GCS15, GT00, GZT08, MCZ91, TSCJ94, ZG10, vR13, GZ15] and references therein for other models with retarded arguments used in this context.

The purpose of this subsection is to show that the Galerkin scheme developed in this article can also be easily adapted to deal with DDEs with multiple delays, as well as with non-autonomous, forced DDEs. For this purpose, Eq. (6.5) is placed in a quasi-periodic regime by choosing the parameter values $\alpha = 2.1, \beta = 1.05, \gamma = 3, \kappa = 10, \tau_1 = 0.95$, and $\tau_2 = 5.13$.

We outline now the necessary modifications to the nonlinear term $G(y)$ in the Galerkin system (5.19) for the case of multiple discrete delays. As a direct generalization of the case with a single discrete delay, given in footnote 13, the nonlinearity F in (2.4) takes the form $F(x(t), x(t-\tau_1), \cdots, x(t-\tau_p))$, where $0 < \tau_1 < \cdots < \tau_p =: \tau$. In this more general situation, the identity (5.15) becomes

$$(6.6) \qquad \left\langle \Pi_N \mathcal{F}\Big(\sum_{n=0}^{N-1} y_n(t)\mathcal{K}_n^\tau\Big), \mathcal{K}_j^\tau \right\rangle_{\mathcal{H}} = F\left(\sum_{n=0}^{N-1} y_n(t), \sum_{n=0}^{N-1} y_n(t)K_n(-\frac{\tau_1}{\tau}), \cdots, \right.$$
$$\left. \sum_{n=0}^{N-1} y_n(t)K_n(-\frac{\tau_{p-1}}{\tau}), \sum_{n=0}^{N-1} y_n(t)K_n(-1)\right),$$

which leads to the corresponding change in the formula (5.21) for the nonlinear vector field in Eq. (5.19). Note also that the forcing term is dealt with in Remark 5.1.

Again, we compare the DDE's attractor in the left panel of Fig. 6 with the attractor obtained from an associated Galerkin system, in the right panel. Despite the complexity of the DDE's attractor, a 40-dimensional Galerkin system can already provide a very accurate reconstruction of this attractor. The need for a higher dimensionality of the Galerkin approximation in this case arises from the presence of incommensurable frequencies in the periodically forced model, namely the seasonal cycle and the internal frequencies [Dij05, JNG94, GR00, GR02, TSCJ94].

---

[14]The nonlinearity used in [TSCJ94] consists of a sigmoid made of two tanh segments joined continuously by a line segment, cf. [MCZ91, Eq. (9)]; this sigmoid is simplified here to be just a hyperbolic tangent function.
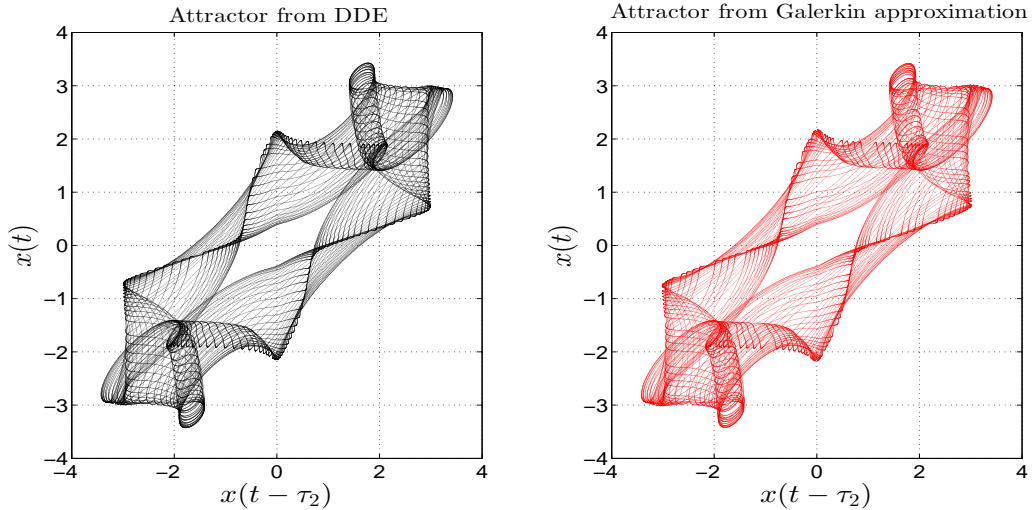
FIGURE 6. The attractor associated with Eq. (6.5) (left panel) and its approximation obtained from a 40D-Galerkin approximation (right panel).

APPENDIX A. PROOF OF THEOREM 3.1

The identity (3.5) follows from the definition of K_n given in (3.3) and the properties of the Legendre polynomials L_n given by (3.13) and (3.15). The normalization $K_n(1) = 1$ results directly from the identity (3.5) and the normalization $L_n(1) = 1$ as recalled in (3.12).

The orthogonality property of $(K_n, K_n(1))$ is proved in [Koo84, Theorem 3.1]. The formula (3.10) about the norm can be checked directly by using (3.5) and the orthogonality property of the Legendre polynomials recalled in (3.11).

Finally, it remains to show that $\left\{ \frac{\mathcal{K}_n}{\|\mathcal{K}_n\|_{\mathcal{E}}} \right\}$ forms a Hilbert basis of $\mathcal{E}$ which due to the orthonormality property boils down to show that the linear space spanned by the $\mathcal{K}_n$'s is dense in $\mathcal{E}$. In that respect, by recalling that $\mathcal{K}_n = (K_n, K_n(1))$ and that K_n are polynomials of degree n orthogonal under the inner product $\langle \cdot, \cdot \rangle_{\mathcal{E}}$, it follows that any function in the subspace

$$\mathcal{S}_n := \{(f, f(1)) : f \text{ is a polynomial of degree } n\}$$

admits a unique expansion in terms of $\{(K_0, K_0(1)), \cdots, (K_n, K_n(1))\}$. Note also that $\bigcup_{n=0}^{\infty} \mathcal{S}_n$ is dense in the subspace

$$\mathcal{C} := \{(f, f(1)) : f \text{ is a continuous function on } [-1, 1]\},$$

which is itself dense in

$$\widetilde{\mathcal{C}} := \{(f, a) : f \text{ is a continuous function on } [-1, 1], \text{ and } a \in \mathbb{R}\}.$$

Since the last subspace $\widetilde{\mathcal{C}}$ is dense in $\mathcal{E}$, the desired result follows.   □

APPENDIX B. PROOF OF PROPOSITION 5.1

Note that $a_n := (a_{n,0}, \cdots, a_{n,n-1})^{\text{tr}}$ given in (5.1) can be obtained by rewriting both sides in terms of the Legendre polynomials and comparing the coefficients, which leads to the algebraic equation $\mathbf{T} a_n = b_n$ to be satisfied by a_n. We provide below some details about the derivation of the matrix $\mathbf{T}$ and the vector b given in (5.3).

First note that by the definition of K_n given in (3.3), we get

$$
\text{(B.1)} \qquad \frac{\mathrm{d}K_n}{\mathrm{d}s}(s) = \frac{\mathrm{d}}{\mathrm{d}s}\left[-(1+s)\frac{\mathrm{d}}{\mathrm{d}s}L_n(s)\right] + (n^2+n+1)\frac{\mathrm{d}L_n(s)}{\mathrm{d}s}.
$$

By using then the recurrence formula for the derivative of L_n given in (3.15), we obtain

$$
\begin{aligned}
\text{(B.2)} \qquad \frac{\mathrm{d}K_n}{\mathrm{d}s}(s) ={}& \frac{\mathrm{d}}{\mathrm{d}s}\left[-(1+s)\sum_{k\in I_n}(2k+1)L_k(s)\right] \\
&+ (n^2+n+1)\sum_{k\in I_n}(2k+1)L_k(s) \\
={}& -\sum_{k\in I_n}(2k+1)L_k(s) - (1+s)\sum_{k\in I_n}(2k+1)\frac{\mathrm{d}L_k(s)}{\mathrm{d}s} \\
&+ (n^2+n+1)\sum_{k\in I_n}(2k+1)L_k(s) \\
={}& (n^2+n)\sum_{k\in I_n}(2k+1)L_k(s) \\
&- (1+s)\sum_{k\in I_n}\sum_{j\in I_k}(2k+1)(2j+1)L_j(s),
\end{aligned}
$$

where we recall from (3.16) that $I_n = \{k\in\mathbb{N} : 0\le k\le n-1, k+n \text{ is odd}\}$.

By using again the definition of K_n given in (3.3), the RHS of (5.1) can be rewritten in terms of the Legendre polynomials L_n as follows:

$$
\begin{aligned}
\text{(B.3)} \qquad & \sum_{k=0}^{n-1}a_{n,k}K_k(s) \\
&= \sum_{k=0}^{n-1}a_{n,k}\left[-(1+s)\frac{\mathrm{d}}{\mathrm{d}s}L_k(s) + (k^2+k+1)L_k(s)\right] \\
&= -(1+s)\sum_{k=0}^{n-1}\sum_{j\in I_k}a_{n,k}(2j+1)L_j(s) + \sum_{k=0}^{n-1}a_{n,k}(k^2+k+1)L_k(s).
\end{aligned}
$$

Now, by using (B.2) and (B.3) in the expansion $\frac{\mathrm{d}K_n}{\mathrm{d}s}(s) = \sum_{k=0}^{n-1}a_{n,k}K_k(s)$, we obtain

$$
\begin{aligned}
\text{(B.4)} \qquad & -(1+s)\sum_{k=0}^{n-1}\sum_{j\in I_k}a_{n,k}(2j+1)L_j(s) + \sum_{k=0}^{n-1}a_{n,k}(k^2+k+1)L_k(s) \\
&= (n^2+n)\sum_{k\in I_n}(2k+1)L_k(s) - (1+s)\sum_{k\in I_n}\sum_{j\in I_k}(2k+1)(2j+1)L_j(s).
\end{aligned}
$$

Finally, by using the three-term recurrence formula $(j+1)L_{j+1}(s) = (2j+1)sL_j(s) - jL_{j-1}(s)$ recalled in (3.13), we obtain from Eq. (B.4) that

$$
\begin{aligned}
\text{(B.5)} \quad & -\sum_{k=0}^{n-1}\sum_{j\in I_k} a_{n,k}\left[(j+1)L_{j+1}(s) + (2j+1)L_j(s) + jL_{j-1}(s)\right] \\
& +\sum_{k=0}^{n-1} a_{n,k}(k^2+k+1)L_k(s) = (n^2+n)\sum_{k\in I_n}(2k+1)L_k(s) \\
& -\sum_{k\in I_n}\sum_{j\in I_k}(2k+1)\left[(j+1)L_{j+1}(s) + (2j+1)L_j(s) + jL_{j-1}(s)\right].
\end{aligned}
$$

The system of algebraic equations given in (5.2)–(5.3) can then be obtained by equating the coefficients for $L_{n-1}, \cdots, L_0$ on both sides of the last equation above. The proof is now complete. □

Appendix C. Analytic formulas for nonlinear systems of DDEs

In this section, we extend the analytic formulas of Section 5 to nonlinear systems of delay equations of the form (4.65) based on the vectorized Koornwinder polynomials introduced in subsection 3.3.

Recall that (4.65) is given by

$$
\text{(C.1)} \qquad \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = L_S\mathbf{x}(t) + B\mathbf{x}(t-\tau) + \int_{t-\tau}^{t} C(s-t)\mathbf{x}(s)\mathrm{d}s + \mathbf{F}\left(\mathbf{x}(t), \int_{t-\tau}^{t}\mathbf{x}(s)\mathrm{d}s\right),
$$

where $\mathbf{F}: \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, and L_S, B, and C are as given in (4.47).

Recall also that the unknown function u_N in the corresponding Galerkin system (4.28) takes values in the Nd-dimensional subspace $\mathcal{H}_N$ defined in (4.26). We can thus rewrite u_N in terms of the first Nd vectorized Koornwinder polynomials (see (3.32)) as follows:

$$
\text{(C.2)} \qquad u_N(t) = \sum_{j=1}^{Nd} y_j(t)\mathbb{K}_j^\tau, \qquad t \geq 0,
$$

where

$$
\text{(C.3)} \qquad y_j(t) = \frac{\langle u_N(t), \mathbb{K}_j^\tau\rangle_{\mathcal{H}}}{\|\mathcal{K}_{j_q}^\tau\|_{\mathcal{H}_1}^2}, \qquad j = 1, \cdots, Nd,
$$

with j_q determined by (3.29) and $\mathcal{H}_1$ being the product space $L^2([-\tau,0];\mathbb{R}) \times \mathbb{R}$ equipped with the inner product given by (2.2).

In the following, we first deal with a special case in subsection C.1 by assuming that the linear operator C in the system (C.1) is time-independent. Necessary changes for the time-dependent case is then outlined in subsection C.2.

C.1. **The case when C in (C.1) is time-independent.** In this case, all of the three linear operators L_S, B and C are $d \times d$ matrices. By the same type of computation as given in Section 5 for the scalar

case, we obtain the following Nd-dimensional ODE system for $(y_1, \cdots, y_{Nd})$:

$$
\begin{aligned}
\frac{\mathrm{d}y_j}{\mathrm{d}t} = \frac{1}{\|\mathcal{K}_{j_q}\|_{\mathcal{E}}^2} \sum_{n=1}^{Nd} &\left[\frac{2}{\tau} \sum_{k=0}^{n_q-1} a_{n_q,k} \left(\delta_{j_q,k} \|\mathcal{K}_{j_q}\|_{\mathcal{E}}^2 - 1 \right) \delta_{n_r,j_r} \right. \\
&\left. + (L_S)_{j_r,n_r} + K_{n_q}(-1) B_{j_r,n_r} + \tau(2\delta_{n_q,0} - 1) C_{j_r,n_r} \right] y_n(t) \\
&+ \frac{1}{\|\mathcal{K}_{j_q}\|_{\mathcal{E}}^2} F_{j_r} \left(\sum_{n=1}^{Nd} y_n(t) \mathbf{K}_n(1), \ \tau \sum_{n=1}^{Nd} (2\delta_{n_q,0} - 1) \mathbf{K}_n(1) y_n(t) \right),
\end{aligned}
$$

(C.4)

where $j = 1, \cdots, Nd$, and where given an integer $l \in \{1, \cdots, Nd\}$, the associated integers $l_q \in \{0, \cdots, N-1\}$ and $l_r \in \{1, \cdots, d\}$ are determined by (3.29)-(3.30). The coefficients $a_{i,j}$ are determined by (5.1), and $\mathcal{E}$ denotes the product space $L^2([-1,1); \mathbb{R}) \times \mathbb{R}$ equipped with the inner product given by (3.9).

The above Galerkin system can be put into the following compact form:

(C.5)
$$
\boxed{\frac{\mathrm{d}y}{\mathrm{d}t} = Ay + G(y),}
$$

where Ay denotes the linear part of Eq. (C.4), and $G(y)$ the nonlinear part. Namely, A is the $Nd \times Nd$ matrix whose elements are given by

(C.6)
$$
\boxed{
\begin{aligned}
A_{j,n} = \frac{1}{\|\mathcal{K}_{j_q}\|_{\mathcal{E}}^2} &\left[\frac{2}{\tau} \sum_{k=0}^{n_q-1} a_{n_q,k} \left(\delta_{j_q,k} \|\mathcal{K}_{j_q}\|_{\mathcal{E}}^2 - 1 \right) \delta_{n_r,j_r} \right. \\
&\left. + (L_S)_{j_r,n_r} + K_{n_q}(-1) B_{j_r,n_r} + \tau(2\delta_{n_q,0} - 1) C_{j_r,n_r} \right],
\end{aligned}
}
$$

where $j, n = 1, \cdots, Nd$, and the nonlinear vector field $G \colon \mathbb{R}^{Nd} \to \mathbb{R}^{Nd}$, is given component-wisely by

(C.7)
$$
\boxed{
G_j(y) = \frac{1}{\|\mathcal{K}_{j_q}\|_{\mathcal{E}}^2} F_{j_r} \left(\sum_{n=1}^{Nd} y_n(t) \mathbf{K}_n(1), \ \tau \sum_{n=1}^{Nd} (2\delta_{n_q,0} - 1) \mathbf{K}_n(1) y_n(t) \right).
}
$$

C.2. **The case when C in (C.1) is time-dependent.** We outline below necessary modifications to (C.6) when $C \in L^2([-\tau, 0); \mathbb{R}^{d \times d})$.

Note that the contribution from the term $\int_{t-\tau}^{t} C(s-t)\mathbf{x}(s)\mathrm{d}s$ to the component $A_{j,n}$ given in (C.6) is:

(C.8)
$$
\alpha_{j,n} = \left\langle \int_{-\tau}^{0} C(\theta)\mathbf{K}_n^\tau(\theta)\mathrm{d}\theta, \ \mathbf{K}_j^\tau(0) \right\rangle.
$$

Note also that $\int_{-\tau}^{0} C(\theta)\mathbf{K}_n^\tau(\theta)\mathrm{d}\theta$ is a d-dimensional column vector, whose k^{th} component is given by $\int_{-\tau}^{0} C_{k,n_r}(\theta) K_{n_q}^\tau(\theta)\mathrm{d}\theta$ where $1 \le k \le d$.

It follows then

(C.9)
$$
\alpha_{j,n} = \int_{-\tau}^{0} C_{j_r,n_r}(\theta) K_{n_q}^\tau(\theta)\mathrm{d}\theta.
$$

Since each component of C belongs to $L^2([-\tau, 0); \mathbb{R})$, by the identity (3.24) given in Lemma 3.1, we have

$$(C.10) \qquad C_{i,j} = \sum_{l=0}^{\infty} \frac{\langle C_{i,j}, K_l^\tau \rangle_{L^2}}{\tau \|\mathcal{K}_l^\tau\|_{\mathcal{H}_1}^2} K_l^\tau, \quad 1 \leq i, j \leq d.$$

It follows then

$$\alpha_{j,n} = \sum_{l=0}^{\infty} \frac{\langle C_{j_r,n_r}, K_l^\tau \rangle_{L^2}}{\tau \|\mathcal{K}_l^\tau\|_{\mathcal{H}_1}^2} \int_{-\tau}^{0} K_l^\tau(\theta) K_{n_q}^\tau(\theta) \mathrm{d}\theta.$$

Since $\langle K_i^\tau, K_j^\tau \rangle_{\mathcal{H}_1} = \delta_{i,j}$, we get

$$\frac{1}{\tau} \int_{-\tau}^{0} K_i^\tau(\theta) K_j^\tau(\theta) \mathrm{d}\theta = (\delta_{i,j} \|\mathcal{K}_i^\tau\|_{\mathcal{H}_1}^2 - 1), \quad i, j \in \mathbb{N},$$

leading thus to

$$(C.11) \qquad \begin{aligned} \alpha_{j,n} &= \sum_{l=0}^{\infty} \frac{\langle C_{j_r,n_r}, K_l^\tau \rangle_{L^2}}{\|\mathcal{K}_l^\tau\|_{\mathcal{H}_1}^2} (\delta_{l,n_q} \|\mathcal{K}_l^\tau\|_{\mathcal{H}_1}^2 - 1) \\ &= \langle C_{j_r,n_r}, K_{n_q}^\tau \rangle_{L^2} - \sum_{l=0}^{\infty} \frac{\langle C_{j_r,n_r}, K_l^\tau \rangle_{L^2}}{\|\mathcal{K}_l^\tau\|_{\mathcal{H}_1}^2} \\ &= \langle C_{j_r,n_r}, K_{n_q}^\tau \rangle_{L^2}, \end{aligned}$$

where we have used $\sum_{l=0}^{\infty} \frac{\langle C_{j_r,n_r}, K_l^\tau \rangle_{L^2}}{\|\mathcal{K}_l^\tau\|_{\mathcal{H}_1}^2} = 0$, which follows from the identity (3.22) in Lemma 3.1.

So the only modification to (C.6) is to replace the term $\tau(2\delta_{n_q,0} - 1)C_{j_r,n_r}$ by $\langle C_{j_r,n_r}, K_{n_q}^\tau \rangle_{L^2}$. Note that when $C_{j_r,n_r} \in L^2([-\tau, 0); \mathbb{R})$ is a constant function, the latter is reduced into the former since

$$\int_{-\tau}^{0} K_n^\tau(\theta) \mathrm{d}\theta = \tau(2\delta_{n,0} - 1), \quad n \in \mathbb{N}.$$

To summarize, when C is time-dependent, the Nd-dimensional Galerkin approximation associated with (C.1) still takes the form (C.5) with the nonlinear part $G(y)$ given by (C.7), and the elements of the $Nd \times Nd$ matrix A, given by

$$(C.12) \qquad \boxed{\begin{aligned} A_{j,n} = \frac{1}{\|\mathcal{K}_{j_q}\|_{\mathcal{E}}^2} &\left[\frac{2}{\tau} \sum_{k=0}^{n_q-1} a_{n_q,k} \left(\delta_{j_q,k} \|\mathcal{K}_{j_q}\|_{\mathcal{E}}^2 - 1 \right) \delta_{n_r,j_r} \right. \\ &\left. + (L_S)_{j_r,n_r} + K_{n_q}(-1) B_{j_r,n_r} + \langle C_{j_r,n_r}, K_{n_q}^\tau \rangle_{L^2} \right]. \end{aligned}}$$

References

[AMR97]   M. Alfaro, F. Marcellán, and M. L. Rezola, *Estimates for Jacobi-Sobolev type orthogonal polynomials*, Appl. Anal. **67** (1997), no. 1-2, 157–174.

[BB78]   H. T. Banks and J. A. Burns, *Hereditary control problems: Numerical methods based on averaging approximations*, SIAM J. Control Optim. **16** (1978), 169–208.

[BGV82]   K Bhattacharya, M Ghil, and IL Vulis, *Internal variability of an energy-balance model with delayed albedo effects*, Journal of the Atmospheric Sciences **39** (1982), no. 8, 1747–1773.

[BH89]   D. S. Battisti and A. C. Hirst, *Interannual variability in a tropical atmosphere-ocean model: Influence of the basic state, ocean geometry and nonlinearity*, J. Atmos. Sci. **46** (1989), 1687–1712.

[BHS83]   J.A. Burns, T.L. Herdman, and H.W. Stech, *Linear functional differential equations as semigroups on product spaces*, SIAM Journal on Mathematical Analysis **14** (1983), no. 1, 98–116.

[BK79]   H. T. Banks and F. Kappel, *Spline approximations for functional differential equations*, Journal of Differential Equations **34** (1979), no. 3, 496–522.

[Bré10]   H. Brézis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, 2010.

[BRI84]   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.

[CF80]   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.

[CGH12]   M. D. Chekroun and N. E. Glatt-Holtz, *Invariant measures for dissipative dynamical systems: abstract results and applications*, Commun. Math. Phys. **316** (2012), no. 3, 723–761.

[CH98]   T. Cazenave and A. Haraux, *An Introduction to Semilinear Evolution Equations*, Oxford Lecture Series in Mathematics and its Applications, vol. 13, The Clarendon Press, Oxford, 1998.

[CKG11]   M. D. Chekroun, D. Kondrashov, and M. Ghil, *Predicting stochastic systems by noise sampling, and application to the El Niño-Southern Oscillation*, Proc. Natl. Acad. Sci. USA **108** (2011), 11766–11771.

[CMP77]   S. N. Chow and J. Mallet-Paret, *Integral averaging and bifurcation*, J. Differential Equations **26** (1977), no. 1, 112–159.

[CMZ90]   M. A. Cane, M. Münnich, and S. E. Zebiak, *A study of self-excited oscillations of the tropical ocean-atmosphere system. Part I: Linear analysis*, J. Atmos. Sci. **47** (1990), 1562–1577.

[CNK+14]   M. D. Chekroun, J. D. Neelin, D. Kondrashov, J. C. McWilliams, and M. Ghil, *Rough parameter dependence in climate models: The role of Ruelle-Pollicott resonances*, PNAS **111** (2014), no. 5, 1684–1690.

[CZ95]   R.F. Curtain and H. Zwart, *An Introduction to Infinite-Dimensional Linear Systems Theory*, vol. 21, Springer, 1995.

[DC02]   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.

[DG05]   H. A. Dijkstra and M. Ghil, *Low-frequency variability of the large-scale ocean circulation: A dynamical systems approach*, Rev. Geophys. **43, RG3002** (2005), no. RG3002.

[DHL14]   J. I. Diaz, A. Hidalgo, and Tello L., *Multiple solutions and numerical analysis to the dynamic and stationary models coupling a delayed energy balance model involving latent heat and discontinuous albedo with a deep ocean*, Proc. R. Soc. A **470** (2014), no. 20140376.

[Dij05]   H. A. Dijkstra, *Nonlinear Physical Oceanography: A Dynamical Systems Approach to the Large-Scale Ocean Circulation and El Niño*, Springer, New York/Berlin, 2005.

[DvGVLW95]   O. Diekmann, S. A. van Gils, S. M. Verduyn Lunel, and H.-O. Walther, *Delay Equations: Functional, Complex, and Nonlinear Analysis*, Applied Mathematical Sciences, vol. 110, Springer-Verlag, New York, 1995.

[ER85]   J.-P. Eckmann and D. Ruelle, *Ergodic theory of chaos and strange attractors*, Rev. Modern Phys. **57** (1985), 617–656.

[FM95]   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.

[GAD+02]   M. Ghil, M. R. Allen, M. D. Dettinger, K. Ide, D. Kondrashov, M. E. Mann, A. W. Robertson, A. Saunders, Y. Tian, F. Varadi, and P. Yiou, *Advanced spectral methods for climatic time series*, Rev. Geophys. **40** (2002), no. 1, 1–41.

[GC87]   M. Ghil and S. Childress, *Topics in Geophysical Fluid Dynamics: Atmospheric Dynamics, Dynamo Theory and Climate Dynamics*, Applied Mathematical Sciences, Springer-Verlag, New York/Berlin/London/Paris/ Tokyo, 1987.

[GCS15]   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), no. 20150097.

[Ghi01]   M. Ghil, *Hilbert problems for the geosciences in the 21st century*, Nonlin. Processes Geophys. **8** (2001), 211–222.

[GR00]   M. Ghil and A. W. Robertson, *Solving problems with GCMs: General circulation models and their role in the climate modeling hierarchy*, General Circulation Model Development: Past, Present and Future (D. Randall, ed.), Academic Press, San Diego, 2000, pp. 285–325.

[GR02]   ———, *"Waves" vs."particles" in the atmosphere's phase space: A pathway to long-range forecasting?*, Proc. Natl. Acad. Sci. USA **99, Suppl. 1** (2002), no. Suppl. 1, 2493–2500.

[GT00]   E. Galanti and E. Tziperman, *ENSO's phase locking to the seasonal cycle in the fast-SST, fast-wave, and mixed-mode regimes*, J. Atmos. Sci. **57** (2000), 2936–2950.

[GZ15]   M. Ghil and I. Zaliapin, *Understanding ENSO variability and its extrema: A delay differential equation approach*, Observations, Modeling and Economics of Extreme Events (M. Chavez, M. Ghil, and J. Urrutia-Fucugauchi, eds.), American Geophysical Union, in press, 2015.

[GZT08]   M. Ghil, I. Zaliapin, and S. Thompson, *A delay differential model of ENSO variability: parametric instability and the distribution of extremes*, Nonlin. Processes Geophys. **15** (2008), no. 3, 417–433.

[HVL93]   J. K. Hale and S. M. Verduyn Lunel, *Introduction to Functional-Differential Equations*, Applied Mathematical Sciences, vol. 99, Springer-Verlag, New York, 1993.

[Ism05]   M. E. H. Ismail, *Classical and Quantum Orthogonal Polynomials in One Variable*, Encyclopedia of Mathematics and its Applications, vol. 98, Cambridge University Press, Cambridge, 2005.

[IT86]   K. Ito and R. Teglas, *Legendre-tau approximations for functional-differential equations*, SIAM J. Control Optim. **24** (1986), no. 4, 737–759.

[JNG94]   F. F. Jin, J. D. Neelin, and M. Ghil, *El Niño on the Devil's Staircase: Annual subharmonic steps to chaos*, Science **264** (1994), 70–72.

[JNG95]   N. Jiang, D. Neelin, and M. Ghil, *Quasi-quadrennial and quasi-biennial variability in the equatorial Pacific*, Clim. Dyn. **12** (1995), no. 2, 101–112.

[Kap86]   F. Kappel, *Semigroups and delay equations*, Semigroups, Theory and Applications, Vol. II (Trieste, 1984), Pitman Res. Notes Math. Ser., vol. 152, Longman Sci. Tech., Harlow, 1986, pp. 136–176.

[Koo84]   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.

[Kri08]   T. Krisztin, *Global dynamics of delay differential equations*, Periodica Mathematica Hungarica **56** (2008), no. 1, 83–95.

[KS78]   F. Kappel and W. Schappacher, *Autonomous nonlinear functional differential equations and averaging approximations*, Nonlinear Analysis: Theory, Methods & Applications **2** (1978), no. 4, 391–422.

[KS87]   F. Kappel and D. Salamon, *Spline approximation for retarded systems and the Riccati equation*, SIAM J.Control Optim. **25** (1987), 1082–1117.

[KS14]   B. Krauskopf and J. Sieber, *Bifurcation analysis of delay-induced resonances of the El-Niño Southern Oscillation*, Proc. R. Soc. A **470** (2014), no. 20140348.

[Kua93]   Y. Kuang, *Delay Differential Equations with Applications in Population Dynamics*, Academic Press, 1993.

[Kun82]   K. Kunisch, *Approximation schemes for the linear-quadratic optimal control problem associated with delay equations*, SIAM J. Control and Optimization **20** (1982), no. 4, 506–540.

[KWVdD78]   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.

[KZ86]   F. Kappel and K. Zhang, *Equivalence of functional-differential equations of neutral type and abstract Cauchy problems*, Monatshefte für Mathematik **101** (1986), no. 2, 115–133.

[LS10]   M. Lakshmanan and D. V. Senthilkumar, *Dynamics of Nonlinear Time-delay Systems*, Springer, Heidelberg, 2010.

[Lun04]   A. Lunardi, *Linear and nonlinear diffusion problems*, Lecture Notes, 2004.

[Mac89]   N. MacDonald, *Biological Delay Systems: Linear Stability Theory*, Cambridge Studies in Mathematical Biology, vol. 8, Cambridge University Press, Cambridge, 1989.

[MCZ91]   M. Münnich, M. A. Cane, and S. E. Zebiak, *A study of self-excited oscillations of the tropical ocean-atmosphere system. Part II: Nonlinear cases*, J. Atmos. Sci. **48** (1991), 1238–1248.

[MN07]   W. Michiels and S.-I. Niculescu, *Stability and Stabilization of Time-Delay Systems: An Eigenvalue-based Approach*, Advances in Design and Control, vol. 12, SIAM, Philadelphia, PA, 2007.

[MPS96]   J. Mallet-Paret and G.R. Sell, *The Poincaré–Bendixson theorem for monotone cyclic feedback systems with delay*, Journal of Differential Equations **125** (1996), no. 2, 441–489.

[Nay08]   A. H. Nayfeh, *Order reduction of retarded nonlinear systems–the method of multiple scales versus center-manifold reduction*, Nonlinear Dyn. **51** (2008), 483–500.

[NY89]   S.-I. Nakagiri and M. Yamamoto, *Controllability and observability of linear retarded systems in banach spaces*, International Journal of Control **49** (1989), no. 5, 1489–1504.

[Paz83]   A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Applied Mathematical Sciences, vol. 44, Springer-Verlag, New York, 1983.

[RCC+14]   L. Roques, M. D. Chekroun, M. Cristofol, S. Soubeyrand, and M. Ghil, *Parameter estimation for energy balance models with memory*, Proc. R. Soc. A **470** (2014), no. 2169, 20140349.

[Smi11]   H. Smith, *An Introduction to Delay Differential Equations with Applications to the Life Sciences*, Texts in Applied Mathematics, vol. 57, Springer, New York, 2011.

[Spr07]   J. C. Sprott, *A simple chaotic delay differential equation*, Physics Letters A **366** (2007), no. 4, 397–402.

[Ste89]   G. Stepan, *Retarded Dynamical Systems: Stability and Characteristic Functions*, Pitman Research Notes in Mathematics Series, vol. 210, Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1989.

[STW11]   J. Shen, T. Tang, and L.-L. Wang, *Spectral Methods: Algorithms, Analysis and Applications*, Springer Series in Computational Mathematics, vol. 41, Springer, Heidelberg, 2011.

[Sze75]   G. Szegő, *Orthogonal Polynomials*, fourth ed., American Mathematical Society, Providence, R.I., 1975.

[TSCJ94]   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.

[Uva69]   V. B. Uvarov, *The connection between systems of polynomials that are orthogonal with respect to different distribution functions*, USSR Computat. Math. Math. Phys. **9** (1969), 25–36.

[vR13]   T. Živković and K. Rypdal, *ENSO dynamics: Low-dimensional-chaotic or stochastic?*, J. Geophys. Res. Atmos. **118** (2013), 2161–2168.

[Vya12]   C. P. Vyasarayani, *Galerkin approximations for higher order delay differential equations*, J. Comput. Nonlinear Dynamics **7** (2012), 031004.

[Wan09]   Xiaoming Wang, *Approximating stationary statistical properties*, Chinese Annals of Mathematics, Series B **30** (2009), no. 6, 831–844.

[WC05]   P. Wahi and A. Chatterjee, *Galerkin projections for delay differential equations*, ASME. J. Dyn. Syst., Meas., Control **127** (2005), 80–87.

[Web76]   G. F. Webb, *Functional differential equations and nonlinear semigroups in L^p-spaces*, J. Diff. Equat. **20** (1976), 71–89.

[WWP$^+$94] W Wischert, A Wunderlin, A Pelster, M Olivier, and J Groslambert, *Delay-induced instabilities in nonlinear feedback systems*, Physical Review E **49** (1994), no. 1, 203.

[ZG10]   I. Zaliapin and M. Ghil, *A delay differential model of ENSO variability, Part 2: Phase locking, multiple solutions, and dynamics of extrema*, Nonlin. Processes Geophys. **17** (2010), 123–135.

(MDC) Department of Atmospheric & Oceanic Sciences, University of California, Los Angeles, CA 90095-1565, USA
    E-mail address: mchekroun@atmos.ucla.edu

(MG) Geosciences Department and Laboratoire de Météorologie Dynamique (CNRS and IPSL),École Normale Supérieure, F-75231 Paris Cedex 05, France; and Department of Atmospheric & Oceanic Sciences, University of California,Los Angeles, CA 90095-1565, USA
    E-mail address: ghil@atmos.ucla.edu

(HL) Department of Atmospheric & Oceanic Sciences, University of California, Los Angeles, CA 90095-1565, USA; and Department of Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, USA
    E-mail address: hhliu@vt.edu

(SW) Department of Mathematics, Indiana University
    E-mail address: showang@indiana.edu