2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

A projection approach for model reduction of large-scale time-delay systems, with application to a boundary controlled PDE

Wim Michiels, Elias Jarlebring and Karl Meerbergen

***Abstract*— We present a model order reduction method which allows the construction of a reduced, delay free model of a given dimension for linear time-delay systems. The method builds on the equivalent representation of the time-delay system as an infinite-dimensional linear problem. It combines ideas from a finite-dimensional approximation via a spectral discretization on the one hand, and a Krylov-Padé model reduction approach on the other hand. The method exhibits a good spectral approximation of the original model, in the sense that the smallest characteristic roots are well approximated and the non-converged eigenvalues of the reduced model have a favorable location, and it preserves moments at zero and at infinity. The model reduction approach is illustrated by means of a PDE model for a heated rod with delay in the boundary control.**

I. Introduction

We consider a time-delay system with a single input and a single output of the form

$$\left\{ \begin{array}{lcl} \dot{x}(t) & = & A_0 x(t) + \sum_{i=1}^m A_i x(t-\tau_i) + Bu(t), \\ y & = & Cx(t) + Du(t), \end{array} \right. \quad (1)$$

where $x(t) \in \mathbb{R}^n$ is the state variable at time t, $u \in \mathbb{R}^1$ is the input, $y \in \mathbb{R}^1$ is the output and $\tau_i,\ i = 1, \ldots, m$, represent time-delays. We assume that

$$0 < \tau_1 < \cdots < \tau_m.$$

The transfer function of the system (1) is given by

$$\gamma(s) := C \left(sI - A_0 - \sum_{i=1}^m A_i e^{-s\tau_i} \right)^{-1} B + D. \quad (2)$$

The general problem we consider is to approximate the system (1) with a standard linear dynamical system without delay, in our context conveniently written as

$$\left\{ \begin{array}{lcl} G\dot{z}(t) & = & z(t) + Hu(t), \\ y(t) & = & Fz(t) + Du(t), \end{array} \right. \quad (3)$$

where $z(t) \in \mathbb{R}^{k+1}$. As usual in a model reduction setting, we wish to find a reduced model of a given dimension, which is typically much smaller than the dimension of the original model, i.e., $k \ll n$. More precisely, in this paper, we will present an efficient algorithm for computing G, H, F and D of small dimension such that the reduced system (3) approximates the original system (1) both in terms of approximation of the characteristic roots and in terms of derivatives of the transfer function at the origin and at infinity.

W. Michiels, E. Jarlebring and K. Meerbergen are with the Department of Computer Science, Katholieke Universiteit Leuven, Belgium, firstname.lastname@cs.kuleuven.be

Results related to the considered problem, the reduction of the infinite-dimensional system (1) to the delay free system (3) are rare in the literature, and, in fact, many problems related to such model reduction of time-delay systems are generally considered to be unsolved [17]. We note however that the availability of accurate reduced order models is favorable for many purposes [1]. For example, it may make the numerical simulation of large-scale systems computationally feasible. A reduced model as derived in the paper is also very tractable in the context of control design, not only because of the low dimension, but also because of the standard state space representation, on which many control design techniques build.

In the derivation of the proposed method we start by rewriting the system (1) as an equivalent infinite-dimensional linear system, as in [7]. The discretization of this system leads to a standard finite-dimensional linear problem, which is more suitable for model reduction purposes. The followed discretization approach is based on a spectral approximation, inspired by [5] where the corresponding eigenvalue problem was addressed. The accuracy depends on the choice of interpolation points in the interval $[-\tau_m, 0]$. Different choices of these points lead to different discretizations. We will choose the points in such a way that the accuracy of the eigenvalues is optimized and, at the same time, structure and sparsity can be introduced in the system matrices. Furthermore we prove that the transfer function of the discretized system matches several moments with the original transfer function (2). In the next step, where we project the (large) discretized system on a subspace, we guarantee that these moment matching properties are carried over to the reduced model by using a Padé via Krylov like model reduction method. In addition, by exploiting the structure of the problem during the construction of the Krylov space, as in [13], the process can be made dynamic in the sense that the number of discretization points in the spectral approximation does not need to be chosen beforehand, and the model reduction process can always be resumed if the accuracy of the reduced model is not sufficient.

For model reduction of linear systems based on moment matching, the Padé via Lanczos method and its variations are probably best known [8], [11], [10], [3]. These methods build a two-sided Krylov subspace with the system matrix and the input and output vectors as starting vectors. Krylov methods are also used for balanced truncation type of methods, see e.g. [12], and for parameterized model order reduction, see e.g. [2]. The advantage of two-sided methods is that both the input and the output are taken into account in the reduced

models, which leads to matching twice as many moments for a given dimension of the reduced model compared to the case where only the input is taken into account. In many applications only one-sided methods are used. In particular, one-sided Krylov-Padé methods have become popular for the solution of large scale finite element models in structures and vibrations, which often lead to polynomial eigenvalue problems, see, e.g., [18], [4]. The proposed approach to derive a reduced model for (1) relies on a one-sided Krylov-Padé method, because this allows us to fully exploit the structure of the problem. In addition, the proposed method leads to a good approximation of the smallest characteristic roots of the time-delay system. Since the rightmost characteristic roots are typically among the smallest ones [16], this makes the reduced models suitable for control design purposes. Two-sided methods use operations with the transpose of the matrices. Matrix operations with the transpose cannot exploit the special structure of the matrices, which implies that two sided methods do not preserve the structure and do not have the same dynamic properties as the one-sided method.

In the context sketched above, we recall that for linear systems of the form (3), the moments at zero are $-FH+D, -FGH, -2FG^2H, \ldots$ The moments at infinity, called Markov moments, are $D, FG^{-1}H, FG^{-2}H, \ldots$ In this paper, we will derive a reduced model from the Krylov sequence

$$\{G^{-1}H, H, GH, G^2H, \ldots G^{k-2}H\},$$

which matches the first two moments at infinity and $k-1$ moments at zero. The reason why moments are added at infinity is that the spectral discretization satisfies this property, and we want to keep it in the reduced model.

The structure of the paper is as follows. In Section II, we formulate the equivalent infinite-dimensional problem, outline the spectral discretization and discuss the properties of the discretized system, with the emphasis on moment matching. In Section III we present the dynamic Arnoldi method for the infinite-dimensional linear systems, and show how moment matching properties can be imposed on the reduced model. In Section IV, we illustrate the method and its properties by means of a numerical example.

Due to space limitations, we have omitted the proofs of some of the technical results. These can be found in the publicly available report [15].

II. Finite-dimensional approximation

An approach to analyze the time-delay system (1) is to rewrite it in a linear infinite-dimensional form. The corresponding operators can be discretized, yielding an approximation of (1), involving large matrices and no delays. This is briefly summarized in the next paragraph. In §II-B we discuss some properties of the discretized problem. As the main result of the section we prove that the discretized problem fulfills a moment matching property, which will play an important role in the derivation of the reduced model

A. A spectral discretization

Consider the space $X := \mathbb{R}^n \times \mathcal{L}_2([-\tau_m, 0], \mathbb{R}^n)$, equipped with the inner product

$$< (y_0, y_1), (z_0, z_1) >_X = < y_0, z_0 >_{\mathbb{R}^n} + < y_1, z_1 >_{\mathcal{L}_2} .$$

Let $\mathcal{A}: \ X \to X$ be the derivative operator defined by

$$\begin{aligned} \mathcal{D}(\mathcal{A}) = \{ z = (z_0, z_1) \in X : z_1 \in \mathcal{C}([-\tau_m, \ 0], \ \mathbb{C}^n), \\ z_1' \in \mathcal{C}([-\tau_m, 0], \mathbb{C}^n), \ z_0 = z_1(0) \}, \end{aligned}$$

$$\mathcal{A}z = \begin{pmatrix} A_0 z_0 + \sum_{i=1}^m A_i z_1(-\tau_i), & z_1' \end{pmatrix}, \ z \in \mathcal{D}(\mathcal{A}) \tag{4}$$

and let the operators $\mathcal{B}: \ \mathbb{C} \to X$ and $\mathcal{C}: \ X \to \mathbb{C}$ be given by

$$\begin{aligned} \mathcal{B}u &= \begin{pmatrix} Bu, & 0 \end{pmatrix}, \quad u \in \mathbb{C}, \\ \mathcal{C}z &= Cz_0, \qquad\qquad z = (z_0, z_1) \in X. \end{aligned}$$

We can now rewrite (1) as

$$\begin{cases} \dot{z}(t) &= \mathcal{A}z(t) + \mathcal{B}u(t), \\ y(t) &= \mathcal{C}z(t) + Du(t), \end{cases} \tag{5}$$

where $z(t) \in \mathcal{D}(A) \subset X$.

We outline how system (5) can be discretized using a *spectral method* (see, e.g. [19], [5]). Given a positive integer N, we consider a mesh Ω_N of $N+1$ distinct points in the interval $[-\tau_m, \ 0]$:

$$\Omega_N = \{\theta_{N,i}, \ i = 1, \ldots, N+1\}, \tag{6}$$

where

$$-\tau_m \leq \theta_{N,1} < \ldots < \theta_{N,N} < \theta_{N,N+1} = 0.$$

This allows us to replace the continuous space X with the space X_N of discrete functions defined over the mesh Ω_N, i.e. any function $\phi \in X$ is discretized into a block vector $x = [x_1^T \cdots \ x_{N+1}^T]^T \in X_N$ with components

$$x_i = \phi(\theta_{N,i}) \in \mathbb{C}^n, \quad i = 1, \ldots, N+1.$$

We let $\mathcal{P}_N x, \ x \in X_N$, be the unique $\mathbb{C}^n$ valued interpolating polynomial of degree smaller than or equal to N, satisfying

$$\mathcal{P}_N x(\theta_{N,i}) = x_i, \quad i = 1, \ldots, N+1.$$

In this way we can approximate the operator $\mathcal{A}$ over X with the matrix $A_N: \ X_N \to X_N$, defined as

$$\begin{cases} (\mathcal{A}_N \ x)_i = & (\mathcal{P}_N x)' (\theta_{N,i}), \ i = 1, \ldots, N, \\ (\mathcal{A}_N \ x)_{N+1} = & A_0 \mathcal{P}_N x(0) + \sum_{i=1}^m A_i \mathcal{P}_N x(-\tau_i). \end{cases} \tag{7}$$

Using the Lagrange representation of $\mathcal{P}_N x$,

$$\mathcal{P}_N x = \sum_{k=1}^{N+1} l_{N,k} \ x_k,$$

where the Lagrange polynomials $l_{N,k}$ are real valued polynomials of degree N satisfying

$$l_{N,k}(\theta_{N,i}) = \begin{cases} 1 & i = k, \\ 0 & i \neq k, \end{cases}$$

we get an explicit form for the matrix A_N,

$$A_N = \begin{bmatrix} d_{1,1} & \dots & d_{1,N+1} \\ \vdots & & \vdots \\ d_{N,1} & \dots & d_{N,N+1} \\ a_1 & \dots & a_{N+1} \end{bmatrix} \in \mathbb{R}^{(N+1)n \times (N+1)n}, \quad (8)$$

where

$$\begin{cases} d_{i,k} &= l'_{N,k}(\theta_{N,i}) I_n, \\ a_k &= A_0 l_{N,k}(0) + \sum_{i=1}^{m} A_i l_{N,k}(-\tau_i). \end{cases}$$

In the same way we can approximate $\mathcal{B}$ and $\mathcal{C}$ by

$$B_N = [0 \ \cdots \ 0 \ 1]^T \otimes B, \quad C_N = [0 \ \cdots \ 0 \ 1] \otimes C$$

and we arrive at the finite-dimensional approximation of (1),

$$\begin{cases} \dot{z}(t) = A_N z(t) + B_N u(t), \ z(t) \in \mathbb{R}^{(N+1)n \times 1}, \\ y(t) = C_N z(t) + D u(t). \end{cases} \quad (9)$$

Accordingly, we can approximate (2) by the transfer function of (9), given by

$$\gamma_N(s) := C_N \left(sI - A_N\right)^{-1} B_N + D. \quad (10)$$

B. Properties

The discretized system (9) has the favorable property that several derivatives at the origin and the first derivative at infinity of the transfer function original (2) and the corresponding approximation (10) coincide.

Theorem 1: The transfer functions (2) and (10) satisfy,

$$\left.\frac{d^i \gamma_N(s)}{ds^i}\right|_{s=0} = \left.\frac{d^i \gamma(s)}{ds^i}\right|_{s=0}, \quad i = 0, \dots, N, \quad (11)$$

and

$$\left.\frac{d^i \gamma_N(s^{-1})}{ds^i}\right|_{s=0} = \left.\frac{d^i \gamma(s^{-1})}{ds^i}\right|_{s=0}, \quad i = 0, 1, \quad (12)$$

that is, the moments of $\gamma(s)$ and $\gamma_N(s)$ at zero match up to the Nth moment, and the moments at infinity match up to the first moment.

It is important to note that the properties described by Theorem 1 are *independent* of the choice of the grid points. Hence, other desired properties can be imposed by an optimal choice of the distribution of the grid points.

In what follows we choose the nonzero grid points as scaled and shifted zeros of U_N, the Chebyshev polynomial of the second kind and order N, i.e. the grid points are specified as

$$\theta_{N,i} = \frac{\tau_m}{2}(\alpha_{N,i} - 1), \quad \alpha_{N,i} = -\cos\frac{\pi i}{N+1}, \ i = 1, \dots, N+1. \quad (13)$$

With the choice of the Chebyshev grid (13) the convergence of the individual eigenvalues of A_N to corresponding characteristic roots is fast. More specifically, in [5] it is proven that spectral accuracy (approximation error $O(N^{-N})$) is obtained. An additional property of using a Chebyshev grid, observed in extensive numerical simulations, is that the eigenvalues of A_N, which have not yet converged to corresponding characteristic roots, are located to the left of the eigenvalues that have already converged (see, e.g., the plots in [5]). Other arguments for choosing the grid points (13) are given in the following section.

III. Constructing a reduced-order model

We now know that the discretized system (9) has the nice approximation property that many moments are matched. It is however not a solution to our main problem to construct a small reduced model since the state space dimension of (9) is $n(N+1)$, i.e., even much larger than the state space dimension of the original time-delay system (1).

However, the discretized system (9) is a standard form. We could hence conceptually reduce the dimension of the discretized system by applying a standard Krylov based model reduction technique on (9). This would involve explicitly constructing the large matrices in (9). We will now see that this can be avoided and an efficient implementation becomes possible, where many properties of (9) are exploited. Moreover, the construction is dynamic in the sense that the value of N in (10) does not need to be fixed beforehand.

More specifically, in §III-A we derive an equivalent representation of (9) and (10), where the matrices have a sparse structure. In §III-B we dynamically construct a Krylov space. In §III-C we project the system matrices on this subspace and outline how moment matching properties can be guaranteed.

The technical derivation of the results makes use of the representation of polynomials related to the spectral discretization in appropriately defined Chebyshev bases. In what follows we denote by T_i the Chebyshev polynomial of the first kind and order i, and U_i is the Chebyshev polynomial of the second kind and order i, with $i \geq 0$.

A. A sparse reformulation of the problem

In the derivation we will use a slightly different formulation of the discretization. Note that the eigenvalue problem

$$(sI - A_N)x = 0, \quad x \in \mathbb{C}^{(N+1)n}, \ x \neq 0, \quad (14)$$

where A_N is given by (7), can directly be obtained by requiring that there exists a polynomial of degree N,

$$(\mathcal{P}_N x)(t) = \sum_{k=0}^{N} l_{N,k}(t) \ x_k,$$

which satisfies the conditions

$$\begin{cases} s\mathcal{P}_N x(\theta_{N,i}) = (\mathcal{P}_N x)'(\theta_{N,i}), \quad i \in \{1, \dots, N\}, \\ s\mathcal{P}_N x(0) = A_0 \mathcal{P}_N x(0) + \sum_{i=1}^{m} A_i \mathcal{P}_N x(-\tau_i). \end{cases} \quad (15)$$

Then the vector x in (14) is obtained as $x = [x_0^T \ \cdots \ x_N^T]^T$. Hence, an eigenvalue problem equivalent to (14) can be obtained by expressing $\mathcal{P}_N x$ in another basis and imposing the same conditions. We now represent $\mathcal{P}_N x$ in a basis of Chebyshev polynomials:

$$(\mathcal{P}_N x)(t) = \sum_{i=0}^{N} c_i T_i\left(2\frac{t}{\tau_m} + 1\right), \quad (16)$$

where $c_i \in \mathbb{C}^{N \times 1}$ for $i = 0, \dots, N$. By requiring that this polynomial satisfies the conditions (15) we obtain an equivalent sparse eigenvalue problem for (14), as expressed in the following lemma (Theorem 2.1 from [13]).

Lemma 2: If the grid points in the spectral discretization of (5) are chosen as (13), then the eigenvalue problem (14) is equivalent with

$$(s\Pi_N - \Sigma_N)\, c = 0,\ s \in \mathbb{C},\ c \in \mathbb{C}^{(N+1)n},\ c \neq 0, \quad (17)$$

where

$$\Pi_N = \frac{\tau_m}{4} \begin{bmatrix} \frac{4}{\tau_m} & \frac{4}{\tau_m} & \frac{4}{\tau_m} & \cdots & & \cdots & \frac{4}{\tau_m} \\ 2 & 0 & -1 & & & & \\ & \frac{1}{2} & 0 & -\frac{1}{2} & & & \\ & & \frac{1}{3} & 0 & \ddots & & \\ & & & \frac{1}{4} & \ddots & -\frac{1}{N-2} & \\ & & & & \ddots & 0 & -\frac{1}{N-1} \\ & & & & & \frac{1}{N} & 0 \end{bmatrix} \otimes I \quad (18)$$

and

$$\Sigma_N = \begin{bmatrix} R_0 & R_1 & \cdots & R_N \\ & I_n & & \\ & & \ddots & \\ & & & I_n \end{bmatrix}, \quad (19)$$

with

$$R_i = A_0 T_i(1) + \sum_{k=1}^{m} A_k T_i\left(-2\frac{\tau_k}{\tau_m} + 1\right),\ i = 0, \ldots, N.$$

A comparison between (14) and (17), taking into account the interpretation of the vectors c and d as coefficients in polynomial bases, learns that

$$A_N = (S_N \otimes I)(\Pi_N^{-1}\Sigma_N)(S_N^{-1} \otimes I), \quad (20)$$

where the matrix $S_N \in \mathbb{R}^{(N+1)\times(N+1)}$ maps coefficients of a polynomial of degree N in the Chebyshev basis

$$\left\{ T_i\left(2\frac{t}{\tau_m} + 1\right) :\ i = 0, \ldots, N \right\} \quad (21)$$

onto the corresponding coefficients in the Lagrange basis,

$$\{l_{N,i}(t) :\ i = 1, \ldots, N+1\},$$

defined on the grid (13). The relation (20) leads to an alternative formulation of the transfer function $\gamma_N(s)$, as expressed in the following theorem.

Theorem 3: If the grid points in the spectral discretization of (5) are chosen as (13), then we can express

$$\gamma_N(s) = F_N\left(sG_N - I\right)^{-1} H_N + D, \quad (22)$$

where

$$G_N = \Sigma_N^{-1}\Pi_N, \quad (23)$$

$$H_N = \begin{bmatrix} R_0^{-1}\left(I - \frac{\tau_m}{2}R_1\right)R_0^{-1}B \\ \frac{\tau_m}{2}R_0^{-1}B \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (24)$$

and

$$F_N = [CR_0\ CR_1\ \ \cdots\ CR_N], \quad (25)$$

with Σ_N and Π_N defined by (19) and (18).

The following property of the matrices in (22) plays an important role in the next paragraphs.

Proposition 4: Assume that $N_1, N_2 \in \mathbb{N}$ with $N_1 < N_2$. Then the matrices $\Sigma_{N_1}, \Pi_{N_1}, F_{N_1}, H_{N_1}$ in Theorem 3 are submatrices of $\Sigma_{N_2}, \Pi_{N_2}, F_{N_2}, H_{N_2}$.

Proposition 4 allows an adaptive construction of the approximation. An increase of the number of grid points, N, can be dealt with by extending the corresponding matrices..

B. Dynamic construction of a Krylov space

The model reduction technique presented in the paper is based on projecting the large and sparse matrices F_N, G_N and H_N, defined in Theorem 3, on an appropriately defined subspace. Instrumental to this we use the dynamic construction of a Krylov space of G_N, presented in [13]. This construction on its turn is inspired by methods for polynomial eigenvalue problems that exploit structure to reduce the storage cost of the Krylov vectors [4], [9], [14]. In what follows we summarize this construction (in a slightly adapted form). In §III-C, we derive reduced models based on a projection on the Krylov subspace.

We fix integers k and N, satisfying $1 \leq k \leq N$ and consider the Krylov space

$$\mathcal{K}_k(G_N, b) := \operatorname{span}\{b, G_N b, \ldots, G_N^{k-1} b\}. \quad (26)$$

Due to the special structure of G_N the matrix vector product satisfies the following property [13, Theorem 3.1].

Theorem 5: Assume that $(\sum_{i=0}^{m} A_i)$ is non-singular. Let $i, N \in \mathbb{N}$ with $i \leq N$ and let $Y \in \mathbb{C}^{n\times i}$. Then

$$G_N \operatorname{vec}(Y, 0, 0, \ldots, 0) = \operatorname{vec}(\hat{x}, Z, 0, \ldots, 0),$$

where $Z \in \mathbb{C}^{n\times i}$ is given by

$$Z = YL_i^T, \quad (27)$$

with

$$L_i = \frac{\tau_m}{4} \begin{bmatrix} 2 & 0 & -1 & & & \\ & \frac{1}{2} & 0 & -\frac{1}{2} & & \\ & & \frac{1}{3} & 0 & \ddots & \\ & & & \frac{1}{4} & \ddots & -\frac{1}{i-2} \\ & & & & \ddots & 0 \\ & & & & & \frac{1}{i} \end{bmatrix} \in \mathbb{R}^{i\times i},$$

and $\hat{x} \in \mathbb{C}^{n\times 1}$ satisfies

$$\hat{x} = R_0^{-1}\left(\sum_{i=0}^{k-1} y_i - A_0 \sum_{i=0}^{k-1} z_i - \sum_{j=1}^{m} A_j \left(\sum_{i=0}^{k-1} T_{i+1}(1 - 2\frac{\tau_j}{\tau_m}) z_i\right)\right). \quad (28)$$

Theorem 5 states in a precise way that a vector having all but the in first elements equal to zero is mapped by G_N on a vector having all but the $(i+1)n$ first elements equal to zero. This property is a consequence of the *companion-like* structure of (23).

If we assume that the vector b in (26) has the structure

$$b = [x_0^T \ 0 \ \cdots 0]^T, \quad x_0 \in \mathbb{C}^{n\times 1}, \tag{29}$$

then the vectors $G_N b, \ldots, G_N^{k-1} b$ only have their first $2n$, $3n$, $\ldots$, kn elements different from zero. The Arnoldi algorithm builds the Krylov sequence vector by vector, where, in addition, the vectors are orthogonalized. In step i, the orthogonalization is a linear combination of the $(i+1)$st vector and the previously computed i vectors. Hence, the orthogonalization at the ith iteration does not change the general structure of the $(i+1)$st vector. A dynamic implementation of Arnoldi to construct a basis of (26) with the starting vector (29), where this property is exploited, is described in Algorithm 6. In the description we use notation common for Arnoldi iterations: we let $\underline{\mathcal{H}}_i \in \mathbb{C}^{(i+1)\times i}$ denote the dynamically constructed rectangular Hessenberg matrix and $\mathcal{H}_i \in \mathbb{C}^{i\times i}$ the corresponding $i\times i$ upper part. To simplify the notation we will further denote the Krylov space (26) with starting vector (29) by $\mathcal{K}_k(G_N, x_0)$.

Algorithm 6: **Require:** $k \geq 1$, $x_0 \in \mathbb{C}^{n\times 1}$.

1: Let $v_1 = x_0/\|x_0\|_2$, $V_1 = v_1$, $\underline{\mathcal{H}}_0 =$ empty matrix
2: Factorize $R_0 = \sum_{i=0}^{m} A_i$
3: **for** $i = 1, 2, \ldots, k$ **do**
4: Let $\mathrm{vec}(Y) = v_i$
5: Compute Z according to (27) with sparse L_i
6: Compute $\hat{x}$ according to (28) using the factorization of the inverse computed in Step 2
7: Expand V_i with one block row (zeros)
8: Let $w_i := \mathrm{vec}(\hat{x}, Z)$, compute $h_i = V_i^* w_i$ and then $\hat{w}_i = w_i - V_i h_i$
9: Compute $\beta_i = \|\hat{w}_i\|_2$ and let $v_{i+1} = \hat{w}_i/\beta_i$
10: Let $\underline{\mathcal{H}}_i = \begin{bmatrix} \underline{\mathcal{H}}_{i-1} & h_i \\ 0 & \beta_i \end{bmatrix} \in \mathbb{C}^{(i+1)\times i}$
11: Expand V_i into $V_{i+1} = [V_i, v_{i+1}]$
12: **end for**

Output: $\mathcal{H}_k$, $\underline{\mathcal{H}}_k$ V_k, V_{k+1} v_{k+1}, basis $\mathcal{K}_k(G_N, x_0)$, with $N \geq k$, by extending V_k with $(N+1-k)n$ zero rows.

C. Projection and moment matching properties

We now arrive at the derivation of an approximation of $\gamma_N(s)$, defined by (10) or, equivalently, (22), having a prescribed order k. An approach to do so consists of constructing the Krylov space $\mathcal{K}_k(G_N, x_0)$ by Algorithm 6 and projecting the matrices F_N, G_N, H_N, defined in Theorem 3, on this Krylov space. Assuming $k \leq N$, an orthogonal projection on $\mathcal{K}_k(G_N, x_0)$ yields the following approximation of $\gamma_N(s)$:

$$\gamma^{(k)}(s) := F^{(k)} \ (sG^{(k)} - I)^{-1} \ H^{(k)} + D, \tag{30}$$

where

$$\begin{aligned} F^{(k)} &= F_{k-1} \ V_k, \\ G^{(k)} &= \mathcal{H}_k, \\ H^{(k)} &= V_k^T \ H_{k-1}, \end{aligned} \tag{31}$$

and the matrices

$$V_k = [v_1 \ \cdots \ v_k] \in \mathbb{R}^{kn\times k}, \mathcal{H}_k \in \mathbb{R}^{k\times k}$$

refer to the output of Algorithm 6. It is important to note that the matrices $F^{(k)}$ and $H^{(k)}$ are *submatrices* of $F^{(k+1)}$ and $H^{(k+1)}$. Therefore, they can be constructed in a dynamic way when doing iterations of Algorithm 6, as is the case with the Hessenberg matrix $\mathcal{H}_k$.

With a particular choice of the vector x_0, the transfer function (30) satisfies the following moment matching property with the (original) transfer function (2) of the time-delay system (1).

Theorem 7: Let $N, k \in \mathbb{N}$ with $N \geq k \geq 2$, and let $V_k \in \mathbb{R}^{kn\times k}$. Assume that the columns of the matrix V_k, possibly extended with zero rows, form an orthogonal basis of $\mathcal{K}_k(G_N, R_0^{-1}B)$. Then the transfer function (30) satisfies

$$\left.\frac{d^i \gamma^{(k)}(s)}{ds^i}\right|_{s=0} = \left.\frac{d^i \gamma(s)}{ds^i}\right|_{s=0}, \quad i = 0, \ldots, k-2 \tag{32}$$

and

$$\left.\frac{d^i \gamma^{(k)}(s^{-1})}{ds^i}\right|_{s=0} = \left.\frac{d^i \gamma(s^{-1})}{ds^i}\right|_{s=0}, \quad i = 0, 1. \tag{33}$$

Proof: The proof is performed in two steps. First, we observe that

$$G_N^{-1} H_N = \left[(R_0^{-1}B)^T \ 0 \ \cdots \ 0\right]^T,$$

from which we conclude

$$\begin{aligned} &\mathcal{K}_k(G_N, R_0^{-1}B) \\ &\quad = \mathrm{span}\left\{G_N^{-1}H_N, H_N, G_N H_N, \ldots, G^{k-2} H_N\right\}. \end{aligned} \tag{34}$$

It follows that the transfer function $\gamma^{(k)}(s)$ matches $k-1$ moments at zero and two at infinity with the transfer function $\gamma_N(s)$. Second, this moment matching property carries over to the transfer function $\gamma(s)$ of the delay equation by Theorem 1. ■

The principle behind the proof of Theorem 7, along with an overview of the results obtained so far, is shown in Figure 1.

To conclude the section, we describe the overall model reduction approach in Algorithm 8.

Algorithm 8: 1: Apply Algorithm 6 with $x_0 = R_0^{-1}B$ and construct $G^{(k)} = \mathcal{H}_k$. At the same time dynamically construct $F^{(k)}$ and $H^{(k)}$, defined in (31).

Output: Matrices $(F^{(k)}, G^{(k)}, H^{(k)}, D)$ of the reduced model;

$$\gamma^{(k)}(s) = F^{(k)}(sG^{(k)} - I)^{-1} H^{(k)} + D.$$

IV. Application

We study a heated rod which is cooled using delayed feedback. The algorithm of this paper can be used in such a study by computing an accurate dynamical system of small dimension which does not involve a delay. This is here achieved by first discretizing a heat equation in space and applying the model reduction algorithm to the discretized problem, which is a time-delay system of the form (1).

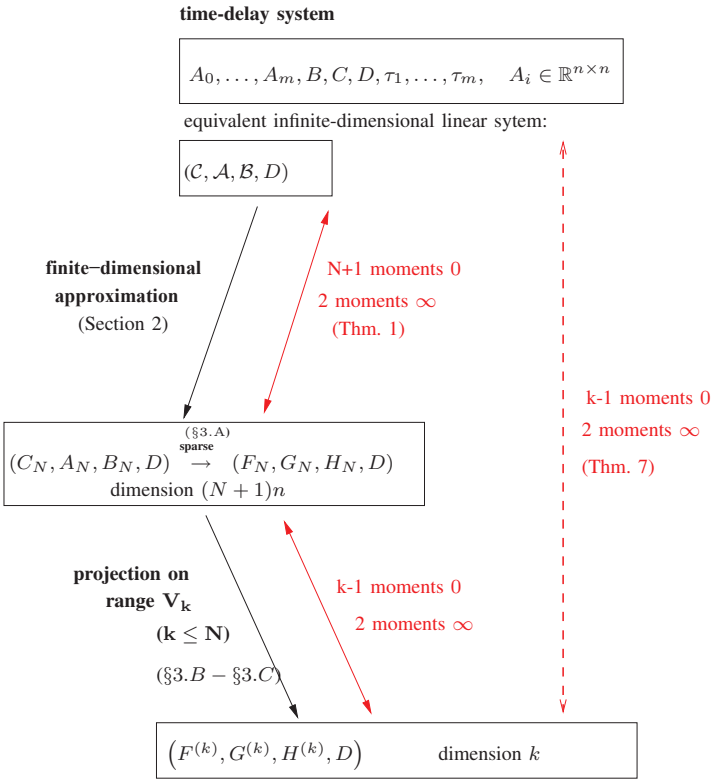

Fig. 1. Overview of Sections II-III. The results on moment matching are shown in red.

The physical model of the heated rod which we consider in this example is given by the partial differential equation

$$\frac{\partial v(x,t)}{\partial t} = \frac{\partial^2 v(x,t)}{\partial x^2} + a_0(x)v(x,t) + a_1(x)v(\pi - x, t-1), \tag{35}$$

with $a_0(x) = -2\sin(x)$, $a_1(x) = 2\sin(x)$ and $v(0,t) = v(\pi,t) = 0$, also used in [13]. The equation is a variant of [6, Example C] and can be interpreted as the heat equation corresponding to a rod with a distributed heating source and a non-local weighted delayed feedback. We discretize the differential equation (35) in space such that the corresponding time-delay system is of dimension n and fit it with the output matrix $C = (1, 1, \ldots, 1)/\|(1, 1 \ldots, 1)\|_2$, i.e., the output is the average temperature of the rod. The model reduction algorithm is applied to the system for

$$B_2 := e_{n/5},$$

i.e., the control $Bu(t)$ is localized at position $x = \pi/5$.

The corresponding time-delay systems of dimension $n = 100$ are now reduced by Algorithm 8. As expected from the theory, the reduced model is accurate both in terms of frequency response and in terms of approximation of the characteristic roots of the system. We illustrate this with figures, which should be interpreted as follows.

The frequency response and the point-wise error in the frequency response are plotted in Figure 2 for $k = 20$, with k the dimension of the reduced model. We clearly see that the approximation is good at $\omega = 0$ where many moments are matched and also decreasing for large frequencies, due to the matching moments at infinity. In Figure 3 we depict the maximum error

$$\sup_{\omega \in [0,\ \infty)} \left| \gamma(j\omega) - \gamma^{(k)}(j\omega) \right| \tag{36}$$

as a function of k. This illustrates the dynamic aspect of the method: the approximation error can always be reduced by increasing the number of iterations of Algorithms 1-2 (i.e., the dimension of the reduced model), while no lower bound is present due to the finite precision of a discretization (this would be the case for a `discretize+reduce` approach). Recall in this context that the application of Algorithm 2 does *nowhere* involve a choice of N. In fact, the discretized system with matrices (C_N, A_N, B_N, D) in the left of Figure 1 only serves as an auxiliary result in the technical derivation, whereas the reduced model can be seen directly as a projection of the infinite-dimensional linear system *equivalent* to the original time-delay system. The latter is indicated with the dashed arrow in the right of Figure 1.

Because the model reduction approach has a connection with a spectral discretization of the linear infinite-dimensional system (5) and Algorithm 6 can be interpreted as an Arnoldi algorithm applied to the operator $\mathcal{A}^{-1}$, we also expect the relevant eigenvalues of the reduced system

to be good approximations of the characteristic roots of the original time-delay system. In Figure 4 we observe that many roots close to the origin are well approximated in the reduced model. In Figure 5 we illustrate the convergence of the first eigenvalues of $G^{(k)}$. The convergence is exponential, which is in accordance with an Arnoldi method.

Finally we note that the rightmost eigenvalues of $(G^{(k)})^{-1}$ coincide with the rightmost characteristic roots, see Figure 4. As a consequence, the reduced models preserve the stability of the system.

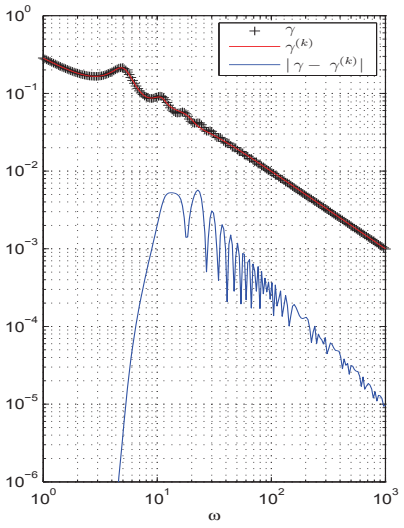

Fig. 2. Frequency response and error. The reduced model is of dimension $k = 20$.

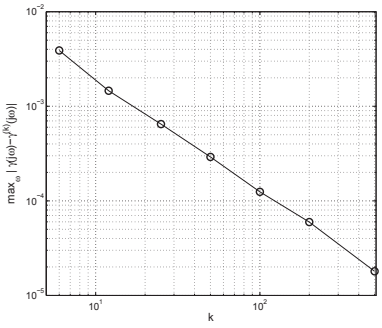

Fig. 3. Maximum error (36) as a function of k.

V. Conclusions

We proposed an approach for deriving reduced order models for infinite-dimensional time-delay systems, which are accurate in terms of the location of the rightmost characteristic roots, imposed by an underlying Arnoldi process, and in terms of approximating the transfer function, as moments at zero and at infinity are matched. As an important aspect, the matrices of the reduced system do not depend on N, the number of discretization points, which allows an interpretation of the reduced system as a projection of the infinite-dimensional system equivalent to the original time-delay system, see Figure 3. Moreover, the implementation is dynamic in the sense that the accuracy of the reduced model can always be increased by resuming the iterations of Algorithm 8.

The approach of the paper can be extended to systems with multiple inputs and multiple outputs provided that 'block versions' of the algorithms are used (in particular, block Arnoldi in Algorithm 6).

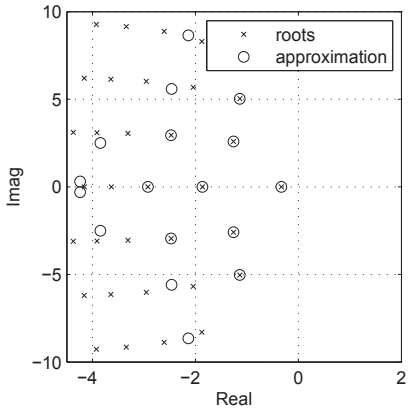

Fig. 4. Characteristic roots and approximations (eigenvalues of $\left(G^{(k)}\right)^{-1}$) for the example in Section IV. The dimension k of the reduced model is equal to 20.

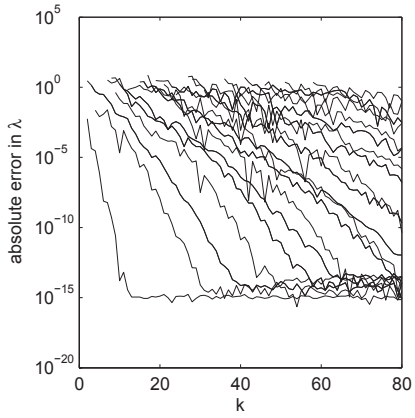

Fig. 5. The exponential convergence of the first characteristic roots.

Acknowledgment

This article presents results of the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture, the Optimization in Engineering Centre OPTEC, the project STRT1-09/33 of the K.U.Leuven Research Foundation, and the project G.0712.11N of the FWO.

REFERENCES

[1] A.C. Antoulas. *Approximation of large-scale dynamical systems*. SIAM, 2005.

[2] A.C. Antoulas, C.A. Beattie, and S. Gugercin. Interpolatory model reduction of large-scale dynamical systems. In J. Mohammadpour and K.M. Grigoriadis, editors, *Efficient Modeling and Control of Large-Scale Systems*, pages 3–58. Springer-Verlag, 2010.

[3] Z. Bai and R. W. Freund. A partial Padé-via-Lanczos method for reduced-order modeling. *Linear Alg. Appl.*, 332–334:139–164, 2001.

[4] Z. Bai and Y. Su. Dimension reduction of second-order dynamical systems via a second-order Arnoldi method. *SIAM J. Matrix Anal. and Applic.*, 26(5):1692–1709, 2005.

[5] D. Breda, S. Maset, and R. Vermiglio. Pseudospectral differencing methods for characteristic roots of delay differential equations. *SIAM J. Sci. Comput.*, 27(2):482–495, 2005.

[6] D. Breda, S. Maset, and R. Vermiglio. Numerical approximation of characteristic values of partial retarded functional differential equations. *Numer. Math.*, 113(2):181–242, 2009. doi:10.1007/s00211-009-0233-7.

[7] R. F. Curtain and H. Zwart. *An introduction to infinite-dimensional linear systems theory*. Springer-Verlag, NY, 1995.

[8] P. Feldman and R. W. Freund. Efficient linear circuit analysis by Padé approximation via the Lanczos process. *IEEE Trans. Computer-Aided Design*, CAD-14:639–649, 1995.

[9] R. W. Freund. Subspaces associated with higher-order linear dynamical systems. *BIT*, 45:495–516, 2005.

[10] K. Gallivan, E. Grimme, and P. Van Dooren. A rational Lanczos algorithm for model reduction. *Numerical Algorithms*, 12:33–63, 1996.

[11] E. Grimme, D. Sorensen, and P. Van Dooren. Model reduction of state space systems via an implicitly restarted Lanczos method. *Numerical Algorithms*, 12:1–31, 1996.

[12] S. Gugercin. An iterative SVD-Krylov based method for model reduction of large-scale dynamical systems. *Linear Alg. Appl.*, 428:1964–1986, 2008.

[13] E. Jarlebring, K. Meerbergen, and W. Michiels. A Krylov method for the delay eigenvalue problem. *SIAM Journal on Scientific Computing*, 32(6):3278–3300, 2010.

[14] K. Meerbergen. The quadratic Arnoldi method for the solution of the quadratic eigenvalue problem. *SIAM J. Matrix Anal. Appl.*, 30(4):1463–1482, 2008.

[15] W. Michiels, E. Jarlebring, and K. Meerbergen. Krylov based model order reduction of time-delay systems. TW Report 568, Department of Computer Science, Katholieke Universiteit Leuven, Belgium, 2010.

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

[17] J. Partington. Model reduction of delay systems. In V. Blondel and A. Megretski, editors, *Unsolved Problems in Mathematical Systems and Control Theory*, pages 29–32. Princeton university press, 2004.

[18] V. Simoncini and F. Perotti. On the numerical solution of $(\lambda^2 A + \lambda B + C)x = b$ and application to structural dynamics. *SIAM J. on Sci. Comput.*, 23(6):1875–1897, 2002.

[19] L. N. Trefethen. *Spectral Methods in MATLAB*. SIAM Publications, Philadelphia, 2000.