

# Approximation of distributed delays

Hao Lu<sup>1</sup>, Michael Di Loreto<sup>\*,1</sup>, Damien Eberard, Jean-Pierre Simon

*Laboratoire Ampère, UMR CNRS 5005, INSA-Lyon, 20 Avenue Albert Einstein, 69621  
Villeurbanne, France*

---

## Abstract

We address in this paper the approximation problem of distributed delays. Such elements are convolution operators with kernel having bounded support, and appear in the control of time-delay systems. From the rich literature on this topic, we propose a general methodology to achieve such an approximation. For this, we enclose the approximation problem in the graph topology, and work with the norm defined over the convolution Banach algebra. The class of rational approximates is described, and a constructive approximation is proposed. Analysis in time and frequency domains is provided. This methodology is illustrated on the stabilization control problem, for which simulations results show the effectiveness of the proposed methodology.

*Key words:* distributed delay, time-delay system, rational, approximation, lumped system, frequency analysis, numerical implementation, stabilization

---

## 1. Introduction

The interest for the use of distributed delays in the stabilization of time-delay systems appears in the pioneering work of Olbrot [1]. To generalize algebraic methods issued from linear systems in finite dimensional spaces to

---

\*Corresponding author

*Email addresses:* hao.lu@insa-lyon.fr (Hao Lu), michael.di-loreto@insa-lyon.fr (Michael Di Loreto), damien.eberard@insa-lyon.fr (Damien Eberard), jean-pierre.simon@insa-lyon.fr (Jean-Pierre Simon)

<sup>1</sup>The authors acknowledge the financial support of the CSC Scholarship from Xi'an Northwestern Polytechnical University (China), and the scholarship program with INSA-Lyon (France).

time-delay systems, Kamen et al. [2] first introduce a general mathematical setting for the control, and in particular for the stabilization, of time-delay systems. This mathematical framework was formalized in [3] by the introduction of the Bézout ring of pseudopolynomials, or independently by [4] with a behavioral approach. In these works, distributed delays appear in the stabilization problem. More generally, they are at the core in spectrum assignment by feedback [5]. Finite spectrum assignment generalizes the principle of Smith’s predictor for dead-time systems [6] to more general systems with delays, which can be stable or unstable. Distributed delays appear also in the characterization of equivalence transformations [7]. Robustness issue and optimization [8], robustness for input-delay systems [9] or finite time control [10] are other topics where distributed delays play a central role.

A distributed delay is a linear input-output convolution operator of the form

$$y(t) = (f * u)(t) = \int_0^{\vartheta} f(\tau)u(t - \tau) d\tau \quad (1)$$

where  $\vartheta$  is a strictly positive and finite real, and kernel  $f(\cdot)$  is a continuous function with support  $[0, \vartheta]$ . Numerical implementation of distributed delay was early investigated, to propose effective algorithms for control. Such an implementation was the starting point of a widely research activity. A first proposition for approximation with finite dimensional systems was proposed in [11]. Reduction and approximation of delay systems, involving lumped delays, were also investigated in [12]. In the work of [13], the authors propose a numerical integral approximation to realize an operator like in (1). Such an approximation writes as a sum of lumped delayed distributions, and unfortunately introduces additional closed-loop poles, and also instability phenomena. See, *e.g.* [14], [15], and references therein. To overcome this problem, various solutions were proposed. In [16], it was outlined that such an approximation has a poor accuracy and an high sensitivity for high frequencies. Hence, the author proposed to add a low-pass filter in the integral approximation. Such a solution was also proposed independently by [17]. Further implementation improvements were proposed in [18], with rational approximation and extension of bilinear transformations. These last papers give positive answers to the open problem of general approximation of distributed delays, outlined in [19].

We propose in this paper a general methodology for numerical implementation of distributed delays. The numerical implementation of an operator

exhibit two sides. The first one involves time discretization of the input-output behavior of the operator. Any distributed delay is a BIBO-stable operator. It is easy to verify that an equivalent discrete time system can be obtained by usual sampling tools, and that this system can be put into a sum of causal lumped distributions. Taking an appropriate sampling period, this equivalent discrete time system can always be made BIBO-stable. Hence, this first part presents no difficulty. We refer for instance to [20] and in references therein for more details on this part. The second part involves the approximation problem in continuous time of such an operator. This is this part we address in this paper. A continuous time approximation need to reproduce with high fidelity the internal dynamics of this operator, for large classes of input signals, but also to generate an arbitrarily close input-output behavior to the original one. For linear systems governed by a convolution, these notions are equivalent to impose simultaneous time and frequency approximations. For this, we explicitly separate two notions, namely input-output approximation and kernel approximation. We will explain why input-output approximation is not suitable for approximation of distributed delays, and we will then focus ourselves on kernel approximation. Kernel approximation can be realized in many different ways. Among most used classes of operators for approximation, we can cite polynomials, rational fractions, or exponentials. See, *e.g.* [21], [22], or [23].

With the objective to substitute the distributed delay by a more tractable system, highlighting rational assumptions, we propose two classes to realize approximation, namely lumped systems and a subclass of distributed delays. With previous objectives, we enclose the approximation problem into the Wiener algebra of BIBO-stable systems, using the graph topology. This corresponds to the weakest topology where feedback is a robust property. Moreover, for stable systems, graph topology and norm topology being the same, we work on norm convergence over this algebra, which is a Banach algebra. This general framework was used first in [24] for approximation of distributed parameter systems by lumped systems. Roughly speaking, working over this algebra, an approximation comes down to realize approximation of the kernel over the Banach algebra  $\mathcal{L}_1(\mathbb{R}_+)$ . This idea also grew in [25], for the approximation of lumped delayed distributions which appear in optimal control. Here, we propose an extension of the classes of approximates, and we show that working in this general setting yields to approximate a distributed delay in both time and frequency domains, for large classes of

input signals. We also propose to highlight the propositions made in [16] or [17], and to bring a mathematical foundation for this solution.

The paper is organized as follows. In Section 2, we define and fully characterize the main properties, in both time and frequency domains, of distributed delays. We characterize in particular a general decomposition of distributed delays on the so-called elementary distributed delays. In Section 3, we explicit our approximation problem, and solve it. This section starts with general comments on approximation, where graph topology is briefly recalled, and where we show that any distributed delay can be approximated over this topology by lumped systems. Then, we move to another class of approximates, using a subclass of distributed delays, easy to implement with stability. We show a density property of this subclass, and analyze the convergence of the approximation which is proposed. Section 4 relates the properties of such an approximation. We outline a proposal for a constructive approximation, and analyze time and frequency properties. A few simulation show the effectiveness of the method, on the stabilization control problem.

## 2. Convolution operators and distributed delays

### 2.1. Convolution algebra

An input-output causal convolution system is a dynamical system described by an equation of the form

$$y(t) = (f * u)(t) = \int_0^t f(\tau)u(t - \tau) d\tau \quad (2)$$

where  $y(\cdot)$ ,  $u(\cdot)$  and  $f(\cdot)$  are said to be the output, input and kernel of the map, respectively. Convolution systems are naturally defined over a commutative algebra since they are closed under addition, multiplication and scalar multiplication, operations that correspond to arbitrary series and parallel interconnections of such systems. For a normed algebra, it is quite interesting to obtain some closure properties for convergence. Hence, we prefer to work on a Banach algebra, that is an algebra for which any convergent sequence of elements in the algebra has a limit in the algebra, and for which the norm has the multiplicativity property. It will be assumed that all linear spaces and algebras are over the complex field. A general algebra of distributions

including a wide class of convolution systems is given by the so-called Callier and Desoer algebra, denoted  $\mathcal{A}$  [26]. We say that  $f \in \mathcal{A}$  if

$$f(t) = \begin{cases} f_a(t) + f_{pa}(t), & t \geq 0 \\ 0, & t < 0 \end{cases} \quad (3)$$

where the complex-valued function  $f_a(\cdot) \in \mathcal{L}_1(\mathbb{R}_+)$ , that is  $f_a$  is a complex valued function, locally integrable on  $\mathbb{R}_+$ , and such that  $\int_0^\infty |f_a(t)| dt < \infty$ . The complex-valued distribution  $f_{pa}$  stands for the purely atomic part and writes

$$f_{pa}(t) = \sum_{n=0}^{\infty} f_n \delta(t - t_n), \quad (4)$$

with  $f_n \in \mathbb{C}$ ,  $n = 0, 1, \dots$ ,  $0 = t_0 < t_1 < t_2 < \dots$ ,  $\delta(t - t_n)$  denotes the Dirac delta distribution centered in  $t_n$ , and  $\sum_{n \geq 0} |f_n| < \infty$ . As shown in Desoer and Vidyasagar [27], it is well known that  $\mathcal{A}$  is a commutative convolution Banach algebra with norm defined by

$$\|f\|_{\mathcal{A}} = \|f_a\|_{\mathcal{L}_1} + \sum_{n=0}^{\infty} |f_n|, \quad (5)$$

and with unit element the Dirac delta distribution  $\delta$ . Denoting  $\hat{f}$  the Laplace transform of  $f$ ,  $\hat{\mathcal{A}}$  denotes the set of Laplace transforms of elements in  $\mathcal{A}$ . The set  $\hat{\mathcal{A}}$  is also a commutative Banach algebra with unit element under pointwise addition and multiplication, for the norm

$$\|\hat{f}\|_{\hat{\mathcal{A}}} = \|f\|_{\mathcal{A}}, \quad \forall f \in \mathcal{A}.$$

The algebra  $\mathcal{A}$  gives a general mathematical framework for the analysis of distributed delays, as it will be explained in the next subsection. Before this, let us recall the concept of Bounded Input-Bounded Output stability [27].

**Definition 1.** A convolution system in the form (2) is said to be BIBO stable if  $f \in \mathcal{A}$ .

More generally, we may be interested by  $\mathcal{L}_p$ -stability, for  $1 \leq p \leq \infty$ . We say that (2) is  $\mathcal{L}_p$ -stable if for any  $u(\cdot) \in \mathcal{L}_p(\mathbb{R}_+)$ , that is  $u(\cdot)$  locally integrable over  $\mathbb{R}_+$  and

$$\|u\|_{\mathcal{L}_p} = \left( \int_{\mathbb{R}_+} |u(t)|^p dt \right)^{1/p} < \infty,$$

the output  $y(\cdot)$  is also in  $\mathcal{L}_p(\mathbb{R}_+)$ . Since

$$\|y\|_{\mathcal{L}_p} = \|f * u\|_{\mathcal{L}_p} \leq \|f\|_{\mathcal{L}} \|u\|_{\mathcal{L}_p}, \quad (6)$$

$\mathcal{L}_p$ -stability is equivalent to BIBO stability [27].

## 2.2. Distributed delays

Let  $\mathbb{I}_{a,b} = [a, b]$  be the bounded closed interval in  $\mathbb{R}_+$ , for some reals  $a$  and  $b$ ,  $0 \leq a < b$ . Notations  $\mathbb{I}_{0,\infty}$  or  $\mathbb{R}_+$  stand for  $[0, \infty[$ . We define  $\mathcal{K}(\mathbb{I}_{a,b})$  as the set of complex valued functions  $g(\cdot)$  of the form

$$g(t) = \begin{cases} g_{\mathbb{I}_{a,b}}(t), & t \in \mathbb{I}_{a,b} \\ 0, & \text{elsewhere} \end{cases} \quad (7)$$

where

$$g_{\mathbb{I}_{a,b}}(t) = \sum_{i \geq 0} \sum_{j \geq 0} c_{ij} t^j e^{\lambda_i t}, \quad (8)$$

for some  $c_{ij}$  and  $\lambda_i$  in  $\mathbb{C}$ , and the sums are finite. In other words,  $g_{\mathbb{I}_{a,b}}$  is a finite linear combination of exponential-polynomials type functions, and it is in particular a continuous function. For any real valued function in  $\mathcal{K}(\mathbb{I}_{a,b})$ , if some  $\lambda_i \in \mathbb{C}$  appears in the sum, then so does its conjugate  $\bar{\lambda}_i$ , and the associated coefficients  $c_{ij}$  are complex conjugates. Hence, any real valued function in  $\mathcal{K}(\mathbb{I}_{a,b})$  is a function generated by real linear combinations of  $t^j e^{\sigma_i t}$ ,  $t^j e^{\sigma_i t} \sin(\beta_k t)$  and  $t^j e^{\sigma_i t} \cos(\beta_k t)$ , for some real numbers  $\sigma_i$ ,  $\beta_k$ , the sums being finite. The formal definition of distributed delay is made below.

**Definition 2.** A distributed delay is a causal convolution system, with kernel  $f$  in  $\mathcal{K}(\mathbb{I}_{\vartheta_1, \vartheta_2})$ , for some bounded real numbers  $0 \leq \vartheta_1 < \vartheta_2$ .

In other words, a distributed delay can be written like an input-output convolution operator of the form

$$y(t) = (f * u)(t) = \int_{\vartheta_1}^{\vartheta_2} f_{\mathbb{I}_{\vartheta_1, \vartheta_2}}(\tau) u(t - \tau) d\tau, \quad (9)$$

with notations introduced in (8). The set of distributed delays, denoted by  $\mathcal{G}$ , is a ring. Obviously, for real valued signals, the kernel  $f(\cdot)$  will be a real valued function in  $\mathcal{K}(\mathbb{I}_{\vartheta_1, \vartheta_2})$ . In the previous definition, we restrict ourselves to define a distributed delay like a convolution operator with kernel in  $\mathcal{K}(\mathbb{I}_{\vartheta_1, \vartheta_2})$ . This restriction is not conservative as we will show in Section 3.2.

Actually, all distributed delays which appeared in the literature are particular cases of this definition. This definition is based on a rational construction, as this appears explicitly using Laplace transforms. Any distributed delay  $\mathcal{G}$  admits a Laplace transform, corresponding to the finite Laplace transform of its kernel  $f \in \mathcal{K}(\mathbb{I}_{\vartheta_1, \vartheta_2})$ ,

$$\hat{y}(s) = \hat{f}(s)\hat{u}(s), \quad \hat{f}(s) = \int_{\vartheta_1}^{\vartheta_2} f_{\mathbb{I}_{\vartheta_1, \vartheta_2}}(\tau) e^{-s\tau} d\tau, \quad (10)$$

where  $\hat{f} \in \hat{\mathcal{G}}$  is an entire function, *i.e.* holomorphic on the whole complex plane. The notion of elementary distributed delay will greatly simplify the approximation problem. Let us define the complex valued function  $\theta_\lambda(\cdot) \in \mathcal{K}(\mathbb{I}_{0, \vartheta})$ , for some  $\lambda \in \mathbb{C}$  and  $\vartheta > 0$ , by

$$\theta_\lambda(t) = \begin{cases} e^{\lambda t}, & t \in [0, \vartheta] \\ 0, & \text{elsewhere} \end{cases} \quad (11)$$

and its Laplace transform

$$\hat{\theta}_\lambda(s) = \frac{1 - e^{-(s-\lambda)\vartheta}}{s - \lambda}, \quad (12)$$

which is an entire function even in  $s = \lambda$  where  $\hat{\theta}_\lambda(\lambda) = \vartheta$ . In other words,  $\lambda$  is a removable singularity, and consequently  $\hat{\theta}_\lambda(s)$  has no pole. The distributed delay whose kernel is  $\theta_\lambda$  is called an elementary distributed delay. The  $k$ th derivative  $\hat{\theta}_\lambda^{(k)}(s)$  of  $\hat{\theta}_\lambda(s)$  yields

$$\hat{\theta}_\lambda^{(k)}(s) = \int_0^\vartheta (-\tau)^k e^{-(s-\lambda)\tau} d\tau, \quad (13)$$

which is still in  $\hat{\mathcal{G}}$ , and corresponds to the Laplace transform of the function  $\theta_\lambda^k(t) = (-t)^k e^{\lambda t}$  for  $t \in [0, \vartheta]$ , and 0 elsewhere. From previous definitions, we can state the following lemma, which also appeared in [28] using 2D-polynomials, and which will play a central role for approximation.

**Lemma 1.** *Any element in  $\hat{\mathcal{G}}$  can be decomposed into a finite sum of Laplace transform of elementary distributed delays and its successive derivatives.*

PROOF. Take any element in  $\mathcal{G}$ . Its kernel  $g(\cdot)$  lies in  $\mathcal{K}(\mathbb{I}_{\vartheta_1, \vartheta_2})$ , and writes as in (8). By time translation corresponding to the lumped delay  $\vartheta_1$ , it is readily a linear finite combination of elementary distributed delays  $\theta_\lambda(t)$  and of the functions  $\theta_\lambda^k(t)$ , as defined in (11) and (13), with  $\vartheta = \vartheta_2 - \vartheta_1$ .  $\square$

In other terms, previous result tells us that for any  $\hat{g} \in \hat{\mathcal{G}}$ , there exist complex polynomials  $\hat{g}_{ik} \in \mathbb{C}[e^{-\vartheta s}]$  with respect to the variable  $e^{-\vartheta s}$  and  $\lambda_i \in \mathbb{C}$ , in finite number, such that

$$\hat{g}(s) = \sum_{i,k} \hat{g}_{ik}(e^{-\vartheta s}) \hat{\theta}_{\lambda_i}^{(k)}(s), \quad (14)$$

where successive derivatives are iteratively computed by

$$\hat{\theta}_{\lambda_i}^{(k)}(s) = (-1)^k k! \frac{1 - e^{-(s-\lambda_i)\vartheta} - \sum_{n=1}^k \frac{\vartheta^n}{n!} e^{-(s-\lambda_i)\vartheta} (s-\lambda_i)^n}{(s-\lambda_i)^{k+1}}, \quad (15)$$

and  $\hat{\theta}_{\lambda_i}^{(k)}(s)$  are still entire functions, since  $\hat{\theta}_{\lambda_i}^{(k)}(\lambda_i) = \frac{(-1)^k \vartheta^{k+1}}{k+1}$ , for any  $k \geq 0$ . From (14) and taking into account that practical distributed delay is a real valued operator, we see that, grouping the terms via least common multiple, any element  $\hat{g} \in \hat{\mathcal{G}}$  can be put into a fraction

$$\hat{g}(s) = \frac{n(s, e^{-\vartheta s})}{d(s)}, \quad (16)$$

where  $n(s, e^{-\vartheta s}) \in \mathbb{R}[s, e^{-\vartheta s}]$  is a real quasipolynomial with respect to the algebraically independent variables  $s$  and  $e^{-\vartheta s}$ , and  $d(s) \in \mathbb{R}[s]$ . Any element in the right hand side of (14) being an entire function,  $\hat{g}(s)$  is also an entire function. Hence, any zero of  $d(s)$  is also a zero of  $n(s, e^{-\vartheta s})$ . Furthermore, the degree with respect to  $s$  of  $\hat{\theta}_{\lambda}^{(k)}(s)$  in (12) and (15), for any  $k$ , being strictly negative, the degree with respect to  $s$  of  $\hat{g}(s)$  satisfies  $\deg_s n < \deg_s d$ , so that its Laplace transform is strictly proper. As in [3], the ring  $\hat{\mathcal{G}}$  is the ring of those Laplace transforms of distributed delays that are rational in the variable  $s$  and  $e^{-\vartheta s}$ , which are entire and strictly proper (with respect to  $s$ ). This result comes from the assumption on rational kernels in (8). The kernel of any element in  $\mathcal{G}$  is obviously in  $\mathcal{L}_1(\mathbb{R}_+)$ , so that from (6) and the definition of  $\mathcal{A}$ , any distributed delay is BIBO stable. Convolution of two kernels with finite support yields another kernel with finite support, and since exponential-polynomials type functions are closed under classical product, we have in fact that  $\mathcal{G}$  is a normed subalgebra of  $\mathcal{A}$  for the  $\|\cdot\|_{\mathcal{A}}$ -norm. From these definitions and properties, we are now able to formulate the approximation problem of distributed delays.



### 3. Approximation of distributed delays

For this purpose, we define two subspaces in direct sum in  $\mathcal{K}(\mathbb{I}_{\vartheta_1, \vartheta_2})$ , denoted respectively  $\mathcal{K}_s(\mathbb{I}_{\vartheta_1, \vartheta_2})$  and  $\mathcal{K}_u(\mathbb{I}_{\vartheta_1, \vartheta_2})$ , consisting of linear combinations of exponential-polynomials type functions on some finite interval as in (8), with  $\text{Re } \lambda_i < 0$  and  $\text{Re } \lambda_i \geq 0$ , respectively, for all  $i \geq 0$ .

From Lemma 1 and (14), we know that any element  $g \in \mathcal{K}_s(\mathbb{I}_{\vartheta_1, \vartheta_2})$  is a linear combination of elements  $\theta_{\lambda_i}^k$  in  $\mathcal{K}_s(\mathbb{I}_{\vartheta_1, \vartheta_2})$ , for which  $\text{Re } \lambda_i < 0$ . Hence, any element in  $\mathcal{K}_s(\mathbb{I}_{\vartheta_1, \vartheta_2})$  can be numerically implemented with stability using elementary blocks, as illustrated in Fig. 1 for  $\theta_\lambda(t)$ . Note that in practice, since  $\lambda \in \mathbb{C}$ , we should implement it using a real decomposition. Approximation for the subclass  $\mathcal{K}_s(\mathbb{I}_{\vartheta_1, \vartheta_2})$  of distributed delays is not required, since this implementation is realized with stability, and requires, from an implementation point of view, only two pointwise delays, namely 0 and  $\vartheta$ . However, this is no more true for elements in  $\mathcal{K}_u(\mathbb{I}_{\vartheta_1, \vartheta_2})$ , where numerical realization for  $\theta_\lambda(t)$  as in Fig. 1 yields to an unstable system. Note that another realization for  $\hat{\theta}_\lambda(s)$  may be

$$\dot{x}(t) = \lambda x(t) + u(t) - e^{\lambda\vartheta} u(t - \vartheta),$$

which is however still numerically unstable for  $\text{Re } \lambda \geq 0$ . Intuitively, this instability is a consequence of a non exact numerical cancellation of the removable singularity  $s = \lambda$  of  $\hat{\theta}_\lambda(s)$ . Therefore, we will focus our attention on distributed delays whose kernels are in  $\mathcal{K}_u(\mathbb{I}_{\vartheta_1, \vartheta_2})$ .

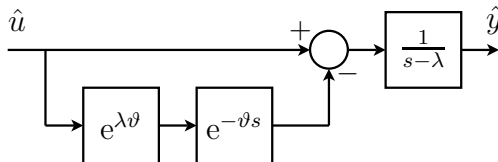


Figure 1: Realization of an element  $\hat{\theta}_\lambda(s) \in \mathcal{K}_s(\mathbb{I}_{0, \vartheta})$ , with  $\text{Re } \lambda < 0$ .

The question can now be stated as follows: How to define a continuous time approximation of a distributed delay whose kernel lies in  $\mathcal{K}_u(\mathbb{I}_{\vartheta_1, \vartheta_2})$ ? We give complete answers to this question in the next two subsections. We start by some general remarks on approximation over  $\mathcal{A}$ , and hence give answers to the previous question.

### 3.1. General comments on approximation

Any distributed delay with kernel  $f \in \mathcal{K}(\mathbb{I}_{\vartheta_1, \vartheta_2})$  is a convolution operator. Its  $\mathcal{L}_\infty$ -induced norm is

$$\|f\|_{\mathcal{A}} = \|f\|_{\mathcal{L}_1}.$$

From (6), we see that this norm is actually an upper bound of all its induced  $\mathcal{L}_p$ -norms, for  $1 \leq p \leq \infty$ . A natural metric is then obtained from the  $\mathcal{A}$ -norm, which in turn can be enclosed in the graph topology. Consider a distribution  $p$  with a coprime factorization  $(n, d)$  in  $\mathcal{A}$ , that is  $\hat{p} = \hat{n}\hat{d}^{-1}$ , and such that there exist  $x$  and  $y$  in  $\mathcal{A}$  satisfying

$$n * x + d * y = \delta, \quad (17)$$

or in the Laplace domain

$$\hat{n}\hat{x} + \hat{d}\hat{y} = 1, \quad \forall s \in \mathbb{C}, \operatorname{Re} s \geq 0. \quad (18)$$

A neighbourhood of  $p$  in the graph topology is the set of all plants of the form  $n_\Delta d_\Delta^{-1}$  where  $(n_\Delta, d_\Delta)$  belongs to some ball in  $\mathcal{A}$  centered at  $(n, d)$ . If  $p$  is itself in  $\mathcal{A}$ , for any positive number  $r$ , a neighbourhood of  $p$  in the graph topology writes

$$\mathcal{B}(p, r) = \{p_\Delta \in \mathcal{A} : \|p - p_\Delta\|_{\mathcal{A}} \leq r\}.$$

In [29, Ch. 7], it was shown that the graph topology is the weakest topology in which feedback stability is a robust property. Graph topology being metrizable and hence first-countable, for BIBO-stable systems, norm topology and graph topology are the same. Let  $\varepsilon$  be a given suitable small positive number. Hence, we say that a distribution  $p_\varepsilon$  in  $\mathcal{A}$  is an approximation of  $p$  in the graph topology if the element  $p_\varepsilon$  is close to  $p$ , that is  $p_\varepsilon \in \mathcal{B}(p, \varepsilon)$ , or equivalently  $\|p - p_\varepsilon\|_{\mathcal{A}} \leq \varepsilon$ . This property can be related to convergence in  $\mathcal{A}$ , but we need first to define the class of operators that will approximate  $p$  in the graph topology. The most commonly used class is the so-called class of lumped systems, which consists in convolution operators whose kernels lie in the set  $\mathcal{Q}$ , defined by

$$\mathcal{Q} = \{g \in \mathcal{A} : g(t) = g_0\delta(t) + g_a(t), \quad g_0 \in \mathbb{C}, \quad g_a \in \mathcal{K}_s(\mathbb{I}_{0, \infty})\}. \quad (19)$$

From [24], we know that any plant  $p$  with coprime factorization  $(n, d)$  over  $\mathcal{A}$  can be approximated by a lumped plant in the graph topology if and only if there exists a real constant matrix  $M$  of rank 1 such that

$$MS_{pa} = 0, \text{ where } S_{pa} = \begin{bmatrix} n_{pa} \\ d_{pa} \end{bmatrix}, \quad (20)$$

and  $(\cdot)_{pa}$  denotes the purely atomic part. Distributed delays having strictly proper transfer functions, we are interested for approximation with the class of strictly proper lumped systems, denoted  $\mathcal{Q}_s$ , set of elements in  $\mathcal{K}_s(\mathbb{I}_{0,\infty})$ . Such elements satisfy ordinary differential equations, and are easily simulated. From [23], we know that the closure of  $\mathcal{Q}_s$  is  $\mathcal{L}_1(\mathbb{R}_+)$ , since every function in  $\mathcal{L}_1(\mathbb{R}_+)$  can be approximated by a sum of exponentials. This yields the following result, that can be seen as a particular case of the above given condition (20).

**Theorem 1.** *Any plant in  $\mathcal{L}_1(\mathbb{R}_+)$  can be approximated by a lumped system in the graph topology.*

PROOF. Let  $f$  be an element in  $\mathcal{L}_1(\mathbb{R}_+)$ . A coprime factorization of  $f$  over  $\mathcal{A}$  is  $n = f$  and  $d = \delta$ . Since  $n_{pa} = 0$ , any real constant matrix  $M = [0 \ \star]$ ,  $(\star)$  denoting an arbitrary real number, satisfies (20). Hence, from [24, Th. 4.1],  $f$  can be approximated by a lumped system in the graph topology.  $\square$

Note that if the element  $f$  in  $\mathcal{L}_1(\mathbb{R}_+)$  is strictly proper, then an approximation in the graph topology realized by a lumped system will be in fact in the set  $\mathcal{Q}_s$ . Any convolution system whose kernel lies in  $\mathcal{L}_1(\mathbb{R}_+)$  can be approximated by a lumped system. Roughly speaking, since any distributed delay is a strictly proper fraction and is BIBO-stable, we know that it can be approximated in the graph topology by a strictly proper lumped system.

**Corollary 1.** *Any distributed delay can be approximated by a lumped system in the graph topology.*

Others approximation classes can be used to numerically realize a distributed delay. In the literature, appears a method based on a numerical integral approximation, but such a method was shown to be at the core of various undesired results, like instability or numerical sensitivity. Let us briefly make some

considerations on such approximations, introducing what we called input-output approximation, and let us show why such approximations are not suitable for our problem. For any distributed delay  $y(t) = (f * u)(t)$ , we say that  $y_{\text{app}}$  is an  $\mathcal{L}_p$  input-output approximation of  $y$ , if for some arbitrarily  $\varepsilon > 0$ ,  $\|y - y_{\text{app}}\|_{\mathcal{L}_p} \leq \varepsilon$  holds. For instance, consider the  $\mathcal{L}_\infty$  input-output approximation based on integral approximation, like Newton-Cotes quadrature methods, which leads to an approximation  $y_{\text{app}}(t)$  of the form

$$y_{\text{app}}(t) = \sum_{k=0}^q c_k f(\tau_k) u(t - \tau_k), \quad (21)$$

where reals  $c_k$  and  $\tau_k$  depend on the applied method [30, Ch. 4]. Various problems arise with this kind of approximation. First note that such an approximation is made for a given  $u(\cdot)$ . Hence, if such an input changes, the properties of this approximation failed, in general. Note also that, as mentioned by [16], the Laplace transform of (21) yields to an input-output approximation transfer function that is proper, since it contains only pointwise delays. From (14) and (16), a distributed having a strictly proper transfer function, this approximation has a poor accuracy for high frequencies, which will have as consequence a high sensitivity to high-frequency plant uncertainties, in particular for the closed-loop system. This fact is illustrated in Fig. 2, where a frequency diagram of  $\hat{\theta}_\lambda(s)$  and an input-output approximation is plotted. This negative result is directly interpretable in the graph topology by the following result.

**Theorem 2.** *Any distributed delay can not be approximated in the graph topology by a purely atomic distribution.*

PROOF. Consider an element in  $\mathcal{G}$  with kernel  $g$ . Then  $g \in \mathcal{L}_1(\mathbb{R}_+)$ . Consider an arbitrary purely atomic distribution with pointwise delays of the form

$$p(t) = \sum_{i \geq 0} p_i \delta(t - t_i)$$

where  $\sum_{i \geq 0} |p_i| < \infty$ . Such a distribution corresponds to the kernel of the approximation using numerical integral approximation like in (21). Then

$$\|g - p\|_{\mathcal{L}_1} = \|g\|_{\mathcal{L}_1} + \sum_{i \geq 0} |p_i|.$$

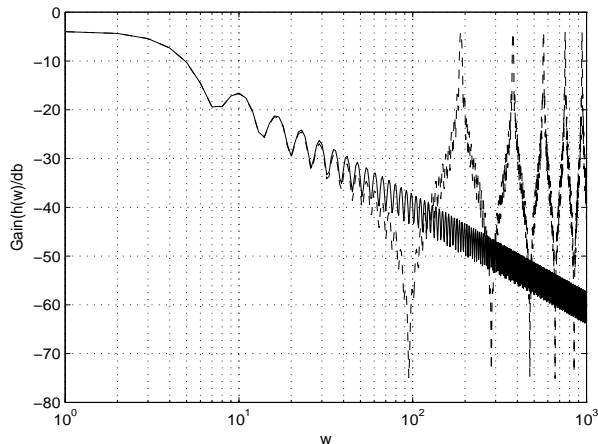


Figure 2: Bode gain diagram of  $\hat{\theta}_1(s)$  and an approximation based on integral approximation (trapezoidal Newton-Cotes method with 30 nodes).

Hence  $\|g - p\|_{\mathcal{A}} \geq \|g\|_{\mathcal{L}_1}$ , and consequently  $p$  can not approximate  $g$  in the graph topology.  $\square$

It is clear that input-output approximation is not suitable for numerical implementation. This was illustrated in [13], [14], or [16]. Hence we focus our attention to the notion of approximation defined in the graph topology. This approximation is equivalent to realize an approximation of the kernel of the convolution, so we may also call it kernel approximation, to discern with respect to input-output approximation.

### 3.2. Kernel approximations

Approximation in the graph topology for an element  $f$  in  $\mathcal{K}_s(\mathbb{I}_{\vartheta_1, \vartheta_2})$  requires to find, for some given  $\varepsilon > 0$ , an element  $f_{\text{app}}$ , defined in general in  $\mathcal{A}$ , such that  $f_{\text{app}} \in \mathcal{B}(f, \varepsilon)$ . Element  $f$  being in particular in  $\mathcal{L}_1(\mathbb{R}_+)$ , adding purely atomic part in approximation can not improve approximation, so we realize it over the Banach algebra  $\mathcal{L}_1(\mathbb{R}_+)$ . Indeed, if an approximation writes

$$f_{\text{app}} = f_a^{\text{app}} + f_{pa}^{\text{app}},$$

where  $f_a^{\text{app}}$  and  $f_{pa}^{\text{app}}$  stand for the atomic and purely atomic parts of  $f_{\text{app}}$ , respectively, the following decomposition holds

$$\|f - f_{\text{app}}\|_{\mathcal{A}} = \|f - f_a^{\text{app}}\|_{\mathcal{A}} + \|f_{pa}^{\text{app}}\|_{\mathcal{A}}.$$

The purely atomic part of the approximation is in an independent sum, and can then be reduced to zero to reduce the approximation error. We have seen that an approximation can be obtained by lumped systems, with rational and stable transfer functions. Since we want to obtain an approximation which is global with respect to the time, we are interested in uniform convergent sequence of functions in  $\mathcal{L}_1(\mathbb{R}_+)$  to  $f$ . Consider a sequence  $f_n$  of functions which uniformly converges to  $f$ , that is, for any  $\varepsilon > 0$ , there exists  $n$  in  $\mathbb{N}$  such that for any  $\kappa \geq n$ ,  $\|f - f_\kappa\|_{\mathcal{A}} \leq \varepsilon$ . So, for any  $\kappa \geq n$ ,  $f_\kappa$  is a kernel approximation of  $f$  in the graph topology. We propose next a method to approximate an element in  $\mathcal{K}_u(\mathbb{I}_{\vartheta_1, \vartheta_2})$  by elements in  $\mathcal{K}_s(\mathbb{I}_{\vartheta_1, \vartheta_2})$ . Since these last elements are easily implementable, this will give us an effective approximation. In other words, we approximate a distributed delay by another distributed delays. We start by considering the case of an elementary distributed delay.

**Lemma 2.** *Any distributed delay with kernel  $\theta_\lambda(\cdot)$  in  $\mathcal{K}_u(\mathbb{I}_{0, \vartheta})$  can be approximated by distributed delays with kernels in  $\mathcal{K}_s(\mathbb{I}_{0, \vartheta})$  for the graph topology.*

PROOF. Let  $\theta_\lambda(\cdot)$  be the kernel of an elementary distributed delay, with  $\text{Re } \lambda \geq 0$ . Let  $\mu \in \mathbb{C}$  such that  $\text{Re } \mu = \alpha < 0$ . We define the transform

$$\Theta_\lambda(s) = (\alpha s)^{-1} \theta_\lambda(-\alpha^{-1} \ln s), \quad e^{-\vartheta\alpha} \leq s \leq 1. \quad (22)$$

The function  $\Theta_\lambda(\cdot)$  is continuous over  $[e^{-\vartheta\alpha}, 1]$ , and we have

$$\|\Theta_\lambda\|_{\mathcal{L}_1} = \|\theta_\lambda\|_{\mathcal{A}}. \quad (23)$$

By Müntz-Szász theorem, we see that  $\Theta_\lambda(\cdot)$  can be approximated with respect to  $\mathcal{L}_1$ -norm, as closed as desired, by a function of the form

$$\Theta_{\lambda, \text{app}}(s) = (\alpha s)^{-1} \Psi_{\lambda, \text{app}}(s), \quad (24)$$

where  $\Psi_{\lambda, \text{app}}$  is a polynomial in  $s$ . Since norms are preserved in (23), this comes down to approximate  $\theta_\lambda$  as closely as we please, using inverse transform of (22), by a sum of exponentials of the form

$$\psi_{\lambda, \text{app}}(t) = \Psi_{\lambda, \text{app}}(e^{-\alpha t}),$$

which is clearly an element in  $\mathcal{K}_s(\mathbb{I}_{0, \vartheta})$ . □

In other words, an approximation of  $\theta_\lambda$  in  $\mathcal{A}$  can be a function of the form

$$\psi_{\lambda,n}(t) = \sum_{i=1}^n \gamma_{i,n} \theta_{\alpha_i}(t), \quad (25)$$

with  $\alpha_i \in \mathbb{C}$ ,  $\text{Re } \alpha_i < 0$  some arbitrarily complex numbers, and  $\gamma_{i,n}$  some suitable constants. The order of the approximation  $n$  describes the number of parallel distributed elements to be added, to get the desired approximation accuracy. Such an approximation does not increase the degree of the elementary distributed delay, since it is a sum of distributed delays which are strictly proper and have same degree than  $\hat{\theta}_\lambda(s)$ . From the synthesis of an approximation for the elementary distributed delay, we show in the next two results that it allows to construct an explicit approximation for any distributed delay.

**Lemma 3.** *Let  $\psi_{\lambda,n}$  be an element in  $\mathcal{K}_s(\mathbb{I}_{0,\vartheta}) \cap \mathcal{B}(\theta_\lambda, \varepsilon)$ , for a given  $\varepsilon > 0$ , and  $k$  in  $\mathbb{N}$ . Then  $\psi_{\lambda,n}^k \in \mathcal{K}_s(\mathbb{I}_{0,\vartheta}) \cap \mathcal{B}(\theta_\lambda^k, \vartheta^k \varepsilon)$ , where*

$$\psi_{\lambda,n}^k(t) = \begin{cases} (-t)^k \psi_{\lambda,n}(t), & t \in [0, \vartheta] \\ 0, & \text{elsewhere} \end{cases}.$$

PROOF. Let  $\varepsilon > 0$ , and take  $\psi_{\lambda,n}$  an element in  $\mathcal{K}_s(\mathbb{I}_{0,\vartheta}) \cap \mathcal{B}(\theta_\lambda, \varepsilon)$ , that is

$$\psi_{\lambda,n}(t) = \sum_{i=1}^n \gamma_{i,n} \theta_{\alpha_i}(t),$$

with  $\gamma_{i,n}$  in  $\mathbb{C}$ , and  $\|\theta_\lambda - \psi_{\lambda,n}\|_{\mathcal{A}} \leq \varepsilon$ . By Laplace transform and  $k$ th order differentiation, we get

$$\hat{\psi}_{\lambda,n}^{(k)}(s) = \sum_{i=1}^n \gamma_{i,n} \hat{\theta}_{\alpha_i}^{(k)}(s).$$

In the time domain, this last identity corresponds to

$$\psi_{\lambda,n}^k(t) = \sum_{i=1}^n \gamma_{i,n} \theta_{\alpha_i}^k(t) = (-t)^k \psi_{\lambda,n}(t).$$

Clearly  $\psi_{\lambda,n}^k \in \mathcal{K}_s(\mathbb{I}_{0,\vartheta})$ . Hence, it remains to show that  $\psi_{\lambda,n}^k \in \mathcal{B}(\theta_\lambda^k, \vartheta^k \varepsilon)$ . But  $\psi_{\lambda,n} \in \mathcal{B}(\theta_\lambda, \varepsilon)$ , so we have

$$\begin{aligned} \|\theta_\lambda^k - \psi_{\lambda,n}^k\|_{\mathcal{A}} &= \int_0^\vartheta |(-t)^k (\theta_\lambda(t) - \psi_{\lambda,n}(t))| dt \\ &\leq \vartheta^k \|\theta_\lambda - \psi_{\lambda,n}\|_{\mathcal{A}} \leq \vartheta^k \varepsilon, \end{aligned}$$

which completes the proof.  $\square$

For any distributed delay, we have the following general result.

**Theorem 3.** *Let  $\hat{g} \in \hat{\mathcal{G}}$  be an arbitrary distributed delay, of the form*

$$\hat{g}(s) = \sum_{i,k} \hat{g}_{ik}(e^{-\vartheta s}) \hat{\theta}_{\lambda_i}^{(k)}(s),$$

and let  $\varepsilon_i$  be given positive real numbers. For any elements  $\psi_{\lambda_i, n}$  in  $\mathcal{K}_s(\mathbb{I}_{0, \vartheta}) \cap \mathcal{B}(\theta_{\lambda_i}, \varepsilon_i)$ , we define

$$\hat{g}_{\text{app}}(s) = \sum_{i,k} \hat{g}_{ik}(e^{-\vartheta s}) \hat{\psi}_{\lambda_i, n}^{(k)}(s).$$

Then  $g_{\text{app}}$  lies in  $\mathcal{K}_s(\mathbb{I}_{0, \vartheta}) \cap \mathcal{B}(g, \tilde{\varepsilon})$ , where  $\tilde{\varepsilon} = M \cdot \max_i \varepsilon_i$  for some positive constant  $M$ .

PROOF. Let  $g$  be the kernel of a distributed delay. According to (14), we decompose it as a linear combination of elementary distributed delays and their successive derivatives. From the definition of  $g_{\text{app}}$ , which is clearly in  $\mathcal{G}$ , we have

$$\begin{aligned} \|g - g_{\text{app}}\|_{\mathcal{A}} &= \left\| \sum_{i,k} \hat{g}_{ik}(e^{-\vartheta s}) (\hat{\theta}_{\lambda_i}^{(k)}(s) - \hat{\psi}_{\lambda_i, n}^{(k)}(s)) \right\|_{\mathcal{A}} \\ &\leq \sum_{i,k} \|\hat{g}_{ik}(e^{-\vartheta s})\|_{\mathcal{A}} \|\hat{\theta}_{\lambda_i}^{(k)}(s) - \hat{\psi}_{\lambda_i, n}^{(k)}(s)\|_{\mathcal{A}}. \end{aligned}$$

Since, for any  $i$ ,  $\psi_{\lambda_i, n}$  are in  $\mathcal{B}(\theta_{\lambda_i}, \varepsilon_i)$ , using Lemma 3, we get

$$\begin{aligned} \|g - g_{\text{app}}\|_{\mathcal{A}} &\leq \sum_{i,k} \vartheta^k \|\hat{g}_{ik}(e^{-\vartheta s})\|_{\mathcal{A}} \|\hat{\theta}_{\lambda_i}(s) - \hat{\psi}_{\lambda_i, n}(s)\|_{\mathcal{A}} \\ &\leq \sum_{i,k} \vartheta^k \varepsilon_i \|\hat{g}_{ik}(e^{-\vartheta s})\|_{\mathcal{A}}. \end{aligned}$$

Denote the positive bounded constant  $M = \sum_{i,k} \vartheta^k \|\hat{g}_{ik}(e^{-\vartheta s})\|_{\mathcal{A}}$ . Hence

$$\|g - g_{\text{app}}\|_{\mathcal{A}} \leq \tilde{\varepsilon} = M \cdot \max_i \varepsilon_i,$$

that is  $g_{\text{app}}$  lies in  $\mathcal{B}(g, \tilde{\varepsilon})$ .  $\square$



By suitable choices for  $\varepsilon_i$ , the upper bound  $\tilde{\varepsilon}$  can be reduced arbitrarily, so that we can find  $g_{\text{app}}$  as close as we please of  $g$ . Note that we particularize the proofs of previous results with approximations with kernels in  $\mathcal{K}_s(\mathbb{I}_{0,\vartheta})$ , but they can be trivially extended to others approximations over  $\mathcal{A}$ . Previous results state that from the approximation of the elementary distributed delay, we can realize a kernel approximation in the graph topology for any distributed delay in  $\mathcal{G}$ . Previous results state that, for any  $f \in \mathcal{K}(\mathbb{I}_{\vartheta_1,\vartheta_2})$  and  $\varepsilon > 0$ , there exists an element  $f_{\text{app}} \in \mathcal{K}_s(\mathbb{I}_{\vartheta_1,\vartheta_2})$  such that

$$\|f - f_{\text{app}}\|_{\mathcal{A}} \leq \varepsilon. \quad (26)$$

Said differently, we have the following corollary.

**Corollary 2.** *The set  $\mathcal{K}_s(\mathbb{I}_{\vartheta_1,\vartheta_2})$  is dense in  $\mathcal{K}(\mathbb{I}_{\vartheta_1,\vartheta_2})$  for the graph topology.*

PROOF. Obvious from Lemma 1 and Theorem 3. □

In the proof of Lemma 2, convergence of polynomial approximation is uniform, so we get here a uniform convergence of this approximation. The assumption that kernel of distributed delay lie in  $\mathcal{K}(\mathbb{I}_{\vartheta_1,\vartheta_2})$  is not restrictive. Indeed, if this is not the case, any continuous function  $g$  in  $[\vartheta_1, \vartheta_2]$  can be approximated with respect to  $\mathcal{L}_1$ -norm as close as we please by a function in  $\mathcal{K}(\mathbb{I}_{\vartheta_1,\vartheta_2})$  [23]. Previous results addressed the approximation of distributed delays using kernels in  $\mathcal{K}_s(\mathbb{I}_{\vartheta_1,\vartheta_2})$ , which is based on uniform convergence of polynomials to any continuous function over  $\mathbb{I}_{\vartheta_1,\vartheta_2}$ . Other approximation can be proposed. Indeed, any element  $f$  in  $\mathcal{K}(\mathbb{I}_{\vartheta_1,\vartheta_2})$  can also be uniformly approximated by stepwise continuous functions in  $\mathcal{L}_1(\mathbb{I}_{0,\infty})$ . This can be seen as a particular case of the previous approximation. For this, consider

$$\theta_{\lambda,\text{app}}(t) = \sum_{i=0}^n \gamma_i \psi(t - t_i), \quad (27)$$

where  $t_i = \frac{i\vartheta}{n}$ , and  $\psi(\cdot)$  is a function in  $\mathcal{K}_s(\mathbb{I}_{0,\infty})$ . Then

$$\begin{aligned} \|\theta_{\lambda} - \theta_{\lambda,\text{app}}\|_{\mathcal{A}} &= \int_0^{\vartheta} |\theta_{\lambda}(t) - \theta_{\lambda,\text{app}}(t)| dt + \int_{\vartheta}^{\infty} |\theta_{\lambda,\text{app}}(t)| dt \\ &= \sum_{k=0}^{n-1} \int_{k\frac{\vartheta}{n}}^{(k+1)\frac{\vartheta}{n}} \left| \theta_{\lambda}(t) - \sum_{i=0}^k \gamma_i \psi(t - t_i) \right| dt + \int_{\vartheta}^{\infty} |\theta_{\lambda,\text{app}}(t)| dt. \end{aligned}$$

From Lemma 2, we see that approximation (27) uniformly converges to  $\theta_\lambda(\cdot)$  in the graph topology by a suitable choice of coefficients  $\gamma_i$ . Taking  $\psi(t) = e^{-\alpha t} h(t)$ , where  $\alpha > 0$  and  $h(\cdot)$  stands for the Heaviside function, we obtain as a particular case the result obtained in [16], and separately in [17], where a low pass filter is added in the integral approximation with lumped delayed distributions. Indeed, Laplace transform of (27) is of the required form

$$\hat{\theta}_{\lambda, \text{app}}(s) = \frac{1}{s + a} \sum_{i=0}^n \gamma_i e^{-st_i}$$

for  $a > 0$ . Coefficients  $\gamma_i$  can be obtained for instance from some numerical integral approximation, to guarantee approximation over  $\mathcal{A}$ , or equivalently over  $\mathcal{L}_1(\mathbb{I}_{0, \infty})$ .

#### 4. Discussion on approximation

We particularize in this section a proposal for approximation, and we analyze the main properties of the error in the time and frequency domains. These properties will be still valid for any other approximation in the graph topology over  $\mathcal{A}$ .

##### 4.1. Proposal for a constructive approximation

We propose in this subsection an effective and constructive approximation using elements in  $\mathcal{K}_s(\mathbb{I}_{\vartheta_1, \vartheta_2})$ . We start this proposal by the particular case of  $\theta_0(\cdot)$ . We denote  $C_n^k = \frac{n!}{k!(n-k)!}$ ,  $\psi_0(\xi) = \theta_0(-\alpha^{-1} \ln \xi)$ , for  $\xi \in ]0, 1]$ , and  $\psi_0(0) = 0$ .

**Lemma 4.** *Consider the sequence  $\theta_{0,n}(\cdot)$  in  $\mathcal{K}_s(\mathbb{I}_{0, \vartheta})$  described by*

$$\theta_{0,n}(t) = \sum_{k=0}^n C_n^k \psi_0\left(\frac{k}{n}\right) e^{-\alpha kt} (1 - e^{-\alpha t})^{n-k}, \quad t \in \mathbb{I}_{0, \vartheta}$$

and 0 elsewhere. Then  $\theta_{0,n}(\cdot)$  uniformly converges to  $\theta_0(\cdot)$  for the  $\mathcal{A}$ -norm.

PROOF. Define

$$\psi_0(\varrho) = \theta_0(-\alpha^{-1} \ln \varrho), \quad \varrho \in ]0, 1], \quad (28)$$

and  $\psi_0(0) = 0$ . The function  $\psi_0(\cdot)$  has a bounded step discontinuity for  $\varrho = e^{-\alpha \vartheta}$ . Since  $e^{-\alpha \vartheta}$  is irrational, there exists  $\kappa$  in  $0, 1, \dots, n-1$  such that

$\kappa/n < e^{-\alpha\vartheta} < (\kappa + 1)/n$ . We take  $\psi_{0,c}(\cdot)$  any continuous function over  $[0, 1]$  satisfying

$$\psi_{0,c}(\varrho) = \psi_0(\varrho), \varrho \in [0, \kappa/n] \cup [(\kappa + 1)/n, 1],$$

that will remove this discontinuity and will be close as desired to  $\psi_0(\cdot)$  for the  $\mathcal{L}_1(\mathbb{I}_{0,1})$ -norm. In particular  $\psi_{0,c}(\cdot)$  will satisfy

$$\psi_{0,c}(k/n) = \psi_0(k/n), k = 0, 1, \dots, n.$$

We now approximate as close as we please the function  $\psi_{0,c}$  (or equivalently  $\psi_0$ ) by Bernstein polynomials

$$\psi_{0,\text{app}}(\varrho) = \sum_{k=0}^n C_n^k \psi_0\left(\frac{k}{n}\right) \varrho^k (1 - \varrho)^{n-k}.$$

By the inverse transform of (28), we obtain

$$\theta_{0,n}(t) = \psi_{0,\text{app}}(e^{-\alpha t}),$$

which is in  $\mathcal{K}_s(\mathbb{I}_{0,\vartheta})$  since  $\psi_0(0) = 0$ . Furthermore, from Lemma 2, we know that  $\theta_{0,n}(\cdot)$  converges uniformly, by construction, to  $\theta_0(\cdot)$ .  $\square$

Previous lemma gives an approximation of  $\theta_0$  in  $\mathcal{B}(\theta_0, \varepsilon_n)$ , where positive upper bound  $\varepsilon_n$  can be reduced arbitrarily when  $n$  increases. This algorithm can be easily generalized to approximations in  $\mathcal{K}(\mathbb{I}_{0,\infty})$  or of the form (27), by modifying the domain of definition for (28). For the more general case of  $\theta_\lambda(\cdot)$ , we propose the following constructive solution.

**Lemma 5.** *The sequence of functions defined over  $\mathcal{K}_s(\mathbb{I}_{0,\vartheta})$  by*

$$\theta_{\lambda,n}(t) = \frac{1}{(1 - e^{-\alpha\vartheta})^n} \sum_{k=0}^n C_n^k \Phi_\lambda\left(\frac{k}{n}\right) (e^{-\alpha t} - e^{-\alpha\vartheta})^k (1 - e^{-\alpha t})^{n-k},$$

for  $t \in \mathbb{I}_{0,\vartheta}$ , and

$$\Phi_\lambda(\mu) = \theta_\lambda(-\alpha^{-1} \ln((1 - e^{-\alpha\vartheta})\mu + e^{-\alpha\vartheta})), \mu \in [0, 1],$$

uniformly converges to  $\theta_\lambda(\cdot)$ .

PROOF. From Lemma 2, define

$$\Phi_\lambda(\mu) = \theta_\lambda(-\alpha^{-1} \ln((1 - e^{-\alpha\vartheta})\mu + e^{-\alpha\vartheta}))$$

with  $\mu \in [0, 1]$ . Polynomial approximation  $\Phi_{\lambda, \text{app}}(\cdot)$  can be obtained from Bernstein polynomials, that is

$$\Phi_{\lambda, \text{app}}(\mu) = \sum_{k=0}^n C_n^k \Phi_\lambda\left(\frac{k}{n}\right) \mu^k (1 - \mu)^{n-k}. \quad (29)$$

The inverse transformation for  $\Phi_{\lambda, \text{app}}$  yields to

$$\theta_{\lambda, n}(t) = \frac{1}{(1 - e^{-\alpha\vartheta})^n} \sum_{k=0}^n C_n^k \Phi_\lambda\left(\frac{k}{n}\right) (e^{-\alpha t} - e^{-\alpha\vartheta})^k (1 - e^{-\alpha t})^{n-k},$$

which uniformly converges to  $\theta_\lambda(\cdot)$ .  $\square$

This approximation writes like a sum of elements in  $\mathcal{K}_s(\mathbb{I}_{0, \vartheta})$  and  $\theta_0(\cdot)$ . With Lemma 4 and approximation of  $\theta_0(\cdot)$ , such an approximation is defined over  $\mathcal{K}_s(\mathbb{I}_{0, \vartheta})$ , and writes like in (25). Some basic considerations can be made on the order of this approximation. We know from [21] or [22] that, given  $\varepsilon > 0$ , there exists  $\eta$  such that for any  $\mu_1, \mu_2 \in [0, 1]$ ,  $|\mu_1 - \mu_2| \leq \eta$  implies  $|\Phi_\lambda(\mu_1) - \Phi_\lambda(\mu_2)| \leq \frac{\varepsilon}{2\vartheta}$ , and that

$$|\Phi_\lambda(\mu) - \Phi_{\lambda, \text{app}}(\mu)| \leq \frac{\varepsilon}{2\vartheta} + \frac{\|\Phi_\lambda\|_{\mathcal{L}_\infty}}{2\vartheta\eta^2 n}.$$

There exists  $\beta$  a positive bounded real number, such that  $\left| \mu_1^{\frac{\lambda}{\alpha}} - \mu_2^{\frac{\lambda}{\alpha}} \right| \leq \beta |\mu_1 - \mu_2|$ , with for instance  $\beta \geq \frac{e^{\lambda\vartheta} - 1}{1 - e^{-\alpha\vartheta}}$ . This in turn implies that  $\eta \leq \frac{\varepsilon}{2\vartheta\beta} e^{-2\lambda\vartheta}$ . Taking the maximal admissible value for  $\eta$ , we finally obtain that the approximation order  $n$  satisfies

$$n \geq \frac{4\vartheta^3 \beta^2}{\varepsilon^3} e^{5\lambda\vartheta}.$$

From this simple consideration coming from the use of Bernstein polynomials, we obtain, to guarantee a norm upper bound  $\varepsilon$  for the error, a condition on the order of the approximation. This condition is however quite conservative, and in practice, the order can be chosen iteratively, as illustrated in Fig. 3 and Fig. 4.

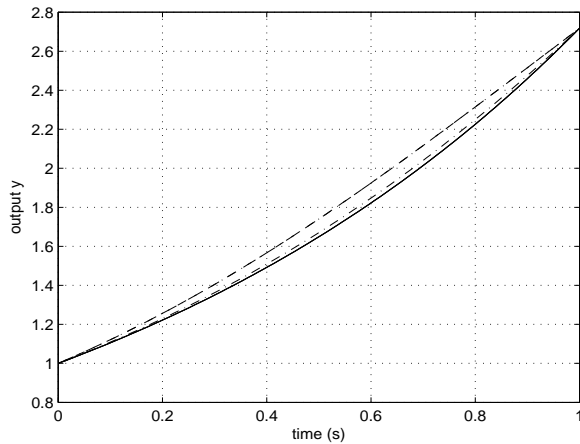


Figure 3: Kernel approximations  $\theta_{1,\text{app}}(t)$  using exponentials in the time domain of the kernel  $\theta_1(t)$  (continuous line), for orders  $n = 5$  (dashed) and  $n = 10$  (dot-dashed).

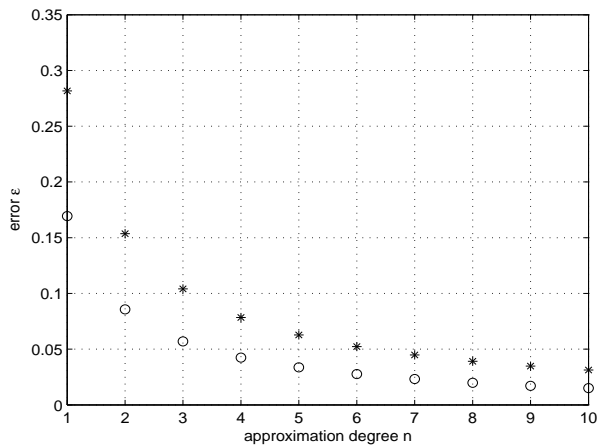


Figure 4: Representation of the norm error  $\|\theta_1 - \theta_{1,n}\|_{\mathcal{L}^\alpha}$  with respect to the order  $n$  of the approximation, for  $\alpha = 1$  (star) and  $\alpha = 1/5$  (circle).

#### 4.2. Properties in time and frequency domains

Let  $f_{\text{app}}$  be a kernel approximation in  $\mathcal{B}(f, \varepsilon)$ , for a given  $\varepsilon > 0$  and  $f \in \mathcal{K}(\mathbb{I}_{\vartheta_1, \vartheta_2})$ . The corresponding output is  $y_{\text{app}}(t) = (f_{\text{app}} * u)(t)$ . From (6),

for any  $1 \leq p \leq \infty$  and  $u \in \mathcal{L}_p(\mathbb{R}_+)$ ,

$$\|y - y_{\text{app}}\|_{\mathcal{L}_p} \leq \varepsilon \|u\|_{\mathcal{L}_p}. \quad (30)$$

In other words, the output error  $e(t) = y(t) - y_{\text{app}}(t)$  can be made arbitrarily small for the  $\mathcal{L}_p$ -norm by a suitable choice of the arbitrary bound  $\varepsilon$ . Such a property includes the case of persistent inputs in  $\mathcal{L}_1(\mathbb{R}_+)$ . For the particular case  $p = 2$ , we verify that the proposed approximation holds also in the frequency domain, since

$$\sup_{\text{Re } s \geq 0} |\hat{f}(s) - \hat{f}_{\text{app}}(s)| = \|\hat{f} - \hat{f}_{\text{app}}\|_{\mathcal{H}_\infty} \leq \varepsilon. \quad (31)$$

**Theorem 4.** *Let  $f \in \mathcal{A}$  be a given distribution, and  $f_{\text{app}}$  an approximation in  $\mathcal{B}(f, \varepsilon)$ , for some small  $\varepsilon > 0$ . Then*

$$\|\hat{f} - \hat{f}_{\text{app}}\|_{\mathcal{H}_\infty} = \sup_{\omega \in \mathbb{R}} |\hat{f}(j\omega) - \hat{f}_{\text{app}}(j\omega)| \leq \varepsilon,$$

and for all  $\omega$  in  $\mathbb{R}$ ,

$$|\arg(\hat{f}(j\omega)) - \arg(\hat{f}_{\text{app}}(j\omega))| \leq \varepsilon.$$

PROOF. Since  $f_{\text{app}} \in \mathcal{B}(f, \varepsilon)$ , and

$$\|\hat{f} - \hat{f}_{\text{app}}\|_{\mathcal{H}_\infty} \leq \|\hat{f} - \hat{f}_{\text{app}}\|_{\mathcal{A}},$$

module inequality is trivial. To show that approximation holds also for the phase angle, we denote  $\hat{f}_{\text{app}} = \hat{f} + \hat{e}_{\text{app}}$ . Hence

$$\left| |\hat{f}(j\omega)| |1 - e^{j(\varphi(\omega) - \varphi_{\text{app}}(\omega))}| - |\hat{e}_{\text{app}}(j\omega)| \right| \leq \varepsilon,$$

where  $\varphi(\omega) = \arg(\hat{f}(j\omega))$  and  $\varphi_{\text{app}}(\omega) = \arg(\hat{f}_{\text{app}}(j\omega))$ . This implies that

$$|\hat{f}(j\omega)| |1 - e^{j(\varphi(\omega) - \varphi_{\text{app}}(\omega))}|$$

can be made as small as we please for all  $\omega$ , which in turn implies that  $\varphi(\omega)$  and  $\varphi_{\text{app}}(\omega)$  are arbitrarily close for the  $\mathcal{L}_\infty$ -norm.  $\square$

This frequency property is illustrated in Fig. 5, where an approximation with order  $n = 5$  in  $\mathcal{H}_s(\mathbb{I}_{0,1})$  of  $\hat{\theta}_1(j\omega)$  is plotted. The approximation of distributed delay in the graph topology over  $\mathcal{A}$  yields to an approximation in both time and frequency domains. This is a strong property, that turns to be central in control problems.

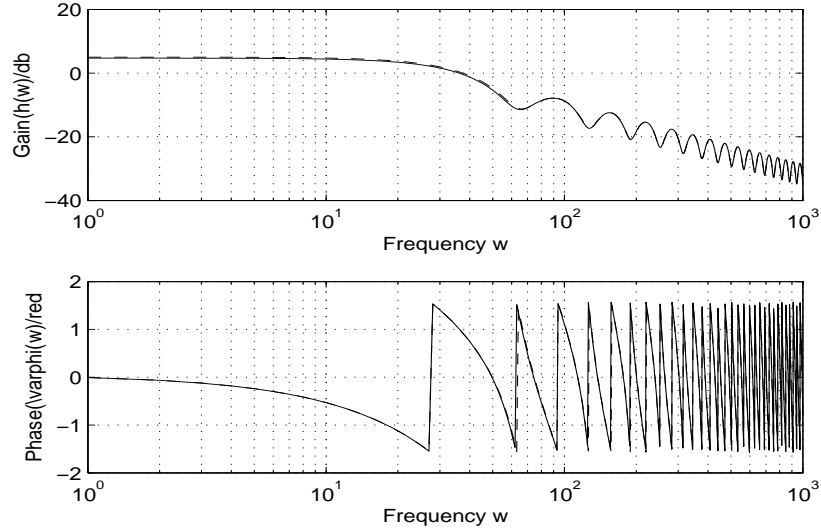


Figure 5: Bode diagram of  $\hat{\theta}_1(j\omega)$  and its kernel approximation  $\hat{\theta}_{1,\text{app}}(j\omega)$ , with order  $n = 5$ .

## 5. Application to control of time-delay systems

Distributed delays appear naturally in the control of time-delay systems. We consider here as control application, the stabilization problem, and we illustrate the approximation method to realize the distributed time operator.

### 5.1. Stabilization

In the stabilization problem, we determine a control law of a given plant, such that the closed-loop system is stable. For this, consider a distribution  $p$  with a coprime factorization  $(n, d)$  over  $\mathcal{A}$ , and define the closed-loop system described in Fig. 6, with a compensator  $c$  defined over the quotient field of  $\mathcal{A}$ .

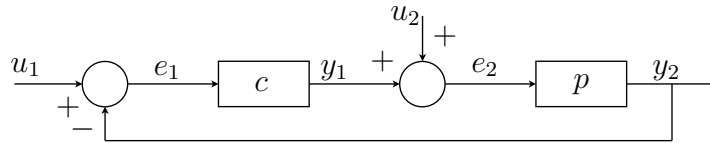


Figure 6: Feedback System.

We assume here that the plant  $p$  admits such a coprime factorization over  $\mathcal{A}$ . This is an assumption, but most dynamical systems fulfill it. The dynamical equations of this system are

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = H(p, c) \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad H(p, c) = \begin{bmatrix} \frac{c}{1+pc} & \frac{-pc}{1+pc} \\ \frac{pc}{1+pc} & \frac{p}{1+pc} \end{bmatrix}. \quad (32)$$

We say that  $c$  stabilizes  $p$ , or the pair  $(p, c)$  is stable, if the matrix  $H(p, c) \in \mathcal{A}^{2 \times 2}$ . All internal signals in the closed-loop are bounded for any bounded exogeneous inputs  $u_1$  and  $u_2$ . A necessary and sufficient condition for  $(p, c)$  to be stable is the existence of some  $n_c$  and  $d_c$  in  $\mathcal{A}$  such that

$$n * n_c + d * d_c = v, \quad (33)$$

with  $v$  a unit in  $\mathcal{A}$ . Since  $(n, d)$  is a coprime factorization, (33) holds, so we conclude that a stabilizing compensator  $c$  has a coprime factorization  $(n_c, d_c)$ . A coprime factorization  $(n_c, d_c)$  over  $\mathcal{A}$  of a stabilizing compensator for  $p$ , includes, in general, some distributed delays. Approximating such a compensator as described in Section 3 yields to the approximated controller  $c_{\text{app}}$  with factorization  $(n_{c,\text{app}}, d_{c,\text{app}})$ , where  $n_{c,\text{app}} \in \mathcal{B}(n_c, \varepsilon_n)$  and  $d_{c,\text{app}} \in \mathcal{B}(d_c, \varepsilon_d)$ , for some given  $\varepsilon_n$  and  $\varepsilon_d$ . Applying this approximation in the control, we would know the conditions on  $c_{\text{app}}$ , such that the pair  $(p, c_{\text{app}})$  is stable. The positive answer comes from a direct application of small gain theorem for BIBO-stability.

**Lemma 6.** *The pair  $(p, c_{\text{app}})$  is stable if*

$$\max(\varepsilon_n, \varepsilon_d) < \|(n \ d)\|_{\mathcal{A}}^{-1}.$$

PROOF. Let  $(n_c, d_c)$  be a coprime factorization over  $\mathcal{A}$  of a stabilizing controller. From (33), we have

$$n * n_c + d * d_c = \delta.$$

The approximated controller  $c_{\text{app}}$  with factorization  $(n_{c,\text{app}}, d_{c,\text{app}})$  yields in closed-loop

$$n * n_{c,\text{app}} + d * d_{c,\text{app}} = \delta - (n \ d) * \begin{pmatrix} n_c - n_{c,\text{app}} \\ d_c - d_{c,\text{app}} \end{pmatrix}. \quad (34)$$



The approximated controller will stabilize the plant  $p$  if and only if the right hand side in (34) is a unit over  $\mathcal{A}$ . Since  $\mathcal{A}$  is a Banach algebra, a sufficient condition is

$$\left\| \begin{pmatrix} n & d \end{pmatrix} * \begin{pmatrix} n_c - n_{c,\text{app}} \\ d_c - d_{c,\text{app}} \end{pmatrix} \right\|_{\mathcal{A}} < 1.$$

Via approximations,  $n_{c,\text{app}} \in \mathcal{B}(n_c, \varepsilon_n)$  and  $d_{c,\text{app}} \in \mathcal{B}(d_c, \varepsilon_d)$ , so that the above inequality yields the sufficient condition

$$\left\| \begin{pmatrix} n & d \end{pmatrix} \right\|_{\mathcal{A}} \max(\varepsilon_n, \varepsilon_d) < 1.$$

□

As a first comment, remark that it is always possible to determine an approximation  $c_{\text{app}}$  such that  $(p, c_{\text{app}})$  is stable. The counterpart will be in the order of approximation, that will increase when the required accuracy vanishes. Small gain theorem, which is still valid over any Banach algebra, gives us a sufficient condition on the approximation accuracy to guarantee robust stability. This condition helps us to determine the order of the approximation. Note that from [31] where the conservativeness of small gain theorem for BIBO-stability was studied, a converse statement for Lemma 6 holds, like in the  $\mathcal{H}_\infty$  case. This highlights the weak conservation of such a condition for robust stabilization. As an application, consider the plant  $\hat{y}(s) = \hat{p}(s)\hat{u}(s)$  given by

$$\hat{p}(s) = \frac{e^{-s}}{s-1}. \quad (35)$$

A coprime factorization writes  $\hat{n} = \frac{e^{-s}}{s+1}$ ,  $\hat{d} = \frac{s-1}{s+1}$ , since

$$\hat{n}(s) 2e^1 + \hat{d}(s) \left( 1 + 2\hat{\theta}_1(s) \right) = 1.$$

Hence, a stabilizing compensator for (35) is

$$u(t) = -2(\theta_1 * u)(t) + 2e^1 y(t).$$

From Lemma 6, a sufficient condition for robust stability with approximation of  $\theta_1$  is that  $\varepsilon_d \leq \frac{e}{3+e}$ . From Fig. 4, we see that a first order may be sufficient for stability purpose. In practice, a sufficient accuracy is obtained for a 5th order approximation over  $\mathcal{K}_s(\mathbb{I}_{0,1})$ , as can be seen in Fig. 3, where step responses are plotted. Note also that if we modify the coprime factorization  $(n, d)$ , we could obtain lower order approximations.

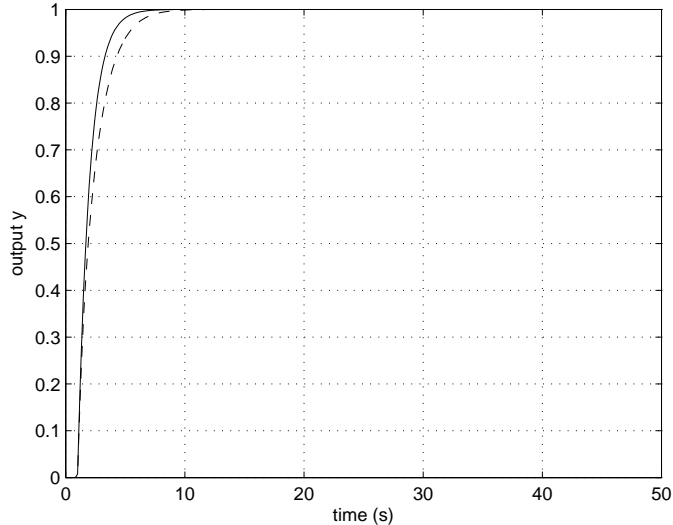


Figure 7: Step responses for the desired closed-loop (continuous line) and for the closed-loop system with approximation of distributed delay (5th order, dash).

## 6. Conclusion

In this paper, we proposed a general framework for rational approximation of distributed delay operators. The properties of this approximation were studied in both time and frequency domains. The effectiveness of this method was shown in simulation on the stabilization and the finite spectrum assignment problems, for general linear time-delay systems. We expect that this work will bring some new light in the understanding of distributed delay approximation, and more generally in approximation for the control of linear infinite dimensional systems. Such an approximation will provide foundations for a renewal of interest in control synthesis with distributed delays. Distributed delays inherit from integral control good robustness properties, and are at the core of numerous results in robustness analysis, optimization and control for time-delay systems.

## References

- [1] A. W. Olbrot, Stabilizability, detectability, and spectrum assignment for linear autonomous systems with general time delays, *IEEE Trans.*

- on Autom. Contr. 23 (1978) 887–890.
- [2] E. W. Kamen, P. P. Khargonekar, A. Tannenbaum, Proper stable bezout factorizations and feedback control of linear time-delay systems, *Int. J. Contr.* 43 (1986) 837–857.
  - [3] D. Brethé, J. J. Loiseau, A result that could bear fruit for the control of delay-differential systems, in: *IEEE MSCA, Chania, Greece*.
  - [4] H. Glüsing-Lüerßen, A behavioral approach to delay-differential systems, *SIAM J. Contr. Optimiz.* 35 (1997) 480–499.
  - [5] A. Z. Manitius, A. W. Olbrot, Finite spectrum assignment problem for systems with delays, *IEEE Trans. on Autom. Contr.* 24 (1979) 541–553.
  - [6] O. J. M. Smith, A controller to overcome dead time, *Inst. Soc. Amer. J.* 6 (1959) 28–33.
  - [7] Z. Artstein, Linear systems with delayed controls: a reduction, *IEEE Trans. on Autom. Contr.* 27 (1982) 869–879.
  - [8] H. Dym, T. T. Georgiou, M. C. Smith, Explicit formulas for optimally robust controllers for delay systems, *IEEE. Trans. on Autom. Contr.* 40 (1995) 656–669.
  - [9] S. Mondié, S. Niculescu, J. J. Loiseau, Delay robustness of closed loop finite assignment for input delay systems, in: *IFAC Workshop on Time-Delay Systems*.
  - [10] M. Di Loreto, Finite time model matching for time-delay systems, in: *IFAC Workshop on Time-Delay Systems, Aquila, Italy*.
  - [11] E. W. Kamen, P. P. Khargonekar, A. Tannenbaum, Stabilization of time-delay systems using finite-dimensional compensators, *IEEE Trans. on Autom. Contr.* 30 (1985) 75–78.
  - [12] J. R. Partington, Some frequency-domain approaches to the model reduction of delay systems, *Annual Reviews in Control* 28 (2004) 65–73.
  - [13] V. Van Assche, M. Dambrine, J. Lafay, Some problems arising in the implementation of distributed-delay control laws, in: *38th IEEE Conference on Decision & Control*, pp. 4668–4672.

- [14] Q.-C. Zhong, Robust Control of Systems with Delays, Ph.D. thesis, Imperial College London, London, 2003.
- [15] O. Santos, S. Mondié, Control laws involving distributed time delays: Robustness of the implementation, in: American Control Conference, pp. 2479–2480.
- [16] L. Mirkin, On the approximation of distributed-delay control laws, Systems & Control Letters 51 (2004) 331–342.
- [17] S. Mondié, W. Michiels, Finite spectrum assignment of unstable time-delay systems with a safe implementation, IEEE Trans. on Autom. Contr. 48 (2003) 2207–2212.
- [18] Q.-C. Zhong, Rational implementation of distributed delay using extended bilinear transformations, in: IFAC World Congress, Prague, Czech Republic.
- [19] J.-P. Richard, Time-delay systems: An overview of some recent advances and open problems, Automatica 39 (2003) 1667–1694.
- [20] Q.-C. Zhong, On distributed delay in linear control laws. part 1: Discrete-delay implementation, IEEE Trans. on Autom. Contr. 49 (2004) 2074–2080.
- [21] N. I. Achieser, Theory of approximation, Frederick Ungar Publishing Corp., New York, 1956.
- [22] E. W. Cheney, Introduction to approximation theory, AMS, 2nd Ed., Chelsea, 1982.
- [23] D. W. Kammler, Approximation with sums of exponentials in  $l_p[0, \infty)$ , J. of Approximation Theory 16 (1976) 384–408.
- [24] M. Vidyasagar, B. D. O. Anderson, Approximation and stabilization of distributed systems by lumped systems, Systems & Control Letters 12 (1989) 95–101.
- [25] Y. Otha, H. Maeda, S. Kodama, Rational approximation of  $l_1$  optimal controllers for siso systems, IEEE Trans. on Autom. Contr. 37 (1992) 1683–1691.

- [26] F. M. Callier, C. A. Desoer, An algebra of transfer functions for distributed linear time-invariant systems, *IEEE Trans. Circuits Syst.* 25 (1978) 651–662.
- [27] C. A. Desoer, M. Vidyasagar, *Feedback systems: Input-output properties*, Academic Press, New-York, 1975.
- [28] D. Brethé, J. J. Loiseau, An effective algorithm for finite spectrum assignment of single-input systems with delays, *J. Math. Computer in Simulations Special issue on time-delay systems*, V. Kolmanovskii and J. P. Richard Eds. (1997).
- [29] M. Vidyasagar, *Control System Synthesis. A Factorization Approach*, MIT Press, Cambridge, Massachussets, 1985.
- [30] A. Ralston, P. Rabinowitz, *A first course in numerical analysis*, Dover Publications, New York, 2nd Ed., 2001.
- [31] M. A. Dahleh, Y. Otha, A necessary and sufficient condition for robust bibo stability, *Systems & Control Letters* 11 (1988) 271–275.