Optimization of the $\mathcal{H}_2$ norm for single delay systems, with application to control design and model approximation

Marco A. Gomez, Alexey V. Egorov,
Sabine Mondié and Wim Michiels

***Abstract*—We propose a novel approach for the optimization of the $\mathcal{H}_2$ norm for time-delay systems, grounded in its characterization in terms of the delay Lyapunov matrix. We show how the partial derivatives of the delay Lyapunov matrix with respect to system or controller parameters can be semi-analytically computed, by solving a delay Lyapunov equation with inhomogeneous terms. It allows us to obtain the gradient of the $\mathcal{H}_2$ norm and in turn to use it in a gradient based optimization framework. We demonstrate the potential of the approach on two classes of problems, the design of robust controllers and the computation of approximate models of reduced dimension. Thereby a major advantage is the flexibility: in the former class of applications the order or structure of the controller can be prescribed, including recently proposed delay-based controllers. For the latter class of applications, approximate models described by both ordinary and delay differential equations (e.g., inhering the structure of the original system) can be synthesized.**

***Index Terms*—Time-delay systems, $\mathcal{H}_2$ norm optimization, delay Lyapunov matrix, model reduction.**

I. Introduction

One of the main objectives of a control design is to achieve given system performance specifications. The $\mathcal{H}_2$ norm, which provides a measure of robustness with respect to noise or external disturbances (see, for instance, [1]) is a widely used performance measure. The fact that under mild conditions it is a smooth function of the system parameters makes it suitable for optimization purposes. In particular, the optimization of $\mathcal{H}_2$ norms has proven to be useful in the synthesis of fixed-order dynamic optimal controllers/reduced order observers for high dimension systems [2]. It has also been successfully applied to the optimal model reduction problem, which consists in approximating a high dimension system by one of lower dimension (see, for example, [3] and [4]).

In recent years, the study of the $\mathcal{H}_2$ norm of delay differential equations has received some attention. See, for instance, [5], [6], and for more detailed explanations [7] and [8]. In [9] the $\mathcal{H}_2$ norm for time-delay systems is computed by using an auxiliary delay free system, obtained by spectral discretization. The important result presented by [10] and extended in [11] is the starting point of the present contribution. There, it is proved, in analogy with the delay free case, that the $\mathcal{H}_2$ norm can be expressed in terms of the delay Lyapunov matrix. There are basically two approaches for computing this matrix. One of them, called the semi-analytic approach, reduces to the computation of the solution of a set of differential equations with boundary conditions of mixed type obtained from the so-called dynamic, symmetry and algebraic properties [12], while the other, suitable when the delays are non-commensurate, is based on the discretization of the corresponding equations that results from taking polynomials as basis functions, [13], [10]. As in this paper we only consider one delay, the presented results are in line with and build on the semi-analytic approach. It is worthy mentioning that the results of this contribution can be trivially extended to the multiple commensurate delays case.

Marco A. Gomez and Sabine Mondié are with the Department of Automatic Control of CINVESTAV-IPN, Mexico city, 07360 Mexico, email: mgomez, smondie@ctrl.cinvestav.mx

Alexey V. Egorov is with St. Petersburg State University, St. Petersburg, 199034 Russia, e-mail: alexey.egorov@spbu.ru.

Wim Michiels is with the Department of Computer Science, KU Leuven, Leuven, 3001 Belgium, e-mail: wim.michiels@cs.kuleuven.be.

In this paper, we propose a general scheme for the optimization of the $\mathcal{H}_2$ norm for time-delay systems based on the results introduced in [10]. Two features makes this $\mathcal{H}_2$ norm characterization in terms of the delay Lyapunov matrix particularly tractable from an optimization perspective. First, as we shall see, the $\mathcal{H}_2$ norm is, under mild conditions, a smooth function of system or controller parameters. Second, an exact, computable objective function is available (instead of a possibly conservative upper bound). The main results of the paper concern the mathematical characterization of the sensitivity of the $\mathcal{H}_2$ norm and the associated delay Lyapunov matrix, along with computational procedures. Equally important is the demonstration of the potential of the obtained results in terms of control design and approximation. In particular, the main contributions can be summarized as follows.

1) The computation of the partial derivatives of the delay Lyapunov matrix with respect to matrix system parameters is obtained from the basic properties of this matrix.
2) The computation of the gradient of the $\mathcal{H}_2$ norm is described, allowing the optimization of the $\mathcal{H}_2$ norm within a (standard) gradient based optimization framework.
3) The approach is applied to the problems of controller synthesis and to the computation of $\mathcal{H}_2$ optimal approximate models of reduced dimension.

Note that both application problems are considered as very challenging. Although the proposed approach is illustrated with the synthesis of static controllers, the methodology can be trivially extended in order to construct dynamic controllers of fixed order. Even for finite-dimensional systems the design of fixed-order controllers with dimension lower than the dimension of the plant typically gives rise to non-convex optimization problems, when using both time-domain and frequency domain methods, see, e.g., [14], and this carries over to the infinite-dimensional delay setting. Also results on the approximation of delay systems are scarce and many problems related to the latter are still considered to be unsolved [15].

For the former application the $\mathcal{H}_2$ norm of the closed loop system is directly minimized as a function of the parameters of the controller. This approach complements the work of [16], [17], where stabilizing controllers are determined based on optimizing the spectral abscissa (the supremum of the real parts of the characteristic roots), and the work reported in [18] on the $\mathcal{H}_\infty$ optimization problem. From an optimization perspective a major difference is that, in contrast to the $\mathcal{H}_2$ norm, the spectral abscissa and the $\mathcal{H}_\infty$ norm are in general non-smooth functions of the system parameters (not everywhere differentiable, even not everywhere Lipschitz continuous). The direct optimization approach allows us to synthesize controllers with a prescribed structure or order (dimension). It is also possible to fix elements of the controller matrices, allowing to impose sparsity or an additional structure, e.g., a PID-like structure or, as we shall illustrate, a delay based controller.

In the latter application, approximate models are computed by minimizing the $\mathcal{H}_2$ norm of the approximation error on the transfer function, which is conceptually similar to the pioneering work of [19] and the references therein. The differences are that in [19] the $\mathcal{H}_\infty$ norm of the error is considered and that sufficient conditions guaranteeing a bound on the error, along with the approximate model, are obtained in an LMI framework. The adopted direct optimization approach has as main advantages a large flexibility in imposing structure on the approximate model and that local optimality of the computed models can be guaranteed, at the price that the underlying optimization problem is in general non-convex.

For sake of completeness, we summarize existing work on model reduction. Results related to the reduction of linear time-delay systems to finite-dimensional LTI (Linear Time Invariant) systems are based on interpolation/moment matching and grounded in either the Krylov framework [20] or the data driven moment Loewner framework [21] (see also [22] for the generalization of moment matching to nonlinear systems). Structure preserving reduction approaches, where the reduced model is also expressed in terms of a delay model, include approaches based on position balancing [23], a feedback interconnection interpretation of the system isolating the delay [24], a generalization of the dominant pole algorithm [25], moment matching [22] and the IRKA (Iterative Rational Krylov Algorithm) algorithm [26].

The manuscript is organized as follows. In the next section, we introduce basic definitions and concepts. The main result of this paper is presented in Section III. There, we provide the formulas for computing the partial derivatives of the delay Lyapunov matrix and the gradient of the $\mathcal{H}_2$ norm. Based on this, the optimization algorithm is introduced. Section IV is devoted to show the practical relevance of the main result: the synthesis of optimal $\mathcal{H}_2$ controllers is addressed in Section IV-A and the $\mathcal{H}_2$ optimal approximation problem is presented in Section IV-B. Finally, we end the paper with some remarks.

Throughout the paper we denote the space of $\mathbb{R}^n$-valued piecewise continuous functions on $[-h, 0]$ by $PC\left([-h,0],\mathbb{R}^n\right)$. The trace of a matrix A is represented by $\mathrm{Tr}(A)$. The symbol $\otimes$ stands for the classical Kronecker product of two matrices A and B [27] and $\mathrm{vec}(A)$ stands for the vectorization of a matrix A. The gradient of a function $f : \mathbb{R}^n \to \mathbb{R}$ is represented by ∇f. The notation $M'(t, \mathbf{p})$ and $\dfrac{\partial M(t,\mathbf{p})}{\partial p_i}$ stand for the matrix obtained by taking the partial derivative of each element of a matrix M with respect to t and with respect to the $i-th$ element of vector $\mathbf{p}$, respectively.

II. Preliminaries

Consider a system of the form

$$\begin{aligned} \dot{x}(t) =& A_0(\mathbf{p})x(t) + A_1(\mathbf{p})x(t-h) + B(\mathbf{p})v(t), \\ y(t) =& C(\mathbf{p})x(t), \end{aligned} \tag{1}$$

where the constant $h > 0$ is the delay, $x \in \mathbb{R}^n$ denotes the state vector, $v \in \mathbb{R}^{n_B}$ and $y \in \mathbb{R}^{n_C}$ represent the input and output of the system, respectively, and the vector $\mathbf{p} = \begin{pmatrix} p_1 & \dots & p_m \end{pmatrix}$ denotes the system parameters. We assume that the matrix functions $A_0(\mathbf{p}),\ A_1(\mathbf{p}) \in \mathbb{R}^{n\times n}$, $B(\mathbf{p}) \in \mathbb{R}^{n\times n_B}$, $C(\mathbf{p}) \in \mathbb{R}^{n_C\times n}$ smoothly depend on the system parameters $\mathbf{p}$. The transfer function is given by

$$G(s,\mathbf{p}) = C(\mathbf{p})\left(sI - A_0(\mathbf{p}) - A_1(\mathbf{p})e^{-sh}\right)^{-1}B(\mathbf{p}). \tag{2}$$

We assume that $t_0 = 0$ and that the initial condition φ of system (1) belongs to the space $PC\left([-h,0],\mathbb{R}^n\right)$, such that existence and uniqueness of solutions is guaranteed. Some basic definitions are presented next.

Definition 1. *System* (1) *is exponentially stable if there exist $\mu \geq 1$ and $\alpha > 0$ such that, for $v(t) = 0$, the solution $x(t,\varphi)$ satisfies*

$$\|x(t,\varphi)\| \leq \mu e^{-\alpha t} \sup_{\theta\in[-h,0]} \|\varphi(\theta)\|,\ t \geq 0.$$

The results presented in this work are focused on the $\mathcal{H}_2$ norm of system (1). We consider strictly proper rational transfer functions, i.e. the matrix D in the standard notation of system (1) is considered to be zero, as otherwise the $\mathcal{H}_2$ norm is not finite. For exponentially stable time-delay systems, it is defined as

$$\|G\|_2 = \left(\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{Tr}\left(G^*(j\omega,\mathbf{p})G(j\omega,\mathbf{p})\right)d\omega.\right)^{1/2}$$

It is well known that the $\mathcal{H}_2$ norm for delay free systems can be computed from the solution of a Lyapunov equation. In [10], it has been proved that this is also possible for the time-delay case by using the delay Lyapunov matrix U, which we define first.

Definition 2. *[12] A parametrized matrix $U(\tau,\boldsymbol{p})$, $\tau \in [-h,h]$, is a continuous delay Lyapunov matrix of system* (1) *associated with a symmetric matrix $W(\boldsymbol{p})$ if it satisfies the following three equalities:*

$$U'(\tau,\boldsymbol{p}) = U(\tau,\boldsymbol{p})A_0(\boldsymbol{p}) + U(\tau-h,\boldsymbol{p})A_1(\boldsymbol{p}),\ \tau\in[0,h], \tag{3}$$

$$U^T(\tau,\boldsymbol{p}) = U(-\tau,\boldsymbol{p}),\ \tau\in[-h,h], \tag{4}$$

$$\begin{aligned} &A_0(\boldsymbol{p})^T U(0,\boldsymbol{p}) + U(0,\boldsymbol{p})A_0(\boldsymbol{p}) \\ &\quad + A_1(\boldsymbol{p})^T U(h,\boldsymbol{p}) + U(-h,\boldsymbol{p})A_1(\boldsymbol{p}) = -W(\boldsymbol{p}), \end{aligned} \tag{5}$$

called the dynamic, symmetry and algebraic properties, respectively.

The following result characterizes the $\mathcal{H}_2$ norm in terms of the delay Lyapunov matrix.

Theorem 1. *[10] If system* (1) *is exponentially stable, then its $\mathcal{H}_2$ norm satisfies*

$$\|G\|_2^2 = \mathrm{Tr}\left(B^T(\boldsymbol{p})U(0,\boldsymbol{p})B(\boldsymbol{p})\right), \tag{6}$$

where $U(\tau,\boldsymbol{p})$ is the delay Lyapunov matrix associated with the matrix $W(\boldsymbol{p}) = C^T(\boldsymbol{p})C(\boldsymbol{p})$.

As a consequence the computational schemes for the $\mathcal{H}_2$ norm proposed in [10] rely on solving equations (3), (4) and (5). The existence and uniqueness of a solution for these equations depends on the so-called Lyapunov condition.

Theorem 2. *[12] The delay Lyapunov matrix of system* (1) *associated with a symmetric matrix $W(\boldsymbol{p})$ exists and is unique if and only if the system satisfies the Lyapunov condition, i.e., the spectrum*

$$\Lambda = \left\{s\in\mathbb{C}|\det\left(sI - A_0(\boldsymbol{p}) - A_1(\boldsymbol{p})e^{-sh}\right) = 0\right\}$$

does not contain a root $\hat{s}$ such that $-\hat{s}$ also belongs to the spectrum.

Remark 1. *The Lyapunov condition is always satisfied for an exponentially stable system. Moreover, if the condition of Theorem 2 holds, the solution of the delay Lyapunov matrix exists and is unique whether the system is stable or not.*

For the time-delay case, the computation of the Lyapunov matrix is not an algebraic problem anymore. Now, by using the semi-analytic approach, the matrix $U(\tau,\mathbf{p})$, $\tau \in [-h,h]$, is computed as the solution of a system of ordinary differential equations with boundary conditions, deduced from equations (3), (4) and (5), [12]. For the sake of completeness, we recall this result. Assume that the Lyapunov condition holds and consider the vector

$$\xi(\tau) = \begin{pmatrix} \mathrm{vec}(U(\tau,\mathbf{p})) \\ \mathrm{vec}(U(\tau-h,\mathbf{p})) \end{pmatrix},\ \tau\in[0,h].$$

As a result of the vectorization of the previously referred system of ordinary differential equations with boundary conditions, the solution of $\xi(\tau)$, $\tau\in[0,h]$, is determined by

$$\xi(\tau) = e^{L(\mathbf{p})\tau}\xi(0),\ \tau\in[0,h], \tag{7}$$

where

$$\xi(0) = -\left(M(\mathbf{p}) + N(\mathbf{p})e^{L(\mathbf{p})h}\right)^{-1} w(\mathbf{p}), \tag{8}$$

with matrices

$$
\begin{aligned}
w(\mathbf{p}) &= \begin{pmatrix} 0 \\ \operatorname{vec}(W(\mathbf{p})) \end{pmatrix}, \\
L(\mathbf{p}) &= \begin{pmatrix} A_0^T(\mathbf{p}) \otimes I & A_1^T(\mathbf{p}) \otimes I \\ -I \otimes A_1^T(\mathbf{p}) & -I \otimes A_0^T(\mathbf{p}), \end{pmatrix}, \\
M(\mathbf{p}) &= \begin{pmatrix} I \otimes I & 0_{n\times n} \otimes 0_{n\times n} \\ A_0^T(\mathbf{p}) \otimes I & A_1^T(\mathbf{p}) \otimes I \end{pmatrix}, \\
N(\mathbf{p}) &= \begin{pmatrix} 0_{n\times n} \otimes 0_{n\times n} & -I \otimes I \\ I \otimes A_1^T(\mathbf{p}) & I \otimes A_0^T(\mathbf{p}) \end{pmatrix}.
\end{aligned} \tag{9}
$$

The existence of the matrix inverse in (8) is guaranteed by the assumption on the Lyapunov condition (see Theorem 2.10 in [12]). The delay Lyapunov matrix $U(\tau, \mathbf{p})$, $\tau \in [-h, h]$, can be obtained by the devectorization of ξ.

Finally, observe that equations (3), (4) and (5) can be seen as a generalization of the algebraic Lyapunov equation of the delay free case.

III. Optimization of the $\mathcal{H}_2$ norm

In this section, we present the main result of our contribution. We first address smoothness properties of the delay Lyapunov matrix and introduce the computation of the partial derivatives with respect to the parameters of the system. Then, we show that the obtained formulae allow us to easily compute the gradient of the squared $\mathcal{H}_2$ norm. Finally, the $\mathcal{H}_2$ norm optimization algorithm is presented.

A. Sensitivity of the delay Lyapunov matrix

It is well known that in general the delay Lyapunov matrix $U(\tau, \mathbf{p})$ is not smooth with respect to τ at $\tau = 0$, having a discontinuity in the derivative. However, in the next lemma, we state that it is smooth with respect to the system parameters for $\tau \in [-h, h]$. This property is key in the subsequent results.

Lemma 1. *Assume that the Lyapunov condition holds for parameters $\boldsymbol{p} = \boldsymbol{p}_0$. Then the function $\boldsymbol{p} \mapsto U(\tau, \boldsymbol{p})$ is smooth at $\boldsymbol{p} = \boldsymbol{p}_0$ for every τ on $[-h, h]$.*

Proof. As the eigenvalues are continuous with respect to the system parameters $\mathbf{p}$, the Lyapunov condition holds for a sufficiently small vicinity of $\mathbf{p}_0$, which means that, by Theorem 2, the delay Lyapunov matrix is well defined in a neighborhood of $\mathbf{p}_0$.

Since the Lyapunov condition holds, implying that $M(\mathbf{p}) + N(\mathbf{p})e^{L(\mathbf{p})h}$ is invertible at $\mathbf{p} = \mathbf{p}_0$ and because matrices $A_0(\mathbf{p})$, $A_1(\mathbf{p})$ and $W(\mathbf{p})$ smoothly depend on parameters $\mathbf{p}$, the initial vector $\xi(0)$ determined by (8) is well defined, and it smoothly depends on $\mathbf{p}$. Combining this result with expression (7), the smoothness of $\xi(\tau)$ with respect to $\mathbf{p}$ at $\mathbf{p} = \mathbf{p}_0$, hence of the delay Lyapunov matrix $U(\tau, \mathbf{p})$ for any $\tau \in [-h, h]$, follows. □

The fact that the delay Lyapunov matrix smoothly depends on $\mathbf{p}$ allows us to state the next lemma.

Lemma 2. *The delay Lyapunov matrix $U(\tau, \boldsymbol{p})$ associated with a symmetric matrix $W(\boldsymbol{p})$ satisfies the following equations:*

$$
\frac{\partial U'(\tau,\boldsymbol{p})}{\partial p_i} = \frac{\partial U(\tau,\boldsymbol{p})}{\partial p_i} A_0(\boldsymbol{p}) + \frac{\partial U(\tau-h,\boldsymbol{p})}{\partial p_i} A_1(\boldsymbol{p}) + U(\tau,\boldsymbol{p}) \frac{\partial A_0(\boldsymbol{p})}{\partial p_i} + U(\tau-h,\boldsymbol{p}) \frac{\partial A_1(\boldsymbol{p})}{\partial p_i}, \ \tau \in [0,h], \tag{10}
$$

$$
\frac{\partial U^T(\tau,\boldsymbol{p})}{\partial p_i} = \frac{\partial U(-\tau,\boldsymbol{p})}{\partial p_i}, \ \tau \in [-h,h], \tag{11}
$$

$$
A_0^T(\boldsymbol{p}) \frac{\partial U(0,\boldsymbol{p})}{\partial p_i} + \frac{\partial U(0,\boldsymbol{p})}{\partial p_i} A_0(\boldsymbol{p}) + A_1^T(\boldsymbol{p}) \frac{\partial U(h,\boldsymbol{p})}{\partial p_i} + \frac{\partial U(-h,\boldsymbol{p})}{\partial p_i} A_1(\boldsymbol{p}) = -\frac{\partial W(\boldsymbol{p})}{\partial p_i} - \frac{\partial A_0^T(\boldsymbol{p})}{\partial p_i} U(0,\boldsymbol{p}) - U(0,\boldsymbol{p}) \frac{\partial A_0(\boldsymbol{p})}{\partial p_i} - \frac{\partial A_1^T(\boldsymbol{p})}{\partial p_i} U(h,\boldsymbol{p}) - U(-h,\boldsymbol{p}) \frac{\partial A_1(\boldsymbol{p})}{\partial p_i}. \tag{12}
$$

Proof. Equations (10), (11) and (12), follow directly from Lemma 1 and equations (3), (4) and (5), respectively. □

Consider for any $i \in \{1, \ldots, m\}$ and $\tau \in [0, h]$ the matrices

$$
Y_{p_i}(\tau) = \frac{\partial U(\tau, \mathbf{p})}{\partial p_i}, \quad Z_{p_i}(\tau) = \frac{\partial U(\tau - h, \mathbf{p})}{\partial p_i}. \tag{13}
$$

Inspired by Lemma 2.7 in [12], we prove that the sensitivity of the matrix $U(\tau, \mathbf{p})$ with respect to the parameters of the system can be recast as the solution of a standard system of ordinary differential equations with inhomogeneous terms depending on the delay Lyapunov matrix U, and subjected to boundary conditions.

Lemma 3. *The delay Lyapunov matrix $U(\tau, \boldsymbol{p})$, $\tau \in [0, h]$, associated with a symmetric matrix $W(\boldsymbol{p})$ is such that the matrices in (13) satisfy the system of matrix equations*

$$
\begin{aligned}
Y'_{p_i}(\tau) &= Y_{p_i}(\tau) A_0(\boldsymbol{p}) + Z_{p_i}(\tau) A_1(\boldsymbol{p}) + U(\tau,\boldsymbol{p}) \frac{\partial A_0(\boldsymbol{p})}{\partial p_i} + U(\tau-h,\boldsymbol{p}) \frac{\partial A_1(\boldsymbol{p})}{\partial p_i}, && (14) \\
Z'_{p_i}(\tau) &= -A_0^T(\boldsymbol{p}) Z_{p_i}(\tau) - A_1^T(\boldsymbol{p}) Y_{p_i}(\tau) - \frac{\partial A_0^T(\boldsymbol{p})}{\partial p_i} U(\tau-h,\boldsymbol{p}) - \frac{\partial A_1^T(\boldsymbol{p})}{\partial p_i} U(\tau,\boldsymbol{p}), && (15)
\end{aligned}
$$

with boundary conditions

$$
Y_{p_i}(0) = Z_{p_i}(h), \tag{16}
$$

$$
A_0^T(\boldsymbol{p}) Z_{p_i}(h) + Y_{p_i}(0) A_0(\boldsymbol{p}) + A_1^T(\boldsymbol{p}) Y_{p_i}(h) + Z_{p_i}(0) A_1(\boldsymbol{p}) = -\frac{\partial W(\boldsymbol{p})}{\partial p_i} - \frac{\partial A_0^T(\boldsymbol{p})}{\partial p_i} U(0,\boldsymbol{p}) - U(0,\boldsymbol{p}) \frac{\partial A_0(\boldsymbol{p})}{\partial p_i} - \frac{\partial A_1^T(\boldsymbol{p})}{\partial p_i} U(h,\boldsymbol{p}) - U(-h,\boldsymbol{p}) \frac{\partial A_1(\boldsymbol{p})}{\partial p_i}. \tag{17}
$$

Proof. Equation (14) follows from (10). Now, by (11), notice that

$$
Z_{p_i}(\tau) = \left(\frac{\partial U(h-\tau, \mathbf{p})}{\partial p_i} \right)^T, \ \tau \in [0, h].
$$

Then, by using (10) and symmetry property (4), we get equation (15) as follows:

$$
\begin{aligned}
Z'_{p_i}(\tau) = &-A_0^T(\mathbf{p}) \left(\frac{\partial U(h-\tau, \mathbf{p})}{\partial p_i} \right)^T - A_1^T(\mathbf{p}) \left(\frac{\partial U(-\tau, \mathbf{p})}{\partial p_i} \right)^T - \frac{\partial A_0^T(\mathbf{p})}{\partial p_i} U^T(h-\tau, \mathbf{p}) \\
&- \frac{\partial A_1^T(\mathbf{p})}{\partial p_i} U^T(-\tau, \mathbf{p}) = -A_0^T(\mathbf{p}) Z_{p_i}(\tau) - A_1^T(\mathbf{p}) Y_{p_i}(\tau) \\
&- \frac{\partial A_0^T(\mathbf{p})}{\partial p_i} U(\tau-h, \mathbf{p}) - \frac{\partial A_1^T(\mathbf{p})}{\partial p_i} U(\tau, \mathbf{p}).
\end{aligned}
$$

The boundary condition (16) is deduced directly from the definition of the matrices in (13) and equation (17) is obtained by rewriting (12) in terms of $Y_{p_i}(0)$ and $Z_{p_i}(h)$. □

In order to solve system (14)-(15) with boundary conditions (16)-(17) for a solution (Y_{p_i}, Z_{p_i}), we vectorize the system and use the Kronecker product property [27]

$$
\operatorname{vec}(EFH) = (H^T \otimes E) \operatorname{vec}(F), \tag{18}
$$

where E, F and H are matrices of compatible dimensions.

We introduce the vector

$$\xi_{p_i}(\tau) = \begin{pmatrix} \mathrm{vec}(Y_{p_i}(\tau)) \\ \mathrm{vec}(Z_{p_i}(\tau)) \end{pmatrix}, \; \tau \in [0, h],$$

and provide an expression for its computation in the next lemma.

Lemma 4. *Assume that the Lyapunov condition holds, then the next equality is satisfied for $\tau \in [0, h]$, $i = 1, \dots, m$:*

$$\xi_{p_i}(\tau) = e^{L(\boldsymbol{p})\tau}\xi_{p_i}(0) + e^{L(\boldsymbol{p})\tau}\int_0^\tau e^{-L(\boldsymbol{p})s}\frac{\partial L(\boldsymbol{p})}{\partial p_i}\xi(s)ds. \quad (19)$$

Here,

$$\xi_{p_i}(0) = \left(M(\boldsymbol{p}) + N(\boldsymbol{p})e^{L(\boldsymbol{p})h}\right)^{-1}\gamma_i(\boldsymbol{p}), \quad (20)$$

where

$$\gamma_i(\boldsymbol{p}) = -\frac{\partial w(\boldsymbol{p})}{\partial p_i} - \frac{\partial N(\boldsymbol{p})}{\partial p_i}\xi(h) - \frac{\partial M(\boldsymbol{p})}{\partial p_i}\xi(0) - N(\boldsymbol{p})e^{L(\boldsymbol{p})h}\int_0^h e^{-L(\boldsymbol{p})s}\frac{\partial L(\boldsymbol{p})}{\partial p_i}\xi(s)ds,$$

with the vector ξ determined by (7) *and matrices $L(\boldsymbol{p})$, $M(\boldsymbol{p})$, $N(\boldsymbol{p})$, given by* (9)*.*

Proof. By applying identity (18) to equations (14) and (15), we get

$$\xi'_{p_i}(\tau) = L(\mathbf{p})\xi_{p_i}(\tau) + \frac{\partial L(\mathbf{p})}{\partial p_i}\xi(\tau). \quad (21)$$

The solution of the above equation is clearly given by (19). We now show that the initial condition is determined by equation (20). Using (18) in the boundary conditions (16) and (17), we obtain

$$M(\mathbf{p})\xi_{p_i}(0) + N(\mathbf{p})\xi_{p_i}(h) = -\frac{\partial w(\mathbf{p})}{\partial p_i} - \frac{\partial M(\mathbf{p})}{\partial p}\xi(0) - \frac{\partial N(\mathbf{p})}{\partial p}\xi(h). \quad (22)$$

As (19) holds,

$$\xi_{p_i}(h) = e^{L(\mathbf{p})h}\xi_{p_i}(0) + e^{L(\mathbf{p})h}\int_0^h e^{-L(p)s}\frac{\partial L(\mathbf{p})}{\partial p_i}\xi(s)ds.$$

Substituting the above equation into (22), we get

$$\left(M(\mathbf{p}) + N(\mathbf{p})e^{L(\mathbf{p})h}\right)\xi_{p_i}(0) = \gamma_i(\mathbf{p}).$$

Since the Lyapunov condition holds, as we mention before, the matrix $M(\mathbf{p}) + N(\mathbf{p})e^{L(\mathbf{p})h}$ is non-singular and we obtain the desired result. □

Clearly, by devectorization of ξ_{p_i} in equation (19), we get

$$Y_{p_i}(\tau) = \frac{\partial U(\tau, \mathbf{p})}{\partial p_i}, \; \tau \in [0, h], \; i = 1, \dots, m.$$

B. The gradient of the $\mathcal{H}_2$ norm

The gradient of the $\mathcal{H}_2$ norm, which plays a key role in the optimization process, constitutes a measure of the sensitivity of the norm with respect to small parameter variations. The first two terms in the Taylor's series of $\|G(\cdot, \mathbf{p})\|_2^2$ around $\mathbf{p} = \boldsymbol{\rho}$ are given by

$$\|G(\cdot, \boldsymbol{\rho})\|_2^2 + \sum_{i=1}^m \frac{\partial}{\partial p_i}\|G(\cdot, \mathbf{p})\|_2^2\Big|_{\mathbf{p}=\boldsymbol{\rho}}(p_i - \rho_i) = \|G(\cdot, \boldsymbol{\rho})\|_2^2 + \Delta p\left(\nabla\|G(\cdot, \boldsymbol{\rho})\|_2^2\right),$$

with $\Delta p = \mathbf{p} - \boldsymbol{\rho}$. A formula for computing it is presented in the next theorem. It is a straightforward consequence of Theorem 1 and of the computation of the matrix $\frac{\partial U(0,\mathbf{p})}{\partial p_i}$, $i = 1, \dots, m$, which follows from Lemma 4.

Theorem 3. *Let $U(0, \boldsymbol{p})$ be the delay Lyapunov matrix associated with matrix $C^T(\boldsymbol{p})C(\boldsymbol{p})$. The gradient of the squared $\mathcal{H}_2$ norm of system* (1) *with respect to $\boldsymbol{p}$ is given by*

$$\nabla\|G\|_2^2 = \begin{pmatrix} 2\mathrm{Tr}\left(\frac{\partial B^T(\boldsymbol{p})}{\partial p_1}U(0,\boldsymbol{p})B(\boldsymbol{p})\right) + \mathrm{Tr}\left(B^T(\boldsymbol{p})\frac{\partial U(0,\boldsymbol{p})}{\partial p_1}B(\boldsymbol{p})\right) \\ \vdots \\ 2\mathrm{Tr}\left(\frac{\partial B^T(\boldsymbol{p})}{\partial p_m}U(0,\boldsymbol{p})B(\boldsymbol{p})\right) + \mathrm{Tr}\left(B^T(\boldsymbol{p})\frac{\partial U(0,\boldsymbol{p})}{\partial p_m}B(\boldsymbol{p})\right) \end{pmatrix}.$$

In what follows we briefly discuss some computational issues. The computation of the $\mathcal{H}_2$ norm only depends on the matrix $U(\tau, \mathbf{p})$ at $\tau = 0$ (or $\xi(0)$) , therefore its gradient computation only requires $\frac{\partial U(0,\mathbf{p})}{\partial p_i}$, obtained from the devectorization of $\xi_{p_i}(0)$. However, one still needs m derivatives of $U(0, \mathbf{p})$ with respect to the parameters.

Let us now discuss the computational complexity of obtaining the $\mathcal{H}_2$ norm and its gradient.

- The computation of $\xi(0)$ using (8) and the computation of $\xi_{p_i}(0)$, $i = 1, \dots, m$, using (20) rely on solving linear systems of equations with the same system matrix $M(\mathbf{p}) + N(\mathbf{p})e^{L(\mathbf{p})h}$. As a consequence, when using a direct solver, the factorization of this matrix needs to be done only *once*. Since the matrix has dimensions $n^2 \times n^2$, the computational cost of solving the systems of equations amounts to $\mathcal{O}(n^6)$ elementary operations for the factorization phase, and $\mathcal{O}((m+1)n^4)$ operations for the substitution phases (if no problem-specific structure or sparsity is exploited).
- To construct the system matrix in a preliminary phase, the computation of the matrix exponential requires $\mathcal{O}(n^6)$ operations. To obtain the right-hand sides $\gamma_i(\mathbf{p})$, we need to evaluate, for $i = 1, \dots, m$,
$$\int_0^h e^{L(\mathbf{p})(h-s)}\frac{\partial L(\mathbf{p})}{\partial p_i}\xi(s)ds.$$
The most efficient way to do so consists of numerically solving differential equation (21) with zero initial condition at $t = 0$, on the interval $[0, h]$. By using a Runge-Kutta integration scheme with grid points $\theta_k, k = 1, \dots, M$, and by exploiting the Kronecker structure of L in the matrix-vector products ($\mathcal{O}(n^3)$ instead of $\mathcal{O}(n^4)$ operations per matrix-vector product), the dominant cost of evaluating $\gamma_i(\mathbf{p})$, $i = 1, \dots, m$, amounts to $\mathcal{O}\left(mMn^3\right)$ operations.

Note further that both of the above ingredients are amendable for parallellization. When only evaluating the $\mathcal{H}_2$ norm, one should set $m = 0$ in the above expressions for the computational complexity. Hence, as long as $m << n^2$ the additional cost of computing the gradient is negligible.

For problems with high-dimension n, condition $m << n^2$ is usually satisfied in the context of control design and model reduction (see Section IV for the principles) since low-order controllers for delay systems, which are easy to implement, and reduced models of low dimension are desirable. However, the main computational bottleneck remains the evaluation of the $\mathcal{H}_2$ norms. For large-scale sparse problems this necessitate the development of dedicated iterative solvers, for which some preliminary steps have been reported in [28] and [29]. This is however beyond the scope of this paper.

C. General algorithm

The optimization of the $\mathcal{H}_2$ norm of system (1) with respect to the parameters corresponds to the minimization problem:

$$\min_{\mathbf{p}=(p_1,\dots,p_m)}\|G(\cdot, \mathbf{p})\|_2^2. \quad (23)$$

Considering the expression for the gradient of the squared $\mathcal{H}_2$ norm of system (1) given by Theorem 3, it is natural to use gradient based algorithms in order to solve the problem.

For our experiments we use the BFGS (BroydenFletcherGoldfarb-Shanno) quasi-Newton method, that is based on an iterative updating of the Hessian matrix approximation (see [30]). It is implemented as part of the MATLAB function `fminunc` included in the optimization toolbox. The method relies on the availability of a routine returning the objective function and its gradient at a given $\mathbf{p}$.

When using formula (6) a stability constraint on $\mathbf{p}$ needs to be included as the definition of the $\mathcal{H}_2$ norm is for exponentially stable systems. Explicitly adding the constraint of a negative spectral abscissa is to be avoided since the stability regions in the parameter spaces might be non-convex, with possibly highly non-smooth boundaries (e.g., cusps). Furthermore, minimizers are not expected at the border of the stability region since under mild controllability and observability conditions the $\mathcal{H}_2$ norm of G will diverge to infinity whenever the boundary is approached. For these reasons, we include the stability constraint implicitly by initializing the algorithm with a stabilizing value of $\mathbf{p}$ and setting the objective function to infinity if the system is unstable, i.e., we define the objective function as

$$f_{\text{obj}}(\mathbf{p}) = \\ = \begin{cases} \text{Tr}\left(B^T(\mathbf{p})U(0,\mathbf{p})U(\mathbf{p})\right), & \text{if (1) is asymptotically stable,} \\ +\infty, & \text{otherwise.} \end{cases} \quad (24)$$

If during the run of the algorithm a trial step would be generated outside the stability region, then, because of the formulation of the objective function as (24), the line search procedure will automatically reduce the step size, in such a way that non-stabilizing iterations are avoided.

To check stability one can, for example, compute the rightmost root via the QPmR (Quasi-Polynomial Mapping Based Rootfinder) algorithm [31], a spectral discretization as in TRACE-DDE [32], a Krylov method in the large-scale setting [33], or use the delay Lyapunov matrix based criteria [34], [35]. The requirement for stabilizing initial values might need a preliminary stabilization phase, for which the algorithm of [17], relying on minimizing the spectral abscissa, can be used.

Finally, we note that the optimization problem (23) is in general non-convex. Under mild conditions, convergence is guaranteed to a point satisfying the first order necessary optimality conditions, $\nabla||G||_2^2 = 0$ (see [36]). The computed optimizers lie in the connected component of the stability region in which the algorithm is initialized. A practical approach to handle these issues consists of generating multiple initial values, by carrying out the preliminary stabilization phase from a set of user-specified and/or random parameter vectors.

IV. Applications of the $\mathcal{H}_2$ norm optimization

In this section, the applicability of the results previously presented is demonstrated by addressing the control synthesis and the model approximation problems, respectively.

A. Control synthesis

We introduce a methodology for designing a controller that minimizes the $\mathcal{H}_2$ norm of the closed-loop system. The idea basically consists of proposing a controller and finding its parameters as outlined in Section III-C. Consider a time-delay system of the form

$$\begin{aligned} \dot{x}(t) =& A_{0o}x(t) + A_{1o}x(t-h) \\ &+ B_0u(t) + B_1u(t-h) + Bv(t), \\ y(t) =& Cx(t) \end{aligned}$$

where $A_{0o}, A_{1o} \in \mathbb{R}^{n\times n}$, $B_0, B_1 \in \mathbb{R}^{n\times n_B}$, $B \in \mathbb{R}^{n\times n_{B_o}}$ and $C \in \mathbb{R}^{n_C\times n}$. Here, u is the control signal, v is the exogenous disturbance and the output y is the measured signal.

Let us propose a static output feedback of the form

$$u(t) = K_c y(t), \quad (25)$$

where $K_c \in \mathbb{R}^{n_B\times n_C}$. The closed-loop system is

$$\begin{aligned} \dot{x}(t) =& A_0x(t) + A_1x(t-h) + Bv(t), \\ y(t) =& Cx(t), \end{aligned} \quad (26)$$

where

$$A_0 = A_{0o} + B_0K_cC, \text{ and } A_1 = A_{1o} + B_1K_cC.$$

Observe that system (26) is in the form of system (1) and that the elements of the matrix K_c of the controller can be viewed as parameters of system (26). They can be tuned to obtain a minimum of the $\mathcal{H}_2$ norm of the closed-loop system as stated in the following problem:

Problem 1. *Find the matrix K_c of the controller that stabilizes the system and minimize the $\mathcal{H}_2$ norm of the closed-loop system.*

Problem 1 can be solved by minimizing (24), where $\mathbf{p}$ represents the controller parameters as

$$\mathbf{p} = \text{vec}(K_c).$$

We illustrate this by some examples. The notation $G_{cl}(s)$ represents the transfer function of the closed-loop system.

Example 1. *Consider the system (see [16])*

$$\begin{aligned} \dot{x}(t) =& A_0x(t) + B_1u(t-5) + Bv(t), \\ y(t) =& x(t), \end{aligned}$$

where

$$A_0 = \begin{pmatrix} -0.08 & -0.03 & 0.2 \\ 0.2 & -0.04 & -0.005 \\ -0.06 & -0.2 & -0.07 \end{pmatrix},$$

$$B_1 = \begin{pmatrix} -0.1 \\ -0.2 \\ 0.1 \end{pmatrix}, \ B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We propose a controller of the form (25)*:*

$$u(t) = \begin{pmatrix} p_1 & p_2 & p_3 \end{pmatrix} y(t).$$

We consider as initial parameters

$$\boldsymbol{p}_0 = \begin{pmatrix} 0.472 & 0.505 & 0.603 \end{pmatrix},$$

which ensure the stability of the closed-loop system and minimize the spectral abscissa. By optimizing the $\mathcal{H}_2$ norm starting from the given initial parameter values, we obtained

$$\boldsymbol{p} = \begin{pmatrix} 0.53844 & 0.33763 & 0.22559 \end{pmatrix}$$

and the achieved value of the closed-loop system squared $\mathcal{H}_2$ norm is 32.4897. Figure 1 depicts the value of the squared $\mathcal{H}_2$ norm at every iteration and Table I shows the trade-off between the spectral abscissa and the $\mathcal{H}_2$ norm with the initial and the obtained parameter values.

TABLE I
TRADE-OFF BETWEEN THE $\mathcal{H}_2$ NORM AND THE SPECTRAL ABSCISSA.

	Spectral abscissa	Squared $\mathcal{H}_2$ norm
Spectral abscissa minimized ($\mathbf{p}_0$)	-0.15	79.3356
$\mathcal{H}_2$ norm optimized ($\mathbf{p}$)	-0.061543	32.4897

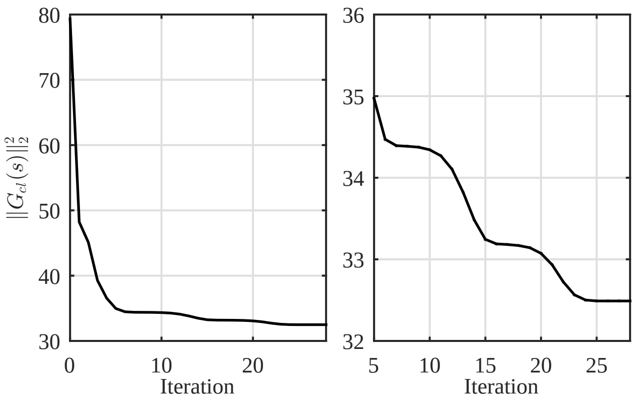

Fig. 1. Values of $\|G_{cl}\|_2^2$ at every iteration, Example 1. The figure on the right shows an enlargement of the figure on the left in the last iterations.

In the next example, we illustrate the applicability of the control design method, even when the delay is an optimization parameter.

Example 2. *Consider the second order system*

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B(u(t) + v(t)), \\ y(t) &= Cx(t), \end{aligned} \tag{27}$$

with matrices

$$A = \begin{pmatrix} 0 & 1 \\ -\nu^2 & -2\delta\nu \end{pmatrix},\ B = \begin{pmatrix} 0 \\ b \end{pmatrix},\ C = \begin{pmatrix} 1 & 0 \end{pmatrix},$$

where ν is the non-damped frequency, δ is the damping factor, and b is the input gain. We introduce a Proportional-Retarded (PR) controller of the form

$$u(t) = -k_p y(t) + k_r y(t-h). \tag{28}$$

This class of controllers, where the delay is a design parameter, has been studied in recent works (see, for instance, [37]). The closed-loop system (27)-(28) *is*

$$\begin{aligned} \dot{x}(t) = &\begin{pmatrix} 0 & 1 \\ -\nu^2 - bk_p & -2\delta\nu \end{pmatrix} x(t) \\ &+ \begin{pmatrix} 0 & 0 \\ bk_r & 0 \end{pmatrix} x(t-h) + \begin{pmatrix} 0 \\ b \end{pmatrix} v(t),\ \ y(t) = Cx(t). \end{aligned} \tag{29}$$

By defining the new time variable $\bar{t} = \dfrac{t}{h}$ and considering $\bar{x}(\bar{t}) = x(t)$, one can rewrite system (29) *as*

$$\begin{aligned} \dot{\bar{x}}(\bar{t}) = &\begin{pmatrix} 0 & h \\ -\nu^2 h - bk_p h & -2\delta\nu h \end{pmatrix} \bar{x}(t) \\ &+ \begin{pmatrix} 0 & 0 \\ bk_r h & 0 \end{pmatrix} \bar{x}(\bar{t}-1) + \begin{pmatrix} 0 \\ bh \end{pmatrix} v(\bar{t}), \\ \bar{y}(\bar{t}) = &C\bar{x}(\bar{t}). \end{aligned} \tag{30}$$

The relation of the $\mathcal{H}_2$ norm of systems (29) *and* (30) *is given as follows:*

$$\|G_{cl}\|_2^2 = \frac{1}{h}\|\bar{G}_{cl}\|_2^2, \tag{31}$$

where $\bar{G}_{cl}$ is the transfer function of the time scaled system (30)*. We use the above equality in order to minimize $\|G_{cl}\|_2^2$.*

We set the numerical values corresponding to the model of a DC servomechanism [37], $\nu = 17.6$, $\delta = 0.0128$, $b = 31$, and take the parameters vector as $\mathbf{p} = \begin{pmatrix} p_1 & p_2 & p_3 \end{pmatrix}$, with $p_1 = h$, $p_2 = k_p h$ and $p_3 = k_r h$. By (31)*, the gradient of $\|G_{cl}\|_2^2$ is given by*

$$\nabla \|G_{cl}\|_2^2 = \frac{1}{p_1} \begin{pmatrix} \dfrac{\partial \|\bar{G}_{cl}\|_2^2}{\partial p_1} - \dfrac{\|\bar{G}_{cl}\|_2^2}{p_1} \\ \dfrac{\partial \|\bar{G}_{cl}\|_2^2}{\partial p_2} \\ \dfrac{\partial \|\bar{G}_{cl}\|_2^2}{\partial p_3} \end{pmatrix}.$$

Figure 2 shows the value of the squared $\mathcal{H}_2$ norm of system (29) *at every iteration, when the algorithm is initialized by $\mathbf{p}_0 = \begin{pmatrix} 0.03 & 0.09 & 0.03 \end{pmatrix}$. At iteration 16 the values of the delay and the gains are $h = 0.01778$, $k_p = 168.8537$ and $k_r = 169.5677$, corresponding to $\|G_{cl}\|_2^2 = 0.018224$. There is no* finite *global minimizer and along the iterations we have $h \to 0$ and $|k_p|,\ |k_r| \to \infty$. In fact, the computed controller tends to a classical PD controller, related to (28) via $\dot{y} \approx \frac{1}{h}(y(t) - y(t-h))$.*

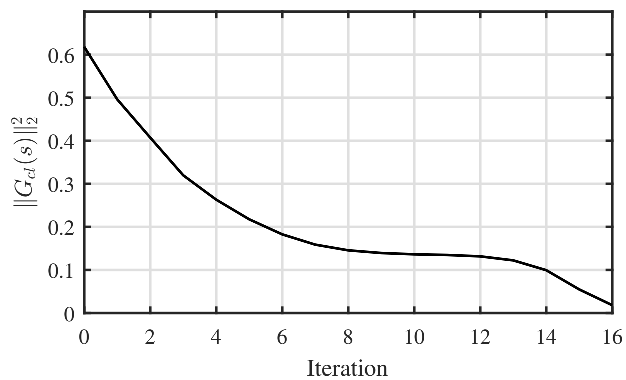

Fig. 2. Values of $\|G_{cl}\|_2^2$ at every iteration, Example 2.

High gains and small delays are to be avoided in an implementation. When fixing $k_p = 22.57$ as in [37], a finite local optimum is found starting with $\begin{pmatrix} p_1 & p_3 \end{pmatrix} = \begin{pmatrix} 0.03 & 0.9 \end{pmatrix}$. The optimized values are $h = 0.05187$ and $k_r = 17.9643$, and the achieved value of $\|G_{cl}\|_2^2$ is 0.049704.

It is worthy mentioning that the approach presented in this subsection can be applied to the design of dynamic controllers of fixed order.

B. $\mathcal{H}_2$ optimal model approximation

When a system is of high dimension, it is for many purposes (analysis, simulation, control design,...) convenient to have a lower order approximation. The $\mathcal{H}_2$ norm can be used as a measure of the approximation error between the two systems. A major advantage of the direct optimization approach is its generality. It allows us, for instance, to preserve a time-delay structure (i.e., the reduced system is also a time-delay system), or compute delay-free reduced order models. As another motivation, only in a very limited number of cases, explicit optimality conditions are available (see, e.g., [38] for the approximation of a delay-free system by a reduced model with transfer function of the form $p_1/(s - p_2^{-s\tau})$).

Consider the transfer function (2) and the one of a reduced dimension r model given by

$$G_r(s, \mathbf{p}_r) = C_r(\mathbf{p}_r)\left(sI - A_{0r}(\mathbf{p}_r) - A_{1r}(\mathbf{p}_r)e^{-sh}\right)^{-1} B_r(\mathbf{p}_r),$$

where $\mathbf{p}_r$ is a parametrization of the reduced order model. Define $G_e(s, \mathbf{p}_r) = G(s, \mathbf{p}) - G_r(s, \mathbf{p}_r)$ and observe that

$$G_e(s, \mathbf{p}_r) = C_e\left(sI_e - A_{0e} - A_{1e}e^{-sh}\right)^{-1} B_e, \tag{32}$$

with

$$I_e = \begin{pmatrix} I_n & 0 \\ 0 & I_r \end{pmatrix}, \; B_e = \begin{pmatrix} B \\ B_r(\mathbf{p}_r) \end{pmatrix},$$
$$C_e = \begin{pmatrix} C & -C_r(\mathbf{p}_r) \end{pmatrix}, \; A_{ie} = \begin{pmatrix} A_i & 0 \\ 0 & A_{ir}(\mathbf{p}_r) \end{pmatrix}, \; i = 0, 1.$$

Now, the parameters of the reduced dimension system are parameters of the extended transfer function (32). The optimal model reduction problem can be stated as follows:

Problem 2. *Find a reduced order system with matrices* $(A_{0_r}, A_{1_r}, B_r, C_r)$ *that minimizes the* $\mathcal{H}_2$ *norm of the transfer function* $G_e(s, \boldsymbol{p}_r)$.

We illustrate the solution of Problem 2 with two examples.

Example 3. *A nonlinear model for describing a refining plant is obtained in [39] and the linear case considering the delay induced by the recycle process is considered in [40]. The linearized model is given by*

$$\begin{aligned} \dot{x}(t) =& A_0 x(t) + A_1 x(t - 0.1) + B v(t), \\ y(t) =& C x(t), \end{aligned} \tag{33}$$

where

$$A_0 = \begin{pmatrix} -4.93 & -1.01 & 0 & 0 \\ -3.20 & -5.30 & -12.8 & 0 \\ 6.40 & 0.347 & -32.5 & -1.04 \\ 0 & 0.833 & 11 & -3.96 \end{pmatrix},$$
$$A_1 = \begin{pmatrix} 1.92 & 0 & 0 & 0 \\ 0 & 1.92 & 0 & 0 \\ 0 & 0 & 1.87 & 0 \\ 0 & 0 & 0 & 0.724 \end{pmatrix}, \; B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

and

$$C = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}.$$

We assume that the output matrix is given by C*. The recycle delay here is considered to be of one minute, that according to the time-scale in [40] is equivalent to set* $h = 0.1$.

We propose a reduced order system of dimension $r = 2$

$$\begin{aligned} \dot{x}_r(t) =& A_{0r} x_r(t) + A_1 x_r(t - 0.1) + B_r v_r(t), \\ y_r(t) =& C_r x(t). \end{aligned} \tag{34}$$

We define the vector of parameters

$$\boldsymbol{p}_r = \begin{pmatrix} \mathrm{vec}(A_{0r})^T & \mathrm{vec}(A_{1r})^T & \mathrm{vec}(B_r)^T & \mathrm{vec}(C_r)^T \end{pmatrix}$$

and compute values of $\boldsymbol{p}_r$ *that minimize* $\|G_e\|_2^2$*, where the transfer function* G_e *is determined by* (32) *with matrices*

$$A_{ie} = \begin{pmatrix} A_i & 0 \\ 0 & A_{ir} \end{pmatrix}, \; B_e = \begin{pmatrix} B \\ B_r \end{pmatrix}, \; C_e = \begin{pmatrix} C & -C_r \end{pmatrix},$$

where $i = 0, 1$, *and* A_0, A_1, B *and* C_2 *are given in* (33)*. The obtained matrices of system* (34) *with initial parameter values*

$$A_{0r_0} = \begin{pmatrix} -3 & -1 \\ -3 & -2 \end{pmatrix}, \; A_{1r_0} = \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix},$$
$$B_{r_0} = \begin{pmatrix} 1.6 & 0.3 \\ 0.15 & 0.7 \end{pmatrix}, \; C_{r_0} = \begin{pmatrix} 0.7 & -0.7 \end{pmatrix},$$

are

$$A_{0r} = \begin{pmatrix} -5.9722 & 0.1903 \\ -8.1593 & -1.3019 \end{pmatrix}, \; A_{1r} = \begin{pmatrix} -4.3358 & 3.1555 \\ -9.8049 & 6.8424 \end{pmatrix},$$
$$B_r = \begin{pmatrix} 1.1664 & 0.8980 \\ 1.9544 & 1.7810 \end{pmatrix}, \; C_r = \begin{pmatrix} 0.2356 & -0.6808 \end{pmatrix},$$

and the achieved value of $\|G_e\|_2^2$ *is* 3.5059×10^{-5}*. Figure 3 shows the value of* $\|G_e\|_2^2$ *at every iteration and Figure 4 the frequency response comparison between system* (33) *and the obtained reduced order system* (34).

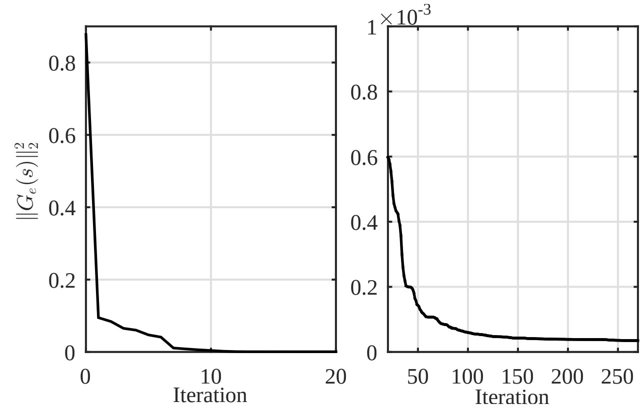

Fig. 3. Values of $\|G_e\|_2^2$ at every iteration, Example 3. The figure on the left shows the first twenty iterations while the figure on the right depicts an enlargement of the last iterations.

Fig. 4. Bode magnitude diagram of systems (33) and (34) with the obtained parameters $\mathbf{p}_r$, Example 3.

V. Conclusions

The computation of the sensitivity of the delay Lyapunov matrix with respect to given system parameters allows us to find the gradient of the $\mathcal{H}_2$ norm of time delay systems. Gradient based optimization algorithm enable us to synthesize fixed-order controllers and to find $\mathcal{H}_2$ optimal approximate models of reduced dimension. The presented examples show the effectiveness of the proposed approach.

A first direction for future work is the extension of these results to other classes of time-delay systems and the integration with recent results on $\mathcal{H}_\infty$ synthesis, aiming at a general $\mathcal{H}_2$-$\mathcal{H}_\infty$ framework for the synthesis of structured controllers. A second direction consists of developing iterative numerical algorithms for the $\mathcal{H}_2$ norm computation and optimization for large-scale sparse problems, for which some preliminary steps are reported in [28].

Acknowledgment

This work was supported by the Project CONACYT 180725, the Programme of Interuniversity Attraction Poles of the Belgian Federal Science Policy Office (IAP P6-DYSCO), by project G0A5317N of the Research Foundation-Flanders (FWO - Vlaanderen), and by the project UCoCoS, funded by the European Unions Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie Grant Agreement No 675080.

References

[1] K. Zhou, J. Doyle, and K. Glover, *Robust and Optimal Control*. Prentice-Hall, 1996.

[2] P. Apkarian, D. Noll, J.-B. Thevenet, and H. Tuan, "A spectral quadratic-SDP method with applications to fixed-order $\mathcal{H}_2$ and $\mathcal{H}_\infty$ synthesis," *European Journal of Control*, vol. 10, pp. 527–538, 2004.

[3] W. Yan and J. Lam, "An approximate approach to $\mathcal{H}_2$ optimal model reduction," *IEEE Transactions on Automatic Control*, vol. 44, no. 7, pp. 1341–1358, 1999.

[4] P. Van Dooren, K. Gallivan, and P.-A. Absil, "$\mathcal{H}_2$-optimal model reduction of MIMO systems," *Applied Mathematics Letters*, vol. 21, pp. 1267–1273, 2008.

[5] L. Mirkin, Z. J. Palmor, and D. Shneiderman, "H^2 optimization for systems with adobe input delays: a loop shifting approach," *Automatica*, vol. 48, pp. 1722–1728, 2012.
[6] W. Michiels, G. Hilhorst, G. Pipeleers, and J. Swevers, "Model order reduction for time-delay systems, with application to fixed-order $\mathcal{H}_2$ optimal controller design," in *Recent results on time-delay systems: analysis and control*, ser. Advances in Delays and Dynamics, E. Witrant, E. Fridman, O. Sename, and L. Dugard, Eds. Springer, 2016.
[7] W. Michiels and S.-I. Niculescu, *Stability, Control, and Computation for Time-delay Systems: An Eigenvalue-based Approach*, ser. Advances in Design and Control. SIAM, 2014, vol. 27.
[8] E. Fridman, *Introduction to time-delay systems: analysis and control*. Birkhauser, 2014.
[9] J. Vanbiervliet, W. Michiels, and E. Jarlebring, "Using spectral discretisation for the optimal $\mathcal{H}_2$ design of time-delay systems," *International Journal of Control*, vol. 84, no. 2, pp. 228–241, 2011.
[10] E. Jarlebring, J. Vanbiervliet, and W. Michiels, "Characterizing and computing the H_2 norm of time-delay system by solving the delay Lyapunov equation," *IEEE Transactions on Automatic Control*, vol. 56, no. 4, pp. 814–825, 2011.
[11] V. A. Sumacheva and V. L. Kharitonov, "Computation of the $\mathcal{H}_2$ norm of the transfer matrix of a neutral type system," *Differential equations*, vol. 50, no. 13, pp. 1752–1759, 2014.
[12] V. L. Kharitonov, *Time-Delay Systems: Lyapunov functionals and matrices*. Basel: Birkhäuser, 2013.
[13] E. Huesca, S. Mondié, and O. Santos, "Polynomial approximations of the Lyapunov matrix of a class of time delay systems," in *8th IFAC Workshop on Time-Delay Systems*, vol. 42, no. 14, 2009, pp. 261 – 266, 8th IFAC Workshop on Time-Delay Systems. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S147466701530570X
[14] D. Henrion and M. Šebek, "Overcoming non-convexity in polynomial robust control design," in *Proceedings of the Symposium of Mathematical Theory of Networks and Systems (MTNS)*, Leuven, Belgium, 2004.
[15] J. Partington, "Model reduction of delay systems," in *Unsolved Problems in Mathematical Systems and Control Theory*, V. Blondel and A. Megretski, Eds. Princeton University Press, 2004, pp. 29–32.
[16] J. Vanbiervliet, K. Berheyden, W. Michiels, and S. Vandewalle, "A nonsmooth optimisation approach for the stabilisation of time-delay systems," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 14, pp. 478–493, 2008.
[17] W. Michiels, "Spectrum based stability analysis and stabilization of systems described by delay differential algebraic equations," *IET Control Theory and Applications*, vol. 5, no. 16, pp. 1829–1842, 2011.
[18] S. Gumussoy and W. Michiels, "Fixed-order H-infinity control for interconnected using delay differential algebraic equations," *SIAM Journal on Control and Optimization*, vol. 49, no. 5, pp. 2212–2238, 2011.
[19] J. Lam, H. Gao, and C. Wang, "$\mathcal{H}_\infty$ model reduction of linear systems with distributed delay," *IEE Proceedings - Control Theory and Applications*, vol. 152, no. 6, pp. 662–674, 2005.
[20] W. Michiels, E. Jarlebring, and K. Meerbergen, "Krylov based model order reduction of time-delay systems," *SIAM Journal on Matrix Analysis and Applications*, vol. 32, no. 4, pp. 1399–1421, 2011.
[21] I. Pontes Duff, P. Vuillemin, C. Poussot-Vassal, C. Briat, and C. Seren, "Approximation of stability regions for large-scale time-delay systems using model reduction techniques," in *Proceeding of the 14th European Control Conference*, Linz, Austria, 2015, pp. 356–361.
[22] G. Scarciotti and A. Astolfi, "Model reduction of neutral linear and nonlinear time-invariant time-delay systems with discrete and distributed delays," *IEEE Transactions on Automatic Control*, vol. 61, no. 6, 2016.
[23] E. Jarlebring, T. Damm, and W. Michiels, "Model reduction of time-delay systems using position balancing and delay lyapunov equations," *Mathematics of Control, Signals and Systems*, vol. 25, no. 2, pp. 147–166, 2013.
[24] N. van de Wouw, W. Michiels, and B. Besselink, "Model reduction for delay differential equations with guaranteed stability and error bound," *Automatica*, vol. 55, pp. 132–139, 2015.
[25] M. Saadvandi, K. Meerbergen, and W. Desmet, "On dominant poles and model reduction of second order time-delay systems," *Applied Numerical Mathematics*, vol. 62, no. 1, pp. 21–34, 2014.
[26] P. Schulze and B. Unger, "Data-driven interpolation of dynamical systems with delay," *System & Control Letters*, vol. 97, pp. 127–131, 2016.
[27] A. Graham, *Kronecker products and matrix calculus with applications*. Ellis Horwood Limited, 1981.
[28] J. Peeters and W. Michiels, "Computing the $\mathcal{H}_2$ norm of large-scale time-delay systems," in *11th IFAC Workshop on Time-Delay Systems*, 2013.
[29] E. Jarlebring and F. Poloni, "Iterative methods for the delay Lyapunov equation with T-sylvester preconditioning," *Preprint arXiv:1507.02100v1*, 2015.
[30] S. S. Rao, *Engineering optimization: theory and practice*. John Wiley & Sons, 2009.
[31] T. Vyhlídal and P. Zítek, "Mapping based algorithm for large-scale computation of quasi-polynomial zeros," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 171–177, 2009.
[32] D. Breda, S. Maset, and R. Vermiglio, "TRACE-DDE: A tool for robust analysis and characteristic equations for delay differential equations," in *Topics in Time Delay Systems: Analysis, Algorithm and Control*, J. L. et al., Ed. Springer, 2009, pp. 145–155.
[33] E. Jarlebring, K. Meerbergen, and W. Michiels, "A Krylov method for the delay eigenvalue problem," *SIAM Journal on Scientific Computing*, vol. 32, no. 6, pp. 3278–3300, 2010.
[34] A. V. Egorov, C. Cuvas, and S. Mondié, "Necessary and sufficient stability conditions for linear systems with pointwise and distributed delays," *Automatica*, vol. 80, pp. 218–224, 2017.
[35] A. V. Egorov, "A finite necessary and sufficient stability condition for linear retarded type systems," in *Proceedings of 55th IEEE Conference on Decision and Control, Las Vegas, USA*, 2016, pp. 3155–3160.
[36] J. Nocedal and S. Wright, *Numerical Optimization*. New York, USA: Springer, 2006.
[37] R. Villafuerte, S. Mondié, and R. Garrido, "Tuning of proportional retarded controllers: theory and experiments," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 3, pp. 983–990, 2013.
[38] I. Pontes Duff, S. Gugercin, C. Beattie, C. Poussot-Vasseal, and C. Serin, "$\mathcal{H}_2$-optimality conditions for reduced time-delay systems of dimension one," in *Proceeding of the 13th IFAC Workshop on Time Delay Systems*, Istanbul, Turkey, 2016, pp. 7–12.
[39] T. J. Williams and R. E. Otto, "A generalized chemical processing model for the investigation of computer control," *Transactions of the American Institute of Electrical Engineers, Part I: Communication and Electronics*, vol. 79, no. 5, pp. 458–473, 1960.
[40] D. Ross, "Control design for time lag systems via quadratic criterion," *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 664–672, 1971.