A Novel Method to Compute the Structured Distance to Instability for Combined Uncertainties on Delays and System Matrices

Francesco Borgioli, Wim Michiels

***Abstract*—This paper presents a novel method to compute the distance to instability for linear time-delay systems with multiple delays, subjected to combined uncertainties on the system matrices and the delay parameters. The method is able to fully exploit structure on and possible inter-dependencies between the uncertainties affecting the coefficient matrices, the property that realistic perturbations are real valued, as well as the structure of the delay equation. The resulting algorithm relies on a characterization in terms of the pseudospectral abscissa (PSA) and on techniques from optimization over manifolds of matrices.**

I. Introduction

A widely used measure for the robustness of stability of a dynamical system is the distance to instability (or stability radius), which reveals the smallest (in terms of magnitude) perturbation that renders the system unstable: for linear time-invariant (LTI) systems in continuous time this would be the smallest perturbation that pushes one or more eigenvalues to the right half of the complex plane (equivalently, that generates a non-negative spectral abscissa, i.e. the real part of the rightmost eigenvalue). If we consider the simplest model $\dot{x} = Ax$, the real and the complex stability radii are respectively defined as the norms of the minimum real-valued and of the minimum complex-valued perturbation of matrix A that gives rise to instability of the system; another important distinction can also be made between structured and unstructured stability radius, with the former one that allows to characterize a more specific perturbation, where for instance only a single block or coefficient of the matrix A is uncertain. The main contributions for the computation of the complex and real stability radii include [6], respectively [16], [19].

In the present paper we focus on a class of dynamical systems described by a set of LTI delay differential equations of retarded type (see the conclusions for generalizations),

$$\dot{x}(t) = A_0 x(t) + \sum_{i=1}^{m} A_i x(t - \tau_i), \tag{1}$$

where $A_i \in \mathbb{R}^{n\times n}$ with $n \geq 2$ and $0 < \tau_1 < \cdots < \tau_m$, and we present a method to compute the real structured stability radius when the system is subject to real-valued uncertainties on the matrices and on the delays simultaneously. In addition, structure can be imposed on the coefficient matrices, independently of each other, in order to potentially perturb a submatrix, i.e. a single block or a single coefficient. Moreover, we consider the potential inter-dependencies between the uncertainties perturbing the nominal matrices A_i: this has a major relevance for those uncertain system parameters which are present in more than one matrix A_i (for an application see Example 1 in Section II). The generality of the setting and the latter property precludes the application of existing techniques.

First, due to the nonlinear dependence of the characteristic equation on the delay parameters, a reformulation in the form $\det(I - M(\lambda)\Delta)) = 0$, with Δ the combined uncertainty, is not possible, while the structured singular value (μ) framework precisely relies on such a reformulation. It should be remarked that when considering the delays as unperturbed but allowing the perturbations on different matrices to be structured, *independently* of each other, as we shall do, the characteristic equation can still be recast in the above form where Δ is block diagonal with real-valued diagonal blocks. Such a structure is however difficult to exploit in the approximating of structured singular values. In the μ-framework the concept of complex distance to instability for a system with multiple-delays is introduced and developed in [13], where the structure of the equation is exploited but the perturbations on individual coefficient matrices are assumed unstructured; in [7], a formula for the real stability radius of either retarded or neutral system with multiple delays is provided, where the spectral norm is used as a measure of the uncertainty magnitude.

Second, since we aim at computing the exact value of the distance to instability, an over-bounding of the uncertainty on the delay or an embedding into a larger set of perturbations, followed by the application of the small gain theorem or other methods for robust control are not possible, as this will lead to possibly conservative bounds. Using this approach a lower-bound on the stability radius of a time-delay system simultaneously subject to real-valued perturbations on the delay values and to real-valued perturbations on the system matrices is presented in [12]. For the same reason methods relying on the construction of Lyapunov functionals are not suitable.

Third, a direct parametric analysis and the explicit construction of stability charts (see, e.g., [18], [3]) are suitable for uncertain parameters, while they are not adopted towards matrix-valued uncertainty.

The key to tackle the various types of uncertainties simultaneously, and to exploit their structure, is a novel iterative algorithm to compute the (real structured) PSA, which is grounded in geometric optimization, i.e. on gradient-ascent methods defined over matrix manifolds. The pseudospectrum is defined as a set in the complex plane to which eigenvalues can be moved when the system is subject to perturbations whose norm is smaller than a prescribed upper bound (see [21] for an overview). The PSA is the real part of the globally rightmost eigenvalue in the pseudospectrum. As a consequence, the distance to instability is the norm of the smallest perturbation that generates a non-negative PSA. Note that this relation has already been used in [10] for the computation of the unstructured complex stability radius of large-scale nonlinear eigenvalue problems.

In this work we start by deriving optimality conditions for the rightmost point of the real structured ε-pseudospectrum, where ε defines the upper bound on the norm of perturbations. We demonstrate that it is possible to express the set of optimal matrix perturbations (i.e. perturbations that induce the globally rightmost point of the pseudospectrum) as a set of low-rank matrices. These surprising properties allow us to develop gradient based optimization algorithms, where we can exploit the low rank property. Using ideas tracing back to [9], the resulting algorithm relies on the construction of a steepest ascent differential equation in the space of perturbations, aimed at maximizing the real part of the rightmost eigenvalue. Although its global convergence cannot be theoretically demonstrated, this tech-

F. Borgioli and W. Michiels are with the Department of Computer Science, KU Leuven, 3001 Leuven, Belgium, e-mail: {first-name.lastname}@kuleuven.be.

nique has already proved to be efficient in the computation of PSA for unstructured complex-valued ([4]) and real-valued ([5]) perturbations of a standard eigenvalue problem. The main advantage of a gradient flow discretization, which is closely related to the application of a first-order optimization method with adaptive stepsize, over methods based on a fixed-point iteration induced by the optimality conditions (see [14] for the complex PSA) is that the monotonicity property of the real part of the rightmost eigenvalue along the flow is favorable in terms of local convergence properties and globalization strategy.

The paper is structured as follows: Section II introduces the notion of distance to instability and its connection with the PSA; Section III illustrates the algorithm for the PSA computation; Section IV illustrate the bisection-Newton's method used to compute the distance to instability. Examples and concluding remarks are respectively shown in Sections V and VI. Some preliminary results, without taking into account dependencies between perturbations, are presented in [1].

II. Problem statement and main assumptions

Let us consider system (1), assuming that its zero solution is exponentially stable. Subsequently we assume that matrices A_i are affected by K *independent real-valued* time-invariant uncertainties δA_j. Uncertainties δA_j may perturb either a whole matrix or a submatrix, for instance a single block or a single coefficient, and as a consequence their dimension may differ from one to another. Furthermore, we assume that each delay value τ_i is potentially subject to a perturbation $\delta\tau_i$ such that $|\delta\tau_i| < \tau_i$. The delay eigenvalue problem (DEP) associated with the perturbed system then reads

$$\underbrace{\left(\lambda I - \sum_{i=0}^{m}\left(A_i + \sum_{j=1}^{K} B_{i,j}\delta A_j C_{i,j}\right)e^{-\lambda(\tau_i+\delta\tau_i)}\right)}_{:=M(\lambda,\Delta)y} y = 0, \quad (2)$$

where $y \in \mathbb{C}^n,\ \lambda \in \mathbb{C}$, matrices A_i are defined as in (1) and, for the sake of compactness, we have adopted $0 = \tau_0 < \tau_1 < \cdots < \tau_m$; each nominal matrix A_i can be perturbed by one or more uncertainties δA_j, with the structures (and the weights) of these perturbations defined by auxiliary matrices $B_{i,j}$ and $C_{i,j}$. Matrices $B_{i,j}$ and $C_{i,j}$ are real-valued and their dimensions depend on the dimension of the associated uncertainty δA_j: then, for each $\delta A_j \in \mathbb{R}^{p_j\times q_j}$, we have $B_{i,j} \in \mathbb{R}^{n\times p_j},\ C_{i,j} \in \mathbb{R}^{q_j\times n}$ for all $i = 0,\dots,m$. Intuitively, if matrix A_i is not affected by δA_j, then we set $B_{i,j} = 0$ and $C_{i,j} = 0$; similarly, as $\tau_0 = 0$, then also $\delta\tau_0 = 0$. The combined perturbation is denoted by $\Delta = (\Delta A, \Delta\tau) = (\delta A_1,\dots,\delta A_K, \delta\tau_1,\dots,\delta\tau_m) \in \mathbb{R}^*\times\mathbb{R}^m$, with $\mathbb{R}^* = \mathbb{R}^{p_1\times q_1}\times\cdots\times\mathbb{R}^{p_K\times q_K}$. To motivate the adopted model we present an example.

Example 1. *We consider a simple model for a turning process as presented in [8],*

$$\ddot{x}(t) + 2\xi\omega\dot{x}(t) + \omega^2 x(t) = \frac{k}{m}\left(x(t-\tau) - x(t)\right), \quad (3)$$

whose corresponding unperturbed DEP (after rewriting into a set of first order differential equations) reads

$$\left(\lambda I - A_0 - A_1 e^{-\lambda\tau}\right) y = 0,$$

where

$$A_0 = \begin{bmatrix} 0 & 1 \\ -\omega^2 - \frac{k}{m} & -2\xi\omega \end{bmatrix},\ A_1 = \begin{bmatrix} 0 & 0 \\ \frac{k}{m} & 0 \end{bmatrix}, \quad (4)$$

with ω the natural frequency, ξ the damping ratio, m the modal mass and k the cutting force coefficient. Considering system (1), we then have $\tau_1 = \tau$. Let us now assume that parameters k and ξ are respectively affected by $\delta k = \delta A_1$ and $\delta\xi = \delta A_2$, and the delay τ_1 also admits a perturbation $\delta\tau_1$. Then, the perturbed DEP reads

$$M(\lambda,\Delta) = \lambda I - \Big(A_0 + B_{0,1}\delta k C_{0,1} + B_{0,2}\delta\xi C_{0,2}\Big) - \Big(A_1 + B_{1,1}\delta k C_{1,1} e^{-\lambda(\tau_1+\delta\tau_1)}\Big), \quad (5)$$

where

$$B_{1,1} = \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix}, C_{1,1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}^T, B_{0,2} = \begin{bmatrix} 0 \\ -2\omega \end{bmatrix}, C_{0,2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}^T$$

and $B_{0,1} = -B_{1,1},\ C_{0,1} = C_{1,1}$.

We aim at computing the real structured distance to instability (stability radius) of problem (2). As a consequence of the continuous dependence of the spectral abscissa w.r.t. the coefficient matrices and the delays [15], a formal definition of distance to instability reads

$$r_{\mathbb{R}} := \inf_{\lambda=\jmath\omega,\ \omega\geq 0}\ \inf_{\Delta}\left\{\|\Delta\| :\ \det\left(M(\lambda,\Delta)\right) = 0\right\}. \quad (6)$$

Intuitively, we look for the minimum perturbation Δ such that the perturbed DEP in (2) has one or more imaginary axis eigenvalues. From this definition, we clearly need to introduce an overall measure on the uncertainties affecting the system. Introducing weights $v_i \in \mathbb{R}_0^+\cup\{+\infty\}$ for $i = 1,\dots,m$, and denoting the Frobenius norm by $\|G\|_F = \sqrt{\sum_i\sum_j G_{ij}^2}$ for matrix $G = [G_{ij}]_{i,j}$, we arrive at the following global measure of the uncertainty,

$$\|\Delta\| := \left\|\begin{bmatrix}\|\delta A_1\|_F & \dots & \|\delta A_K\|_F & v_1|\delta\tau_1| & \dots & v_m|\delta\tau_m|\end{bmatrix}^T\right\|_\infty. \quad (7)$$

With this definition, asking for an ε-bounded set of perturbations, thus such that $\|\Delta\| \leq \varepsilon$, is equivalent to imposing

$$\|\delta A_j\|_F \leq \varepsilon,\ v_i|\delta\tau_i| \leq \varepsilon, \quad j = 1,\dots,K,\ i = 1,\dots,m.$$

As a consequence, setting a weight $v_i = +\infty$ implies that $\delta\tau_i = 0$ and that the τ_i is assumed free from uncertainties.

Given the above measure, for a fixed ε we can now define the structured ε-pseudospectrum Λ_ε and the ε-PSA α_ε as follows,

$$\Lambda_\varepsilon := \bigcup_{\substack{\Delta\in\mathbb{R}^*\times\mathbb{R}^m \\ \|\Delta\|\leq\varepsilon}} \left\{\lambda\in\mathbb{C} : \det M(\lambda,\Delta) = 0\right\}, \quad (8)$$

$$\alpha_\varepsilon := \max\{\Re(\lambda) : \lambda\in\Lambda_\varepsilon\}. \quad (9)$$

Generally, the PSA is defined as the supremum of the real part of the points in the pseudospectrum; since this work deals with DEPs of retarded type, we know from the theory that the supremum coincides with the maximum function (see for instance Chapter 1 of [15]).

Now, given the definition of PSA and its continuity w.r.t. to ε (which we will prove in Section IV), it comes naturally that the distance to instability in (6) can be equivalently defined as follows

$$r_{\mathbb{R}} := \inf\{\varepsilon : \alpha_\varepsilon \geq 0\}. \quad (10)$$

III. Computing the PSA: an iterative method

In this section we illustrate the theoretical details and the implementation of the algorithm to compute the PSA for a DEP like in (2); the computation of the stability radius will then ground on this algorithm. In the rest of the paper we make the following assumption.

Assumption 1. *The globally rightmost eigenvalue of the ε-pseudospectrum Λ_ε is simple.*

The next lemma provides the well known formula to compute the derivative of a simple eigenvalue, and we refer the reader interested in a proof to Lemma 2.7 in [17].

Lemma 1. *Let $F(\lambda,\theta) : \mathbb{C}\times\mathbb{C}^d \longrightarrow \mathbb{C}^{n\times n}$ be a continuously differentiable function depending on a set of parameters θ, and let λ be a simple root of the eigenvalue problems $F(\lambda,\theta)y = 0$ and $x^*F(\lambda,\theta) = 0$, with $x,\ y \in \mathbb{C}^n$ with unit norm, and x^* the conjugate of x, then*

$$\frac{\partial\lambda}{\partial\theta_i} = -\frac{x^*\frac{\partial F(\lambda,\theta)}{\partial\theta_i}y}{x^*\frac{\partial F(\lambda,\theta)}{\partial\lambda}y}, \qquad i = 1,\dots,d. \tag{11}$$

Let $\lambda(\Delta)$ be the rightmost eigenvalue of DEP (2) perturbed with Δ; then we can compute the PSA by solving the optimization problem

$$\begin{array}{ll}\max & \Re\lambda(\Delta)\\ \text{s.t.} & \Delta\in\mathbb{R}^*\times\mathbb{R}^m,\ \|\Delta\|\le\varepsilon.\end{array} \tag{12}$$

The next theorem defines a few properties of the optimal perturbations, namely the matrix and the delay perturbations that generate the globally rightmost point of the pseudospectrum (8), which will be then fully exploited to reduce the search space of optimization problem (12). Before we state the theorem, let us define with k_j the number of nominal matrices A_i affected by uncertainty δA_j.

Theorem 1. *Let $\lambda_{\rm RM}$ be a globally rightmost point of the structured ε-pseudospectrum generated by some perturbation $\Delta = (\Delta A, \delta\tau)$. Let also x, y be respectively the left and the right eigenvectors of $\lambda_{\rm RM}$ (with unitary norms) uniquely defined by condition*

$$\xi := -x^*Dy > 0 \tag{13}$$

where

$$D = I + \sum_{i=0}^{m}\left(A_i + \sum_{j=1}^{K} B_{i,j}\delta A_j C_{i,j}\right)(\tau_i+\delta\tau_i)e^{-\lambda_{\rm RM}(\tau_i+\delta\tau_i)}.$$

Then the following properties hold:

(i) There always exists a set of ε-bounded perturbations $\widetilde{\Delta A} = (\widetilde{\delta A_1},\dots,\widetilde{\delta A_K})$, such that for each $j = 1,\dots,K$

$$\operatorname{rank}(\widetilde{\delta A_j}) \le \rho_j^* := \min(2k_j, p_j, q_j)$$

and for which the rightmost eigenvalue is $\lambda_{\rm RM}$;

(ii) Let us define

$$X = [\Re(x)\ \Im(x)],\ Y = [\Re(y)\ \Im(y)],$$

with $\Re$, $\Im$ the real and the imaginary part respectively, and

$$\Gamma_i = \begin{bmatrix} \Re(e^{-(\tau_i+\delta\tau_i)\lambda_{\rm RM}}) & -\Im(e^{-(\tau_i+\delta\tau_i)\lambda_{\rm RM}}) \\ \Im(e^{-(\tau_i+\delta\tau_i)\lambda_{\rm RM}}) & \Re(e^{-(\tau_i+\delta\tau_i)\lambda_{\rm RM}}) \end{bmatrix}$$

for $i = 0,\dots,m$. If for some $j\in\{1,\dots,K\}$ matrix $\sum_{i=0}^m B_{i,j}^T X\Gamma_i Y^T C_{i,j}^T$ is nonzero, then the optimal perturbation δA_j satisfies

$$\delta A_j = -\varepsilon\frac{\sum_{i=0}^m B_{i,j}^T X\Gamma_i Y^T C_{i,j}^T}{\left\|\sum_{i=0}^m B_{i,j}^T X\Gamma_i Y^T C_{i,j}^T\right\|_F}. \tag{14}$$

(iii) If the derivative $\frac{\partial\Re(\lambda_{\rm RM})}{\partial\delta\tau_i}$, which reads

$$\frac{1}{\xi}\Re\left(x^*\left(A_i+\sum_{j=1}^{K}B_{i,j}\delta A_i C_{i,j}\right)\lambda_{\rm RM}e^{-\lambda_{\rm RM}(\tau_i+\delta\tau_i)}y\right), \tag{15}$$

is nonzero, then the optimal time-delay perturbation is such that $v_i|\delta\tau_i| = \varepsilon$.

Proof. (i) In order to simplify the notation, let us assume, without losing generality, that each uncertainty δA_j perturbs the first k_j nominal matrices. Take any $j\in\{1,\dots,k\}$. Trivially, the relation $\operatorname{rank}(\delta A_j)\le\min\{p_j,q_j\}$ always holds. Then, let us assume that $\rho_j^* = 2k_j$. Let $r_j^*\le 2k_j$ be the rank of the space spanned by vectors in $Z := \{\Re(C_{1,j}y),\Im(C_{1,j}y),\dots,\Re(C_{k_j,j}y),\Im(C_{k_j,j}y)\}$; we define a matrix H_j such that the first r_j^* columns' span includes the space spanned by vectors in Z, and the other columns complete the basis of $\mathbb{R}^{q_j\times q_j}$, so that H_j is orthogonal. Thus, we can write $\delta A_j = \widehat{\delta A_j}H_j^T$ for some matrix $\widehat{\delta A_j}$. We indicate with $H_j^{(s)}$ the s-th column of matrix H_j and, in order to simplify the notation, we write $\widetilde{\tau_i} = \tau_i + \delta\tau_i$. With this factorization, we have

$$\begin{aligned}
0 &= \left(\lambda_{\rm RM}I - \sum_{i=0}^m\left(A_i+\sum_{j=1}^K B_{i,j}\delta A_j C_{i,j}\right)e^{-\lambda_{\rm RM}\widetilde{\tau_i}}\right)y = \\
&\left(\lambda_{\rm RM}I - \sum_{i=0}^m A_i e^{-\lambda_{\rm RM}\widetilde{\tau_i}}\right)y \\
&-\sum_{i=0}^m\sum_{j=1}^K B_{i,j}\widehat{\delta A_j}H_j^T e^{-\lambda_{\rm RM}\widetilde{\tau_i}}\left(\Re(C_{i,j}y)+j\Im(C_{i,j}y)\right) = \\
&\left(\lambda_{\rm RM}I - \sum_{i=0}^m A_i e^{-\lambda_{\rm RM}\widetilde{\tau_i}}\right)y \\
&-\sum_{i=0}^m\sum_{j=1}^K B_{i,j}\widehat{\delta A_j}e^{-\lambda_{\rm RM}\widetilde{\tau_i}}\begin{pmatrix}\langle H_j^{(1)^T},\Re(C_{i,j}y)+j\Im(C_{i,j}y)\rangle\\ \vdots\\ \langle H_j^{(r_j^*)^T},\Re(C_{i,j}y)+j\Im(C_{i,j}y)\rangle\\ 0\\ \vdots\\ 0\end{pmatrix},
\end{aligned}$$

where $\langle\cdot,\cdot\rangle$ indicates the Euclidean inner product between two vectors. Thanks to the last expression, we can get rid of the last $q_j - r_j^*$ columns of H_j. Thus, for every $j = 1,\dots,K$ we define a new matrix $H_j^\star$ by resetting all elements in H_j to zero except for the elements in the first r_j^* columns. Then, we set matrices $\widetilde{\delta A_j} = \widehat{\delta A_j}{H_j^\star}^T$ with rank r_j^* and that preserve $\lambda_{\rm RM}$ and its right eigenvector y.
We still have to prove that $\left\|\widetilde{\delta A_j}\right\|_F \le \|\delta A_j\|_F$; this is equivalent to proving that the Euclidean norm on each row of δA_j is larger than the Euclidean norm of the same row in $\widetilde{\delta A_j}$. Without loss of generality let us investigate the Euclidean norms of the first row in each matrix. Let $\left(\widehat{\delta A_j}\right)_{(1)} = \begin{bmatrix}\beta_1 & \dots & \beta_{q_j}\end{bmatrix}$ be the first row in $\widehat{\delta A_j}$ and the matrices H_j and $H_j^\star$ defined as follows

$$H_j = \begin{bmatrix}p_1^T\ p_2^T\ p_3^T\ \dots\ p_{q_j}^T\end{bmatrix},\ H_j^\star = \begin{bmatrix}p_1^T\ \dots\ p_{r_j^*}^T\ 0\ \dots\ 0\end{bmatrix},$$

with $p_i\in\mathbb{R}^{q_j},\ i = 1,\dots,q_j$ orthonormal vectors. Thus

$$\begin{aligned}\left(\delta A_j\right)_{(1)} &= \left(\widehat{\delta A_j}\right)_{(1)}H_j^T = \beta_1 p_1^T + \dots + \beta_{q_j}p_{q_j}^T,\\ \left(\widetilde{\delta A_j}\right)_{(1)} &= \left(\widehat{\delta A_j}\right)_{(1)}{H_j^\star}^T = \beta_1 p_1^T + \dots + \beta_{r_j^*}p_{r_j^*}^T.\end{aligned}$$

Since $p_1,\dots,p_{q_j}$ are orthonormal, it follows that

$$\left\|\left(\widetilde{\delta A_j}\right)_{(1)}\right\|_2 \le \left\|\left(\delta A_j\right)_{(1)}\right\|_2.$$

(ii) Let us indicate with $(\delta A_j)_{s,t}$ the coefficient in position (s,t) of matrix δA_j, and with $B_{i,j}^{(s)}, C_{i,j(t)}$ respectively the s-th column of $B_{i,j}$ and the t-th row of $C_{i,j}$, for $s = 1,\dots,p_j,\ t = 1,\dots,q_j$. From Lemma 1, using some algebra manipulations, we obtain that

$$\begin{aligned}
\frac{\partial\Re(\lambda_{\rm RM})}{\partial\delta A_j} &= \left(\frac{\partial\Re(\lambda_{\rm RM})}{\partial(\delta A_j)_{s,t}}\right)_{\substack{s=1,\dots,p_j\\ t=1,\dots,q_j}} = \\
&-\frac{1}{\xi}\sum_{i=0}^m\Re\left(x^*B_{i,j}^{(s)}C_{i,j(t)}ye^{-\lambda_{\rm RM}(\tau_i+\delta\tau_i)}\right)_{\substack{s=1,\dots,p_j\\ t=1,\dots,q_j}} = \\
&-\frac{1}{\xi}\sum_{i=0}^m\left(B_{i,j}^{(s)T}\Re\left(xy^*e^{-\lambda_{\rm RM}(\tau_i+\delta\tau_i)}\right)C_{i,j(t)}^T\right)_{\substack{s=1,\dots,p_j\\ t=1,\dots,q_j}} = \\
&-\frac{1}{\xi}\sum_{i=0}^m B_{i,j}^T X\Gamma_i Y^T C_{i,j}^T.
\end{aligned}$$

Since each δA_j is ε-bounded, the following constraints hold,

$$g_j := \sum_{\substack{s=1,\dots,p_j \\ t=1,\dots,q_j}} \left(\delta A_j\right)_{s,t}^2 - \varepsilon^2 \leq 0, \quad j = 1, \dots, K.$$

As λ_{RM} is the globally rightmost point of the pseudospectrum, from the theory of Lagrange multipliers we have that

$$\frac{\partial \Re(\lambda_{\mathrm{RM}})}{\partial \delta A_j} - 2\mu_j \delta A_j = 0, \quad j = 1, \dots, K$$

where $\mu_j \geq 0$ are the multipliers associated with the g_j inequality constraints. If $\sum_{i=0}^{m} B_{i,j} X \Gamma_i Y^T C_{i,j}^T$ is nonzero, then μ_j is strictly positive and g_j is an active constraint; therefore δA_j is a positive multiple of $\frac{\partial \Re(\lambda_{\mathrm{RM}})}{\partial \delta A_j}$ and from the condition $\|\delta A_j\|_F = \varepsilon$ we derive the thesis.
(iii) The same reasoning used in point (*ii*) can be applied here; if $\frac{\partial \Re(\lambda_{\mathrm{RM}})}{\partial \delta \tau_i} \neq 0$, then the constraint on the size of $\delta\tau_i$ must be active. As a consequence, the optimal $\delta\tau_i$ is either $-\frac{\varepsilon}{v_i}$ or $+\frac{\varepsilon}{v_i}$. □

Assertion (*i*) of the theorem thus proves that the PSA is always generated by a set of matrices perturbations whose rank is independent of n, hence for high-dimensional systems we can restrict ourselves to low rank perturbations; even more, the space of perturbations where we look for the maximizer of α_ε can be reduced, and we can define a new optimization problem

$$\begin{aligned} \max \quad & \Re\lambda(\Delta) \\ \text{s.t.} \quad & \Delta \in \mathcal{S} := \mathcal{S}_F \times \mathcal{S}_\tau \end{aligned} \tag{16}$$

where

$$\begin{aligned} \mathcal{S}_F :=& \{\Delta A \in \mathbb{R}^* : \operatorname{rank}(\delta A_j) \leq \rho_j^* := \min(2k_j, p_j, q_j), \\ & \|\delta A_j\|_F \leq \varepsilon,\ j = 1, \dots, K\}, \\ \mathcal{S}_\tau :=& \prod_{i=1}^{m} \left[-\frac{\varepsilon}{v_i}, +\frac{\varepsilon}{v_i}\right]. \end{aligned} \tag{17}$$

Assertion (*ii*) of the theorem also characterizes the optimal matrices perturbations when these have maximum norm and the constraint on the norm is active. However, surprisingly, the optimal perturbations can also correspond to a stationary point of the spectral abscissa in the space of perturbations: in the following we provide an example where the optimal matrix perturbation does not have maximum norm.

Example 2. *Let us consider the following perturbed delay system*

$$\dot{x}(t) = A_1 x(t) + (A_2 + B\delta a C)\, x(t - \tau), \tag{18}$$

where δa is a scalar, the delay τ is set equal to 0.21, $A_1 = -0.0225\ I_4$ (where I_4 is the identity matrix with dimension 4) and

$$A_2 = \frac{0.995}{100} \begin{bmatrix} 0 & 4 & 0 & 0 \\ -301 & -56 & 301 & 224 \\ 0 & 0 & 0 & 16 \\ 1.9531 & 109.375 & -3.9063 & -437.5 \end{bmatrix}.$$

The structured perturbation on matrix A_2 is defined by

$$B = \begin{bmatrix} 0 & 3.1102 & 0 & 0 \end{bmatrix}^T, \quad C = \begin{bmatrix} 1.6 & 0 & -1.6 & 0 \end{bmatrix}.$$

We also assume that the delay τ is unperturbed. In Figure 1 we represent the root locus of the rightmost eigenvalue of system (18) with the uncertain scalar parameter δa varying in modulus between 0 and 0.6. It is easy to observe that the rightmost point of the locus (real part ≈ -0.01) is generated by a perturbation δa whose modulus is between 0.1 and 0.4, while a perturbation such that $\delta a = \pm 0.6$ generates a rightmost eigenvalue with a smaller real part. Hence, for $\varepsilon = 0.6$ the PSA α_ε is not generated by a perturbation of size ε.

In order to find the globally rightmost point of the pseudospectrum we have to compute the global optimum of problem (16). For this

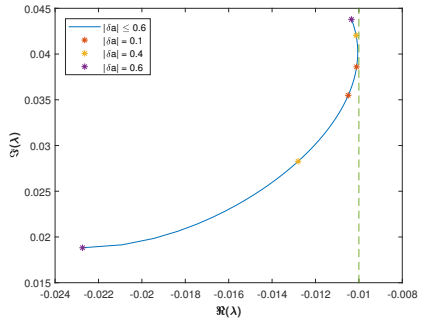


Figure 1: The root locus represents the rightmost eigenvalue of system (18) under different values for the scalar perturbation δa.

we extend a state-of-the art method for computing the PSA for the standard eigenvalue problem (see [5] and the references therein). Even though it cannot be guaranteed that the global optimum is found, the method turns out to be very robust in practice and is considered a standard method for this type of problems. In what follows we focus on the extensions and differences in induced by delay equations.
First, we would like to move on some continuous and differentiable a.e. path in $\mathcal{S}$ along which the spectral abscissa of the perturbed DEPs is monotonically increasing. Hence, we consider the set of perturbations $\Delta = (\Delta A, \Delta\tau) \in \mathcal{S}$ as depending on a parameter s. The next proposition provides a decomposition of $\Delta A(s)$; from now on, we will denote with "dot-quantities" ($\delta\dot{A}_j, \dots$) the derivatives w.r.t. parameter s.

Proposition 1. *Let the function $\mathbb{R}^+ \ni s \mapsto \delta A_j(s) \in \mathbb{R}^{p_j \times q_j}$ be continuous and differentiable a.e. for all s. Then we can always decompose $\delta A_j(s)$ as follows*

$$\delta A_j(s) = -\varepsilon U_j(s) Q_j(s) V_j(s)^T,$$

where $U_j(s) \in \mathbb{R}^{p_j \times \rho_j^*}$, $V_j(s) \in \mathbb{R}^{q_j \times \rho_j^*}$ have orthonormal columns and are s.t. $U_j(s)^T \dot{U}_j(s) = V_j(s)^T \dot{V}_j(s) = 0$, and $Q_j(s) \in \mathbb{R}^{\rho_j^* \times \rho_j^*}$ is s.t. $\|Q_j(s)\|_F \leq 1$ for all $j = 1, \dots, K$.

Proof. In order to simplify the notation in this proof, let us neglect the dependence on the index j. Let us consider the compact singular value decomposition (SVD) of δA

$$\delta A(s) = -\varepsilon \widehat{U}(s) \widehat{\Sigma}(s) \widehat{V}(s)^T,$$

where $\widehat{U}(s), \widehat{V}(s) \in \mathbb{R}^{p_j \times \rho_j^*}$ have orthonormal columns and $\widehat{\Sigma}(s) \in \mathbb{R}^{\rho_j^* \times \rho_j^*}$ is a diagonal matrix such that $\left\|\widehat{\Sigma}(s)\right\|_F \leq 1$. By Theorem 4.3 in [2], the factors of the SVD are also continuous and differentiable a.e. Taking into account the orthonormality of the columns of $\widehat{U}$, we have that

$$\dot{\widehat{U}}(s)^T \widehat{U}(s) + \widehat{U}(s)^T \dot{\widehat{U}}(s) = 0, \ \forall\ s \geq 0,$$

thus $\widehat{U}(s)^T \dot{\widehat{U}}(s)$ is a $\rho_j^* \times \rho_j^*$ skew-symmetric matrix, as well as $\widehat{V}(s)^T \dot{\widehat{V}}(s)$.

Let $Q_1(s)$ be a solution of the following Cauchy problem

$$\dot{Q}_1(s) = -Q_1(s)^{-T} [\widehat{U}(s)^T \dot{\widehat{U}}(s)]^T,\ Q_1(0) = I_{\rho_j^*}, \tag{19}$$

defined in each interval where $\delta A(s)$ is continuous and differentiable. From (19) we can conclude the following: $Q_1^T(s) Q_1(s) = I_{\rho_j^*}$ implies that $Q_1^T(s) \dot{Q}_1(s)$ is skew-symmetric, which implies that $\frac{d}{ds}\left(Q_1^T(s) Q_1(s)\right) = 0$. As for $s = 0$, we have $Q_1(0) = I_{\rho_j^*}$, then $Q_1(s)$ is orthogonal for all $s \geq 0$. If we now define $U(s) := \widehat{U}(s) Q_1(s)^T$ (and we omit parameter s for simplicity), we have

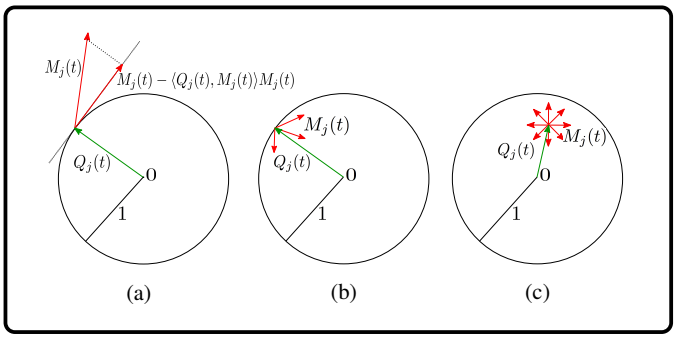


Figure 2: Three possible configurations for the choice of $\dot{Q}_j(t)$ in (21).

$$\begin{aligned} U^T\dot{U} &= Q_1\widehat{U}^T\dot{U} = Q_1\dot{Q}_1^T + Q_1\widehat{U}^T\dot{\widehat{U}}Q_1^T \\ &= -Q_1[\widehat{U}^T\dot{\widehat{U}}]Q_1^{-1} + Q_1\widehat{U}^T\dot{\widehat{U}}Q_1^T = 0. \end{aligned}$$

Thus, we proved that matrix $\widehat{U}(s) \in \mathbb{R}^{p_i\times 2}$ can be decomposed in the product of a matrix $U(s) \in \mathbb{R}^{p_i\times\rho_j^*}$ such that $U(s)^T\dot{U}(s) = 0$ and an orthogonal matrix $Q_1(s) \in \mathbb{R}^{\rho_j^*\times\rho_j^*}$. By applying the same reasoning to matrix $\widehat{V}(s)$ (using factors $V(s)$ and $Q_2(s)$) , we have that (neglecting again parameter s)

$$\delta A = -\varepsilon\widehat{U}\widehat{\Sigma}\widehat{V}^T = -\varepsilon UQ_1\widehat{\Sigma}Q_2V^T = -\varepsilon UQV^T,$$

where $\|Q(s)\|_F \leq 1$ because Frobenius norm is invariant to orthogonal transformations. □

With this result we can express a set of perturbations $\Delta(s)$ as

$$\begin{array}{lll} \delta A_j(s) = & -\varepsilon U_j(s)Q_j(s)V_j(s)^T, & s\in\mathbb{R}^+,\ j=1,\dots,K, \\ \delta\tau_i(s) = & \frac{\varepsilon}{v_i}q_i(s), & s\in\mathbb{R}^+,\ i=1,\dots,m, \end{array} \tag{20}$$

where U_j, V_j, Q_j satisfy the properties introduced in Proposition 1, and in addition we impose $|q_i(s)| \leq 1$ for $i = 1,\dots,m$.
Matrices U_j, V_j, Q_j and functions q_i can be also defined in an equivalent way. Using the technique proposed in [9], let us introduce some arbitrary matrices $R_j(s) \in \mathbb{R}^{p_j\times\rho_j^*}$, $S_j(s) \in \mathbb{R}^{q_j\times\rho_j^*}$, $M_j(s) \in \mathbb{R}^{\rho_j^*\times\rho_j^*}$ for $j = 1,\dots,K$ and functions $r_i(s)$ for $i = 1,\dots,m$. Let us define $\langle A, B\rangle = \text{Trace}(A^TB)$ as the Frobenius inner product of two matrices. Without losing of generality, it is easy to prove that the properties imposed on the decomposition (20) hold for the solutions of the following differential equations (where s is omitted)

$$\left\{\begin{array}{ll} \dot{U}_j = & \left(I_n - U_jU_j^T\right)R_j, \\ \dot{V}_j = & \left(I_n - V_jV_j^T\right)S_j, \\ \dot{Q}_j = & \begin{cases} M_j - \langle M_j, Q_j\rangle Q_j, & \text{if } \|Q_j\|_F = 1,\ \langle M_j, Q_j\rangle > 0 \\ M_j, & \text{otherwise,} \end{cases} \\ \dot{q}_i = & \begin{cases} 0, & \text{if } |q_i| = 1,\ r_iq_i > 0 \\ r_i, & \text{otherwise.} \end{cases} \end{array}\right. \tag{21}$$

The reason why we have two separate cases in the equations defined for $\dot{Q}_j(s)$ (and for $\dot{q}_i(s)$) is the constraint imposed on its norm.

If $\|Q_j(s)\|_F = 1$ and the intended direction of change, $M_j(s)$, points outwards the ball, then following this direction would lead to a solution $Q_j(s)$ such that $\|Q_j(s+\eta)\|_F > 1$ for $\eta > 0$ sufficiently small, violating the constraint. This can be overcome by projection of $M_j(s)$ on the tangent space to the ball, as shown in Figure 2a. If direction $M(s)$ points inwards the ball, it is an allowed direction (Figure 2b). Finally, if $\|Q_j(s)\|_F < 1$ (Figure 2c) there is no constraint on the rate of change, $\dot{Q}_j(s)$.

In what follows we specify matrices $R_j(s)$, $S_j(s)$, $M_j(s)$ and functions $q_i(s)$ in order to individuate a piecewise continuously differentiable path of perturbations $\Delta(s)$ in $\mathcal{S}$ along which the real part of the corresponding rightmost eigenvalue $\lambda(s)$ is monotonically increasing. In order to impose the monotonicity of the spectral abscissa $\Re\lambda(s)$, we exploit Lemma 1 to compute the derivative w.r.t. s of an eigenvalue whose left and right eigenvectors (with unitary norms) x, y are such that $\xi > 0$ as in equation (13): the derivative reads

$$\frac{d\lambda(s)}{ds} = \frac{1}{\xi}\left\{x^*\left(-\sum_{i=0}^{m}\sum_{j=1}^{K}B_{i,j}\delta\dot{A}_j(s)C_{i,j}e^{-\lambda(\tau_i+\delta\tau_i(s))} + \right.\right.$$

$$\left.\left. + \sum_{i=0}^{m}\left(A_i + \sum_{j=1}^{K}B_{i,j}\delta A_j(s)C_{i,j}\right)\lambda\delta\dot{\tau}_i(s)e^{-\lambda(\tau_i+\delta\tau_i(s))}\right)y\right\}.$$

Substituting in the last formula the expressions in (21) and considering the case when $\dot{Q}_j(s) = M_j(s)$ and $\dot{q}_i(s) = r_i(s)$, we obtain

$$\begin{aligned} \Re(\dot{\lambda}) &= \sum_{i=0}^{m}\sum_{j=1}^{K}\frac{\varepsilon}{\xi}\Re\left(x^*B_{i,j}\dot{U}_jQ_jV_j^TC_{i,j}e^{-\lambda(\tau_i+\delta\tau_i)}y\right) \\ &+ \sum_{i=0}^{m}\sum_{j=1}^{K}\frac{\varepsilon}{\xi}\Re\left(x^*B_{i,j}U_j\dot{Q}_jV_j^TC_{i,j}e^{-\lambda(\tau_i+\delta\tau_i)}y\right) \\ &+ \sum_{i=0}^{m}\sum_{j=1}^{K}\frac{\varepsilon}{\xi}\Re\left(x^*B_{i,j}U_jQ_j\dot{V}_j^TC_{i,j}e^{-\lambda(\tau_i+\delta\tau_i)}y\right) \\ &+ \sum_{i=0}^{m}\frac{\varepsilon}{\xi v_i}\Re\left(x^*\left(A_i + \sum_{j=1}^{K}B_{i,j}\delta A_jC_{i,j}\right)\lambda\dot{q}_ie^{-\lambda(\tau_i+\delta\tau_i)}y\right) \\ &+ \sum_{i=0}^{m}\frac{\varepsilon}{\xi v_i}\Re\left(x^*\left(A_i + \sum_{j=1}^{K}B_{i,j}\delta A_jC_{i,j}\right)\lambda\dot{q}_ie^{-\lambda(\tau_i+\delta\tau_i)}y\right) \\ &= \sum_{i=0}^{m}\sum_{j=1}^{K}\frac{\varepsilon}{\xi}\langle X, B_{i,j}(I_n - U_jU_j^T)R_jQ_jV_j^TC_{i,j}Y\Gamma_i^T\rangle \\ &+ \sum_{i=0}^{m}\sum_{j=1}^{K}\frac{\varepsilon}{\xi}\langle X, B_{i,j}U_jM_jV_j^TC_{i,j}Y\Gamma_i^T\rangle \\ &+ \sum_{i=0}^{m}\sum_{j=1}^{K}\frac{\varepsilon}{\xi}\langle X, B_{i,j}U_jQ_jS_j^T(I_n - V_jV_j^T)C_{i,j}Y\Gamma_i^T\rangle \\ &+ \sum_{i=0}^{m}\frac{\varepsilon}{\xi v_i}\underbrace{\Re\left(x^*\left(A_i + \sum_{j=1}^{K}B_{i,j}\delta A_jC_{i,j}\right)\lambda e^{-\lambda(\tau_i+\delta\tau_i)}y\right)}_{:\beta_i}\dot{q}_i \\ &= \sum_{j=1}^{K}\frac{\varepsilon}{\xi}\left\langle\sum_{i=0}^{m}B_{i,j}^TX\Gamma_iY^TC_{i,j}^TV_jQ_j^T, (I_n - U_jU_j^T)R_j\right\rangle \\ &+ \sum_{j=1}^{K}\frac{\varepsilon}{\xi}\left\langle\sum_{i=0}^{m}U_j^TB_{i,j}^TX\Gamma_iY^TC_{i,j}^TV_j, M_j\right\rangle \\ &+ \sum_{j=1}^{K}\frac{\varepsilon}{\xi}\left\langle S_i, (I_n - V_jV_j^T)\sum_{i=0}^{m}C_{i,j}Y\Gamma_i^TX^TB_{i,j}U_jQ_j\right\rangle \\ &+ \sum_{i=0}^{m}\frac{\varepsilon}{\xi v_i}\beta_i\dot{q}_i, \end{aligned}$$

Therefore, since $(I_n - U_jU_j^T)$, $(I_n - V_jV_j^T)$ are positive semi-definite, a non-negative $\dot{\Re}(\lambda)$ is guaranteed by the following choices

$$\begin{aligned} &R_j = \sum_{i=1}^{m}B_{i,j}^TX\Gamma_iY^TC_{i,j}^TV_jQ_j^T, \quad S_j = \sum_{i=1}^{m}C_{i,j}Y\Gamma_i^TX^TB_{i,j}U_jQ_j, \\ &M_j = U_j^T\sum_{i=1}^{m}B_{i,j}^TX\Gamma_iY^TC_{i,j}^TV_j, \quad r_i = \beta_i. \end{aligned} \tag{22}$$

The other cases for differential equations (21) give rise to the same expressions, and that (22) can be interpreted as a (scaled) gradient.

A. Implementation

As a consequence of the definition of differential equations (21) with choices (22), at each fixed s the solutions $U_j(s)$, $V_j(s)$, $Q_j(s)$, $q_i(s)$ define a new perturbed DEP such that $\Re(\lambda(s))$ is monotonically increasing, where $\lambda(s)$ is the rightmost eigenvalue of the problem. In each iteration we discretize differential equations (21), by performing a forward Euler method with adaptive stepsize together with a projection of the new quantities onto the space of perturbations $\mathcal{S}$, in order to satisfy the properties imposed in (20). The adaptive stepsize guarantees increasing monotonicity of the sequence $\{\Re(\lambda_k)\}_{k\geq 1}$: at each iteration the initial stepsize r_n is reduced by a scaling factor h (that we set equal to 2) until monotonicity is guaranteed; vice versa, if r_n already guarantees monotonicity, an extra computation with a doubled stepsize is carried out: the stepsize that guarantees the rightmost eigenvalue is then adopted. The whole structure is shown in Algorithm 1.

From assertion (*ii*) of Theorem 1, optimal perturbations are built using eigenvalue λ_{RM} and eigenvectors x, y: as in [14], this may be the basis for a fixed-point iteration. Here we use it instead to initialize the algorithm as follows: we compute the rightmost eigenvalue of the nominal problem, and its corresponding left and right eigenvectors. Then we initialize matrices X, Y, Γ_i, for $i = 0, \ldots, m$ as in Theorem 1 and compute matrices δA_j as in (14). We apply a SVD decomposition such that $\delta A_j = W_j \Sigma_j Z_j$ and compute ρ_j^*. Then we set $U_j^{(1)}$, $V_j^{(1)}$ respectively equal to the first ρ_j^* columns of W_j and Z_j, and $Q_j^{(1)}$ equal to Σ_j (eventually normalized). The delays $\tau_i^{(1)}$ are simply initialized following the derivative in equation (15). Finally we construct the perturbed DEP (2) and, once again, we compute the rightmost eigenvalue λ_1 and the corresponding normalized eigenvectors x_1, y_1.

The algorithm converges if the norm of the (projected) gradient of the spectral abscissa is below a given tolerance: clearly this condition also holds on local maxima, and the algorithm cannot distinguish between locally and globally rightmost points. However, in practice, a convergence to a locally rightmost point is seldom observed; moreover, we run the algorithm several times using different eigenvalues (located in the right part of the original spectrum) as initialization. This has already proved to be an effective strategy to converge to the globally rightmost point in [4] and [5]. Finally, the solution of DEPs, which is the most expensive operation in each step, is obtained using the algorithm and the software described in [11]. An analysis of the cost of this operation is instead reported in [22].

Remark 1. *In Figure 3 we plot the complex (red) and the real (blue) ε-pseudospectra for the system*

$$\dot{x}(t) = \begin{bmatrix} -5 & 1 \\ 2 & -6 \end{bmatrix} x(t) + \begin{bmatrix} -2 & 1 \\ 4 & -1 \end{bmatrix} x(t-1),$$

where an unstructured uncertainty affects the first matrix and $\varepsilon = 3.5$. We also represent the iterations performed by Algorithm 1 when it is initialized on the three (in the upper half-plane) rightmost eigenvalues of the nominal system: we observe here that the initialization, performing a sort of fixed-iteration step suggested by (14), pushes us towards the pseudospectrum boundary, close to some locally rightmost point; then, thanks to a few iterations of the forward Euler method, we are able to converge to the maximum. This is the typical behavior of the algorithm, also in cases when the pseudospectrum exhibits a locally and a globally rightmost points with "close" real parts: these are in general expressed by "distant" perturbations in the space of uncertainties, which are reached by initializing Algorithm 1 with different eigenvalues.
From this picture, we also see that the PSA is computed starting from the second rightmost eigenvalue; this highlights the importance of restarting the algorithm using different initializations.

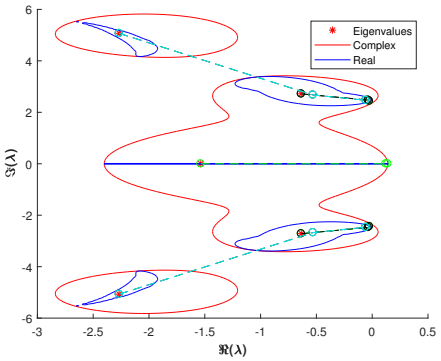


Figure 3: Iterations of Algorithm 1.

Remark 2. *At the beginning of this section we introduced Assumption 1, which was used to demonstrate Theorem 1 (ii) and (iii), whose proofs relied on the concept of derivative. In Theorem 1 (i), we proved that the globally rightmost point can be expressed by a set of perturbations whose rank is independent of n: this still holds for multiple eigenvalues, since eigenvectors x and y are always defined; as a consequence, Algorithm 1 can also be used if Assumption 1 does not hold.*
Let us also consider the case where a perturbed DEP exhibits two different rightmost eigenvalues with "close" real parts and different imaginary parts (the cases where the two real parts exactly coincide form a set of measure zero in the space of uncertainties): since at each iteration we pick the rightmost eigenvalue of the new perturbed DEP (which is not necessarily the same eigenvalue whose real part we were optimizing) and then continue the maximization of its real part, the algorithm is able to "jump" from one "island" of the pseudospectrum to the other. Then, the path we move on remain continuous yet a switch may not be differentiable.

IV. A Newton-bisection method for the stability radius

As stated before we compute the real stability radius of system (1) using the equivalent definition (10), which exploits the continuity of the PSA w.r.t. the matrix coefficients and the delay values; the continuity of the PSA follows from its definition as the maximum of the spectral abscissae of the perturbed DEPs. As stated in Theorem 1.15 in [15], the spectral abscissa of a DEP with retarded dynamics is also continuous w.r.t. the matrix coefficients and the delay values.

Moreover, it is easy to verify that α_ε is a monotone function w.r.t. ε: if $\varepsilon_1 < \varepsilon_2$, $\Lambda_{\varepsilon_1} \subseteq \Lambda_{\varepsilon_2}$ must hold (the set of allowed perturbations is enlarged); then, from the property of maximum function, the PSA α_ε is nondecreasing w.r.t. ε. Hence, the distance to instability is the root ε^* of the equation $\alpha_{\varepsilon^*} = 0$.

We want to use Newton-bisection method to find the unique root of this nonlinear equation. However, in order to use Newton's method, at least almost differentiability of the function is required. The PSA function α_ε is bounded on any interval $[\varepsilon_L, \varepsilon_R]$, thus it is a function of bounded variation; from the monotonicity property and from the Lebesgue's differentiation theorem of functions of bounded variation (see for instance Theorem 3.4 of [20]) we can conclude that α_ε is almost everywhere differentiable w.r.t. ε; in particular, it is differentiable whenever there exists one and one only globally rightmost point (and this is simple, by Assumption 1). The next theorem then provides an explicit expression for the derivative.

Algorithm 1 Computation of the real PSA via forward Euler method applied to equations (21).

Input: Problem data: $A_i, B_i, C_i, \tau_i, \varepsilon,$ tolerance η, scaling factor $h > 1$.

Initialization data: $U_j^{(1)}, Q_j^{(1)}, V_j^{(1)}, \tau_i^{(1)}, \lambda_1, x_1, y_1, r_1$

Set $\gamma = +\infty$, $n = 1$.
while $\gamma > \eta$ **do**
  Compute $X,\ Y,\ \Gamma_i$ as in Theorem 1.
  Compute the derivatives $\dot{U}_j^{(n)},\ \dot{V}_j^{(n)},\ \dot{Q}_j^{(n)}$, for $j = 1, \dots, K$ and $\dot{q}_i^{(n)}$ for $i = 1, \dots, m$ using differential equations (21) with choices (22).
  Compute

$$\begin{cases} \widehat{U}_j^{(n+1)} = U_j^{(n)} + r_n \dot{U}_j^{(n)}, & j = 1, \dots, K \\ \widehat{V}_j^{(n+1)} = V_j^{(n)} + r_n \dot{V}_j^{(n)}, & j = 1, \dots, K \\ \widehat{Q}_j^{(n+1)} = Q_j^{(n)} + r_n \dot{Q}_j^{(n)}, & j = 1, \dots, K \\ \widehat{\tau}_i^{(n+1)} = \tau_i^{(n)} + r_n \dot{q}_i^{(n)}, & i = 1, \dots, m. \end{cases}$$

  Compute the compact QR decompositions such that $\widehat{U}_j^{(n+1)} = Q_j^U R_j^U, \widehat{V}_j^{(n+1)} = Q_j^V R_j^V$.
  For $j = 1, \dots, K$, set $U_j^{(n+1)} = Q_j^U,\ V_j^{(n+1)} = Q_j^V$ and

$$Q_j^{(n+1)} = \begin{cases} \widehat{Q}_j^{(n+1)}, & \text{if } \left\| \widehat{Q}_j^{(n+1)} \right\|_F \leq 1, \\ \dfrac{\widehat{Q}_j^{(n+1)}}{\left\| \widehat{Q}_j^{(n+1)} \right\|_F}, & \text{otherwise.} \end{cases}$$

  Then set $\delta A_j = -\varepsilon U_j^{(n+1)} Q_j^{(n+1)} V_j^{(n+1)T}$.
  For $i = 1, \dots, m$, possibly rescale $\widehat{\tau}_i^{(n+1)}$ to an ε-bounded perturbation of delay values:

$$\begin{cases} \tau_i^{(n+1)} = \min\left(\tau_i + \frac{\varepsilon}{v_i}, \widehat{\tau}_i^{(n+1)}\right), \text{ if } \dot{q}_i^{(n)} \geq 0 \\ \tau_i^{(n+1)} = \max\left(0, \tau_i - \frac{\varepsilon}{v_i}, \widehat{\tau}_i^{(n+1)}\right), \text{ if } \dot{q}_i^{(n)} < 0 \end{cases}$$

  Compute the rightmost eigenvalue $\hat{\lambda}$ of (2), and the normalized left and right eigenvectors $\hat{x}, \hat{y}$.
  Set $r = r_n$.
  while $\Re(\hat{\lambda}) \leq \Re(\lambda_n)$ **do**
    Set $r = r/h$ and repeat steps from 5 to 8 with step r.
    Compute the rightmost eigenvalue $\hat{\lambda}$ of (2), and the normalized left and right eigenvectors $\hat{x}, \hat{y}$.
  end while
  if $r < r_n$ **then**
    Set $\lambda_{n+1} = \hat{\lambda}, x_{n+1} = \hat{x}, y_{n+1} = \hat{y}, r_{n+1} = r$
  else
    Set $\tilde{r} := h r_n$ and repeat steps from 5 to 8 with step $\tilde{r}$.
    Compute the rightmost eigenvalue $\tilde{\lambda}$ of (2), and the normalized left and right eigenvectors $\tilde{x}, \tilde{y}$.
    if $\Re(\hat{\lambda}) \leq \Re(\tilde{\lambda})$ **then**
      Set $\lambda_{n+1} = \tilde{\lambda}, x_{n+1} = \tilde{x}, y_{n+1} = \tilde{y}, r_{n+1} = \tilde{r}$
    else
      Set $\lambda_{n+1} = \hat{\lambda}, x_{n+1} = \hat{x}, y_{n+1} = \hat{y}, r_{n+1} = r_n$
    end if
  end if
  Set $\gamma = \Re(\dot{\lambda}_{n+1})$ and $n \to n + 1$.
end while
return $\alpha_\varepsilon = \Re(\lambda_{n+1})$.

Theorem 2. *Let $\left(\lambda_{\mathrm{RM}}(\varepsilon), x(\varepsilon), y(\varepsilon)\right)$ be respectively the globally rightmost point of Λ_ε and its left and right normalized eigenvectors; let also $\widetilde{\Delta}(\varepsilon) = \left(\widetilde{\Delta A}(\varepsilon), \widetilde{\Delta\tau}(\varepsilon)\right)$ be functions such that Assumption 1 is satisfied for all ε with*

$$\begin{array}{llll} \delta A_j = & \widetilde{\delta A_j}(\varepsilon) = & -\varepsilon \widetilde{U_j}(\varepsilon) \widetilde{Q_j}(\varepsilon) \widetilde{V_j}(\varepsilon)^T, & j = 1, \dots, K \\ \delta\tau_j = & \widetilde{\delta\tau_j}(\varepsilon) = & \frac{\varepsilon}{v_j} \widetilde{q_j}(\varepsilon), & i = 1, \dots, m. \end{array}$$

Then defining the constant of normalization ξ as in Theorem 1 we can express the derivative of the PSA w.r.t ε as follows

$$\begin{aligned} \frac{\partial \alpha_\varepsilon}{\partial \varepsilon} = & \frac{\partial \Re(\lambda_{\mathrm{RM}})}{\partial \varepsilon} = \\ & + \frac{1}{\xi} \Re\Bigg[x^* \Bigg(- \sum_{i=1}^{m} \sum_{j=1}^{K} B_{i,j} \frac{\partial \widetilde{\delta A_j}}{\partial \varepsilon} C_{i,j} e^{\left(-\lambda_{\mathrm{RM}}(\tau_i + \widetilde{\delta\tau_i})\right)} \\ & + \sum_{i=1}^{m} \Bigg(A_i + \sum_{j=1}^{K} B_{i,j} \widetilde{\delta A_j} C_{i,j} \Bigg) e^{\left(-\lambda_{\mathrm{RM}}(\tau_i + \widetilde{\delta\tau_i})\right)} \lambda_{\mathrm{RM}} \frac{\partial \widetilde{\delta\tau_i}}{\partial \varepsilon} \Bigg) y \Bigg] \\ & = \frac{1}{\xi} \Re\Bigg[x^* \Bigg(\sum_{i=1}^{m} \sum_{j=1}^{K} B_{i,j} \widetilde{U_j} \widetilde{Q_j} \widetilde{V_j}^T C_{i,j} e^{\left(-\lambda_{\mathrm{RM}}(\tau_i + \frac{\varepsilon}{v_i}\widetilde{q_i})\right)} \\ & + \sum_{i=1}^{m} \Bigg(A_i - \sum_{j=1}^{K} \varepsilon B_{i,j} \widetilde{U_j} \widetilde{Q_j} \widetilde{V_j}^T C_{i,j} \Bigg) e^{\left(-\lambda_{\mathrm{RM}}(\tau_i + \frac{\varepsilon}{v_i}\widetilde{q_i})\right)} \lambda_{\mathrm{RM}} \frac{\widetilde{q_i}}{v_i} \Bigg) y \Bigg], \end{aligned}$$

where we omit the dependence on ε to simplify the notation. Observe that we obtain the last expression because we are dealing with optimal combined perturbations at every fixed ε.

This formula is again derived from Lemma 1. As Newton's method is a local method, we first look for a good starting value that guarantees its convergence. Then, we first perform some iterations of the bisection method until we find an ε_1 whose associated PSA is close enough to 0 (for instance, s.t. $|\alpha_{\varepsilon_1}| \leq 10^{-1}$). Then, we apply Newton's method whose basic iteration reads

$$\varepsilon_{k+1} = \varepsilon_k - \alpha_{\varepsilon_k} \Bigg/ \left. \frac{\partial \alpha_\varepsilon}{\partial \varepsilon} \right|_{\varepsilon = \varepsilon_k}$$

and employs the derivative defined in Theorem 2. Therefore, if Algorithm 1 exactly computes the PSA (from a numerical point of view), then this method returns an exact value for the structured stability radius. In this sense, the result is *not conservative*, as it is grounded in a necessary and sufficient condition for robust stability.

V. Validation

Example 3. *We consider the system introduced in Example 2 and we assume it is still affected by uncertainty δa as in Equation (18): the distance to instability of this system is $r_{\mathbb{R}} = 0.7399$. Note that, consistently with the results shown in Example 2, $r_{\mathbb{R}} > 0.6$; moreover, the instability is a consequence of the shifting of the second rightmost eigenvalue in the original spectrum.*

Example 4. *Let us consider again the cutting machine case examined in Example 1. This time, we only assume an uncertainty δk on the cutting coefficient and an uncertainty $\delta\tau$ on the delay. We want to bound the uncertainties $|\delta k| \leq 1.5 * 10^6, |\delta\tau_1| \leq 1.5 * 10^{-3}$, thus we consider the characteristic equation*

$$\det\left(\lambda I - A_0 - B_{1,1}\delta k C_{1,1} - (A_1 + B_{2,1}\delta k C_{2,1})\, e^{-\lambda(\tau_1 + \delta\tau_1)}\right) = 0$$

where

$$B_{1,1} = -B_{2,1} = \begin{bmatrix} 0 \\ -\frac{1}{m} \end{bmatrix}, \quad C_{1,1} = C_{2,1} = 1.5 * 10^7 \begin{bmatrix} 1 & 0 \end{bmatrix},$$

and we set $\varepsilon = 0.1,\ v = 0.015^{-1}$, with v the weight associated with perturbation $\delta\tau_1$. Observe that we also "hide" a weight equal to $1.5 * 10^7$ in matrices $C_{1,1}, C_{2,1}$ acting on parameter k: this is necessary because the two nominal parameters k and τ have many orders of magnitude of difference. In Figure 4, we represent the stability region of the system in the k-τ space: for each couple (k, τ) below the blue line the system is stable. The red line represents the level set equal to 0 of the ε-PSA (or, equivalently, the level set equal to ε of the distance to instability); by definition, for each point whose $\alpha_\varepsilon = 0$ there exists a $\Delta = (\delta k, \delta\tau_1)$ with $\|\Delta\| = 0.1$ such that the perturbed system is unstable. The rectangular blocks

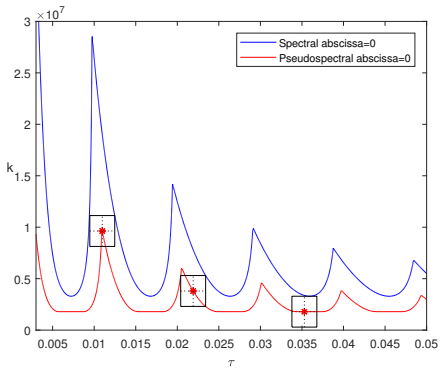


Figure 4: Rectangular blocks represent an ε-bounded set of perturbations for the point marked with a red asterisk; the perturbation has maximum size $1.5 * 10^6$ on the k and $1.5 * 10^{-3}$ on the delay τ.

represent the space of all the ε-bounded perturbations $(\delta k, \delta\tau_1)$ of the point marked with a red asterisk. If we consider the rightmost rectangular block in Figure 4, we see that the PSA is not generated by a maximum-norm perturbation on the delay value: this behavior is equivalent to the one observed in Example 2 for the structured perturbation on a system matrix.

Example 5. *In this example we consider the system from [12]*

$$\dot{x}(t) = A_0x(t) + A_1x(t-\tau_1) + A_2x(t-\tau_2),$$

with

$$A_0 = \begin{bmatrix} -1 & 13.5 & -1 \\ -3 & -1 & -2 \\ -2 & -1 & -4 \end{bmatrix}, A_1 = \begin{bmatrix} -5.9 & 0 & 0 \\ 2 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 7.1 & -70.3 \\ 0 & -1 & 5 \\ 0 & 0 & 6 \end{bmatrix},$$

and $(\tau_1,\ \tau_2) = (0.05,\ 0.2)$. In [12], the delays are assumed uncertain with weights $v_1 = 1,\ v_2 = 0.5$. We want to compute the stability radius when also an uncertainty δA_0 affects matrix A_0, and δA_1 affects the first column of matrix A_1; we thus define

$$B_{0,0} = \|A_0\|_F\, I_3,\ C_{0,0} = I_3,\ B_{1,1} = \|A_1\|_F\, I_3,\ C_{1,1} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}.$$

Choosing matrices $B_{0,0},\ B_{1,1}$ as depending on the Frobenius norm of the perturbed matrices A_0 and A_1, we allow a maximal relative perturbation equal to ε: our method returns the distance to instability $r_{\mathbb{R}} = 9.27826 * 10^{-3}$. Furthermore, if we neglect the perturbations on the matrices and maintain the same uncertainties on the delay values, we obtain $r_{\mathbb{R}} = 2.45684 * 10^{-2}$, which is consistent with the one provided in [12] ($r_{\mathbb{R}} = 2.45769e-02$).

VI. Conclusions and Perspectives

The advantage of the presented approach for computing the PSA and the distance to instability, lies in the very broad class of perturbations considered, and in its ability to exploit their structure and inter-dependencies. Some additional features, which we have not addressed for sake of conciseness, are the following.

First, since low-rank properties of optimal perturbations are fully exploited in the parameterization and the algorithm only requires the computation of selected rightmost eigenvalues (for which fast iterative methods can be used), large-scale problems, which stem for instance from the discretization of PDEs with delay, can be dealt as well.

Second, our implementation supports classes of descriptor systems with delays, where $\dot{x}$ in (1) is replaced with $E\dot{x}$ and E is allowed to be singular. As a motivation, we reconsider system (3) where we now assume that ξ and ω are uncertain. The nonlinear dependence on these parameters (i.e., the product $\xi\omega$) can be overcome by rewriting the equations as (with $x_1 = x,\ x_2 = \dot{x}$ and x_3 a slack variable),

$$\begin{cases} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= \omega x_3(t) + \frac{k}{m}\left(x_1(t-\tau) - x_1(t)\right) \\ 0 &= x_3(t) + 2\xi x_2(t) + \omega x_1(t) \end{cases}$$

In a similar fashion, any polynomial dependence on uncertainty can be handled. Another advantage of descriptor models is their suitability for modeling interconnections of (sub)systems and controllers, without the need for elimination of variables.

Acknowledgments

This work was supported by the project C14/17/072 of the KU Leuven Research Council.

References

[1] F. Borgioli and W. Michiels. Computing the distance to instability for delay systems with uncertainties in the system matrices and in the delay terms. In *Proceedings of the ECC*, Limassol, Cyprus, 2018.

[2] L. Dieci and T. Eirola. On smooth decompositions of matrices. *SIAM Journal on Matrix Analysis and Applications*, 20(3):800–819, 1999.

[3] K. Gu, S. I. Niculescu, and J. Chen. On stability of crossing curves for general systems with two delays. *Journal of Mathematical Analysis and Applications*, 311(1):231–253, 2005.

[4] N. Guglielmi and C. Lubich. Differential equations for roaming pseudospectra: paths to extremal points and boundary tracking. *SIAM Journal on Numerical Analysis*, 49:1194–1209, 2012.

[5] N. Guglielmi and C. Lubich. Low-rank dynamics for computing extremal points of real pseudospectra. *SIAM Journal of Matrix Analysis and Applications*, 34:40–66, 2013.

[6] D. Hinrichsen and A.J. Pritchard. Stability radius for structured perturbations and the algebraic riccati equation. *Systems & Control Letters*, 8:105–113, 1986.

[7] G. Hu and E. J. Davison. Real stability radii of linear time-invariant time-delay systems. *Systems & Control Letters*, 50:209 – 219, 2003.

[8] T Insperger and G. Stépán. Stability of the milling process. *Periodica Polytechnica Mechanical Engineering*, 44(1):47–57, 2000.

[9] O. Koch and C. Lubich. Dynamical low rank approximation. *SIAM J. Matrix Anal. Appl.*, 29(2):434–454, 2007.

[10] K. Meerbergen, E. Mengi, W. Michiels, and R. Van Beeumen. Computation of pseudospectral abscissa for large scale nonlinear eigenvalue problems . *IMA Journal of Numerical Analysis*, 2016. Accepted for publication.

[11] W. Michiels. Spectrum based stability analysis and stabilization of systems described by delay differential algebraic equations. *IET Control Theory and Applications*, 5(16):1829–1842, 2011.

[12] W. Michiels, E. Fridman, and S.I. Niculescu. *Robustness assessment via stability radii in delay parameters*. International Journal of Robust and Nonlinear Control, 2009. In press (published on-line).

[13] W. Michiels, K. Green, T. Wagenknecht, and S.I. Niculescu. Pseudospectra and stability radii for analytic matrix functions with applications to time-delay systems. *Linear Algebra and its Applications*, 418(1):315–335, 2006.

[14] W. Michiels and N. Guglielmi. An iterative method for computing the pseudospectral abscissa for a class of nonlinear eigenvalue problems. *SIAM Journal on Scientific Computing*, 34(4):2366–2393, 2012.

[15] W. Michiels and S.I. Niculescu. *Stability and stabilization of time-delay systems. An eigenvalue based approach*. SIAM, 2007.

[16] L. Qiu, B. Bernhardsson, A. Rantzer, E. Davison, P. Young, and J. Doyle. A formula for computation of the real stability radius. *Automatica*, 31:879–890, 1995.

[17] K. Schreiber. *Nonlinear eigenvalue problems: Newton-type methods and nonlinear Rayleigh functionals*. PhD thesis, TU Berlin, 2008.

[18] R. Sipahi and N. Olgac. Complete stability robustness of third-order LTI multiple time-delay systems. *Automatica*, 41(8):1413–1422, 2005.

[19] J. Sreedhar, P. Van Dooren, and A.L. Tits. A fast algorithm to compute the real structured stability radius. volume 121 of *International Series of Numerical Mathematics*, pages 219–230. 1996.

[20] E.M. Stein and R. Shakarchi. *Real Analysis*. Springer Series in Operations Research. Princeton University Press, 2005.

[21] L.N. Trefethen and M. Embree. *Spectra and pseudospectra: the behavior of nonnormal matrices and operators*. Princeton University Press, 2005.

[22] Z. Wu and W. Michiels. Reliably computing all characteristic roots of delay differential equations in a given right half plane. *Journal of Computational and Applied Mathematics*, 236:2499–2514, 2012.