LINEAR ALGEBRA AND ITS APPLICATIONS

ELSEVIER Linear Algebra and its Applications 351–352 (2002) 381–410
www.elsevier.com/locate/laa

Real and complex stability radii of polynomial matrices

Y. Genin, R. Ştefan, P. Van Dooren*

CESAME, Université Catholique de Louvain, Av. G. Lemaître, 4, B-1348 Louvain-la-Neuve, Belgium

Received 30 January 2001; accepted 27 November 2001

Submitted by D. Hinrichsen

Abstract

Analytic expressions are derived for the complex and real stability radii of non-monic polynomial matrices with respect to an arbitrary stability region of the complex plane. Numerical issues for computing these radii for different perturbation structures are also considered with application to robust stability of Hurwitz and Schur polynomial matrices. © 2002 Elsevier Science Inc. All rights reserved.

Keywords: Complex and real stability radius; Polynomial matrices; Structured matrix perturbations

1. Introduction

The robustness issue is a crucial problem for the application of control theory; for example, one of the basic goals of feedback control is to enhance system robustness (see [4]). Robust stability is also an important topic in linear algebra [2,20] as well as in numerical analysis [19].

A fundamental problem in robustness analysis is to determine the ability of a system matrix to maintain its stability under a certain class of perturbations. A natural robustness measure is the *distance* of a *stable system* $\dot{x} = Ax$ to the *set of unstable systems* of the same form and dimension. The idea of Hinrichsen and Pritchard [9], defining the stability radius as the distance to instability, has proved to be very fruitful in stimulating a large amount of research and in establishing interesting connections

* Corresponding author.

E-mail addresses: genin@csam.ucl.ac.be (Y. Genin), stefan@csam.ucl.ac.be (R. Ştefan), vdooren@csam.ucl.ac.be (P. Van Dooren).

PII: S0024-3795(02)00278-1

(see [7,17,21]). In many applications, it is more convenient to deal with the characteristic polynomial of the closed-loop matrix, as for instance in the single-input or single-output cases (see [3]). A desired property of the (closed-loop) characteristic polynomial is that all its roots are located in a pre-specified, "good" area of the complex plane, $\mathbb{C}_g \subset \mathbb{C}$. Let $\mathbb{K}$ be either the real field $\mathbb{R}$ or the complex field $\mathbb{C}$. A polynomial in the complex variable λ,

$$p(\lambda) = p_0 + p_1\lambda + \cdots + p_k\lambda^k, \quad p_k \neq 0, \ \ p_i \in \mathbb{K}, \ \ i = 0{:}\,k,$$

is said to be $\mathbb{C}_g$-stable (or simply stable) if all its roots are located in the stability region $\mathbb{C}_g$. A natural stability robustness measure is the distance of a *stable* polynomial $p(\lambda)$ to the set of *unstable* polynomials. The stability radius of $p(\lambda)$ is defined as the *norm* of the smallest perturbation

$$\delta p(\lambda) = \delta p_0 + \delta p_1\lambda + \cdots + \delta p_k\lambda^k, \quad \delta p_i \in \mathbb{K}, \ \ i = 0{:}\,k,$$

needed to "destabilize" $p(\lambda)$, i.e. forcing at least one root of $p(\lambda) + \delta p(\lambda)$ to leave the "good" region. The norm of the perturbations will be measured with the help of the norm of a constant matrix (or vector), depending on the polynomial $\delta p(\lambda)$.

A current research problem is to extend the stability radii theory to systems described by equations other than ordinary differential ones. In this respect, the main theme of the present paper is to address the robust stability problem of time-invariant linear systems described by higher order differential or difference equations of the form

$$P_0 + P_1\frac{\mathrm{d}x(t)}{\mathrm{d}t} + \cdots + P_k\frac{\mathrm{d}^k x(t)}{\mathrm{d}t^k} = 0, \quad t \in \mathbb{R}_+, \tag{1}$$

or

$$P_0 + P_1 x(t+1) + \cdots + P_k x(t+k) = 0, \quad t \in \mathbb{Z}_+, \tag{2}$$

where $P_i \in \mathbb{K}^{n\times n}$. Such systems appear frequently in mechanical engineering. Classically, associated with the systems (1) or (2) is the polynomial matrix

$$P(\lambda) = P_0 + P_1\lambda + \cdots + P_k\lambda^k, \quad P_i \in \mathbb{K}^{n\times n}, \ \ i = 0{:}\,k,$$

that is assumed to be square invertible and to have zeros—i.e. the roots of the polynomial $\det P(\lambda)$—inside a given region $\mathbb{C}_g \subset \mathbb{C}$. By extending the stability notion introduced for polynomials, $P(\lambda)$ is said to be $\mathbb{C}_g$-stable (or just stable) if all its zeros are located in the stability region $\mathbb{C}_g$. Similarly, a robust stability measure can be defined as the *norm* of the smallest "destabilizing" perturbation

$$\delta P(\lambda) = \delta P_0 + \delta P_1\lambda + \cdots + \delta P_k\lambda^k, \quad \delta P_i \in \mathbb{K}^{n\times n}, \ \ i = 0{:}\,k.$$

Again, the norm of the perturbations will be measured via the norm of a constant matrix Δ, depending on the coefficients of $\delta P(\lambda)$. A detailed problem formulation will be given in Section 3. It will be shown that the structure of Δ strongly influences the computation of the different stability robustness measures.

The *complex* stability radius theory of polynomial matrices has been investigated by Pappas and Hinrichsen in [16]. They have analyzed the monic case only, but including structured perturbations of the coefficients; moreover, they have obtained computable formulas for different perturbation structures and for arbitrary norms.

The paper is organized as follows. Section 2 is devoted to some prerequisites concerning the *stability radius*. Some particular aspects regarding the scalar polynomial case are emphasized in Section 3, in connection with the problem formulation for different perturbation structures. In Section 4 we are treating the complex case, considering Hölder norms. Closed formulas for the real stability radii of polynomial matrices are then derived in Section 6, with emphasis on the 2-norm case. As a result, it is shown in Section 6 how both real and complex stability radii can be efficiently computed for Hurwitz and Schur polynomial matrices. Some additional comments on computational complexity conclude this section. Future research directions along with some short remarks are finally indicated.

2. Preliminaries and basic results

Consider a partitioning of the complex plane $\mathbb{C}$ into two disjoint sets $\mathbb{C}_\mathrm{g}$ and $\mathbb{C}_\mathrm{b}$ such that $\mathbb{C}_\mathrm{g}$ is open and non-empty, $\mathbb{C} = \mathbb{C}_\mathrm{g} \cup \mathbb{C}_\mathrm{b}$. Recall that $\mathbb{K} \in \{\mathbb{C}, \mathbb{R}\}$ and consider the matrix $A \in \mathbb{K}^{n\times n}$ such that $\Lambda(A) \subset \mathbb{C}_\mathrm{g}$, that is, A is $\mathbb{C}_\mathrm{g}$-stable (or simply stable). The two regions that are typically considered for $\mathbb{C}_\mathrm{g}$ are the open left half plane $\mathbb{C}^- = \{s \in \mathbb{C} : \operatorname{Re} s < 0\}$ and the open unit disc $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$. The *stability radius* of the matrix A, defined as

$$r_{\mathbb{K}}(A, \mathbb{C}_\mathrm{g}) := \inf_{\Delta \in \mathbb{K}^{n\times n}} \left\{ \|\Delta\| : \Lambda(A+\Delta) \cap \mathbb{C}_\mathrm{b} \neq \emptyset \right\}, \tag{3}$$

is the norm of the smallest perturbation Δ forcing at least one eigenvalue of $A + \Delta$ to leave the “good” region $\mathbb{C}_\mathrm{g}$. More details concerning this concept can be found in [9,11].

The size of the perturbation matrix $\Delta \in \mathbb{K}^{m\times l}$ is measured by the induced operator norm

$$\|\Delta\| = \sup_{x\neq 0} \frac{\|\Delta x\|_{\mathbb{K}^m}}{\|x\|_{\mathbb{K}^l}} \tag{4}$$

for arbitrary norms on $\mathbb{K}^l$ and $\mathbb{K}^m$, respectively. In (3), $l = m = n$.

Denote by E the real linear normed space $(\mathbb{K}^l, \|\cdot\|)$. Any linear functional on E can be associated with a vector belonging to the dual of E, $E^* = (\mathbb{K}^l, \|\cdot\|_\mathrm{D})$, where the *dual norm* $\|\cdot\|_\mathrm{D}$ is defined by

$$\|x\|_\mathrm{D} = \max_{v\neq 0} \frac{|x^* v|}{\|v\|}. \tag{5}$$

A vector y is said to be the dual of a vector w if $|y^* w| = \|y\|_\mathrm{D}\|w\|$.

The notation $\|\cdot\|_p$ stands either for the Hölder p-norm of any vector in $\mathbb{K}^n$, $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$, or for the induced operator norm of any linear map $\Delta : \mathbb{K}^l \to \mathbb{K}^m$, $\|\Delta\|_p = \sup_{v\neq 0} \|\Delta v\|_p / \|v\|_p$. The distinction will be clear from the context. Note also that the dual norm of $\|\cdot\|_p$ is $\|\cdot\|_q$, where $1/p + 1/q = 1$. One has that $\|\Delta\|_2 = \sigma_1(\Delta)$, where $\sigma_1(U)$ denotes the largest singular value of the matrix U; in general, the ith singular value of U will be written $\sigma_i(U)$.

Remark 1 (see [9,11]). Let $\partial\mathbb{C}_g$ denote the boundary of $\mathbb{C}_g$. By continuity of the spectrum of a matrix versus perturbations on its entries, it follows that the eigenvalue "leaving" $\mathbb{C}_g$ for $\mathbb{C}_b$ must actually lie on its boundary $\partial\mathbb{C}_g$. Therefore

$$\begin{aligned} r_{\mathbb{K}}(A, \mathbb{C}_g) &= r_{\mathbb{K}}(A, \partial\mathbb{C}_g) \\ &= \inf_{\lambda\in\partial\mathbb{C}_g} \left(\inf_{\Delta\in\mathbb{K}^{m\times l}} \left\{ \|\Delta\| : \det(\lambda I - A - \Delta) = 0 \right\} \right) \\ &= \inf_{\lambda\in\partial\mathbb{C}_g} \left(\inf_{\Delta\in\mathbb{K}^{m\times l}} \left\{ \|\Delta\| : \det(I - \Delta(\lambda I - A)^{-1}) = 0 \right\} \right), \end{aligned} \tag{6}$$

the last equality resulting from the stability of the initial matrix A: $\lambda I - A$ is invertible for $\lambda \in \partial\mathbb{C}_g$. Relation (6) shows that an important issue in stability radius computation is to solve the following linear algebra problem: given a matrix $M \in \mathbb{C}^{l\times m}$ determine

$$\inf_{\Delta\in\mathbb{K}^{m\times l}} \left\{ \|\Delta\| : \det(I - \Delta M) = 0 \right\}. \tag{7}$$

If *both* M and Δ are complex (or real), then the following result holds for arbitrary norms on $\mathbb{K}^l$, $\mathbb{K}^m$ (see also [11, Proposition 3.1]).

Lemma 2. *For all $M \in \mathbb{K}^{l\times m}$ and any operator norm*

$$\inf_{\Delta\in\mathbb{K}^{m\times l}} \left\{ \|\Delta\| : \det(I - \Delta M) = 0 \right\} = \|M\|^{-1}. \tag{8}$$

Moreover, there exists always a rank one "optimal" perturbation Δ_{opt} for which the infimum in (8) *is attained. If $v \in \mathbb{K}^m$ is a unit norm vector such that $\|Mv\|_{\mathbb{K}^l} = \|M\|$, then $\Delta_{\text{opt}} = \|M\|^{-1} v\, u_d^*$, where u_d is the dual of Mv, $\|u_d\| = 1$.*

When Δ is real and M is complex, the problem (7) is more involved. It can be solved with the help of the following theorem, valid only for Euclidean norms ($p = 2$). To our knowledge, there is no other available result for p-Hölder norms. Define the largest *real perturbation value* (or the *real structured singular value*) of M by

$$\mu_{\mathbb{R}}(M) := \left[\inf_{\Delta\in\mathbb{R}^{m\times l}} \left\{ \|\Delta\| : \det(I - \Delta M) = 0 \right\} \right]^{-1}, \quad M \in \mathbb{C}^{l\times m}. \tag{9}$$

Notice that $\mu_{\mathbb{R}}(M) = 0$ if and only if there is no Δ such that $\det(I - \Delta M) = 0$.

By introducing

$$G(\lambda) := (\lambda I - A)^{-1},$$

one can deduce from relation (6), combined with (8) and (9), that

$$r_{\mathbb{C}}(A, \mathbb{C}_{\mathrm{g}}) = \left[\sup_{\lambda \in \partial \mathbb{C}_{\mathrm{g}}} \|G(\lambda)\| \right]^{-1} \quad \text{and} \quad r_{\mathbb{R}}(A, \mathbb{C}_{\mathrm{g}}) = \left[\sup_{\lambda \in \partial \mathbb{C}_{\mathrm{g}}} \mu_{\mathbb{R}}(G(\lambda)) \right]^{-1}. \tag{10}$$

The first equality in (10) has been proved in [9,11], while the second one is due to [11,14]. For Euclidean norms, an explicit formula for $\mu_{\mathbb{R}}$ has been derived by Qiu et al. in [14], and it is presented in Theorem 3. An alternative approach was proposed by Hinrichsen and Pritchard in [11], considering arbitrary pairs of norms, but it proved to be effective for the rank one case only (and in particular when $m = 1$ or $l = 1$). Further, both approaches will be reviewed hereafter with emphasis on properties specifically relevant to our treatment.

For any complex matrix (vector, scalar) $M \in \mathbb{C}^{l \times m}$, let $M_x \in \mathbb{R}^{l \times m}$, $M_y \in \mathbb{R}^{l \times m}$ denote its real and imaginary parts, respectively, that is $M = M_x + \mathrm{j} M_y$. Associate to M the $2l \times 2m$ *real* matrix depending on the real parameter $\gamma \in (0, 1]$:

$$N_M(\gamma) := \begin{bmatrix} M_x & -\gamma M_y \\ \gamma^{-1} M_y & M_x \end{bmatrix}. \tag{11}$$

Then the following result holds.

Theorem 3 [14]. *Let $M \in \mathbb{C}^{l \times m}$. Then*

$$\mu_{\mathbb{R}}(M) = \inf_{\gamma \in (0,1]} \sigma_2 \left(N_M(\gamma) \right) \tag{12}$$

and the function to be minimized on the right-hand side of (12) *is a unimodal function on* (0, 1].

The remarks below are due to Qiu et al. (see [14]).

Remark 4.

1. The minimization in (12) is quite easy since $\sigma_2(\cdot)$ has only one local minimum which is also a global one, except when the infimum is attained for $\gamma \to 0$.
2. The map $M \mapsto \mu_{\mathbb{R}}(M)$ is continuous almost everywhere. At its discontinuity points one has necessarily $M_y = 0$. This leads to a non-continuous function of λ which has also to be maximized on the boundary of the stability region, as shown by (10). This is not a simple numerical problem; the question is discussed in some detail in [18].
3. It can be shown that $\mu_{\mathbb{R}}(M) = \sigma_1(M)$ if and only if the minimal value of σ_2 is attained for $\gamma = 1$.

Remark 5. Assume that the optimum in (12) is attained for some $\gamma_{\mathrm{opt}} \in (0, 1]$. Then the "optimal" perturbation, i.e. the minimum norm real matrix Δ such that $\det(I - \Delta M) = 0$ is given by

$$\Delta = \sigma_{2,\mathrm{opt}}^{-1} \begin{bmatrix} v_x & v_y \end{bmatrix} \begin{bmatrix} u_x & u_y \end{bmatrix}^{\dagger}, \tag{13}$$

where

$$u = \begin{bmatrix} u_x \\ u_y \end{bmatrix} \quad \text{and} \quad v = \begin{bmatrix} v_x \\ v_y \end{bmatrix}$$

are a pair of left and right singular vectors of the matrix $N_M(\gamma_{\mathrm{opt}})$ corresponding to $\sigma_{2,\mathrm{opt}}$, such that $u_x^{\mathrm{T}} u_x = v_x^{\mathrm{T}} v_x$, $u_x^{\mathrm{T}} u_y = v_x^{\mathrm{T}} v_y$ and $u_y^{\mathrm{T}} u_y = v_y^{\mathrm{T}} v_y$. We have used $A^{\dagger}$ to denote the Moore–Penrose (generalized) inverse of the matrix A. Except for special cases (for which we refer to [14] for more details), it follows from (13) that rank $\Delta = 2$.

As already mentioned, there is an alternative method to determine $\mu_{\mathbb{R}}$ (and implicitly $r_{\mathbb{R}}$), proposed by Hinrichsen and Pritchard (see [9,11]). For any $v_1, v_2 \in \mathbb{R}^m$ and $u_1, u_2 \in \mathbb{R}^l$, define the smallest operator norm of all linear maps $\Delta : \mathbb{R}^l \mapsto \mathbb{R}^m$ which take u_1, u_2 onto v_1, v_2 as

$$\delta(u_1, u_2; v_1, v_2) = \inf_{\Delta \in \mathbb{K}^{m \times l}} \left\{ \|\Delta\| : \Delta u_1 = v_1, \Delta u_2 = v_2 \right\}. \tag{14}$$

Note that $\delta = \infty$ if and only if there is no Δ such that $\Delta u_1 = v_1$, $\Delta u_2 = v_2$.

If $\|\Delta\| = \sigma_1(\Delta)$, a closed formula for δ can be obtained on the basis of Theorem 4.3 in [11]. Furthermore, the following result holds for arbitrary pairs of norms on $\mathbb{R}^l$ and $\mathbb{R}^m$, respectively.

Proposition 6. *Let $M \in \mathbb{C}^{l \times m}$, $M = M_x + \mathrm{j} M_y$. Then*

$$\mu_{\mathbb{R}}(M) = \left[\inf_{(v_x, v_y) \neq (0,0)} \delta(M_x v_x - M_y v_y, M_x v_y + M_y v_x; v_x, v_y) \right]^{-1}. \tag{15}$$

Note that the right-hand sides of (12) and (15) are the same for Euclidean norms, but, to our knowledge, there is no *direct* proof of showing this equivalence in the general case.

Let us end this section with some additional remarks. Recall that δ can be determined explicitly in the case of Euclidean norms on $\mathbb{R}^l$ and $\mathbb{R}^m$. However, this approach does not yield an alternative computational scheme for $r_{\mathbb{R}}$ or $\mu_{\mathbb{R}}$: the computational complexity of calculating $r_{\mathbb{R}}$ or $\mu_{\mathbb{R}}$ appears to be too high, due to the optimization over v_x, v_y.

3. Problem formulation

Consider the polynomial matrix

$$P(\lambda) = P_0 + P_1 \lambda + \cdots + P_k \lambda^k, \quad P_i \in \mathbb{K}^{n \times n}, \ \ i = 0{:}\,k,$$

and define the spectrum of P as

$$\Lambda(P) = \{\lambda \in \mathbb{C} : \det P(\lambda) = 0\}.$$

The elements of $\Lambda(P)$ are called the *eigenvalues* or *zeros* of $P(\lambda)$. We shall say that $P(\lambda)$ is $\mathbb{C}_g$-stable (or just stable) if $\Lambda(P) \subset \mathbb{C}_g$ and call $\mathbb{C}_g$ the stability region. The typical regions chosen for $\mathbb{C}_g$ are the open left complex half plane and the open unit disc.

Let us assume that $P(\lambda)$ is $\mathbb{C}_g$-stable, *regular* (i.e. $\det P(\lambda) \not\equiv 0$) and that P_k is *non-singular*. The stability radius of such polynomial matrices is the norm of the smallest perturbation

$$\delta P(\lambda) = \delta P_0 + \delta P_1 \lambda + \cdots + \delta P_k \lambda^k, \quad \delta P_i \in \mathbb{K}^{n\times n}, \ \ i = 0{:}\,k,$$

needed to "destabilize" $P(\lambda)$, and hence forcing at least one zero of $P(\lambda) + \delta P(\lambda)$ to leave the stability region $\mathbb{C}_g$.

Assume that we measure the perturbations via some appropriate norm of a constant matrix Δ depending on the coefficients of $\delta P(\lambda)$. Precise definitions of Δ will be given later. Then the *stability radius* of $P(\lambda)$ with respect to $\mathbb{C}_g$ has the expression

$$r_{\mathbb{K}}(P, \mathbb{C}_g) = \inf_{\Delta} \{\|\Delta\| : \exists \lambda \in \mathbb{C}_b \text{ s.t. } \det(P(\lambda) + \delta P(\lambda)) = 0\}. \tag{16}$$

By convention $r_{\mathbb{K}} = +\infty$ if there is no $\delta P(\lambda)$ such that $\det(P(\lambda) + \delta P(\lambda)) = 0$ for some $\lambda \in \mathbb{C}_b$. Let

$$\det P(\lambda) = a_{nk}\lambda^{nk} + \cdots + a_1\lambda + a_0.$$

We rule out the case $a_{nk} = \det P_k = 0$. In order to see this, let us first deal with the scalar case, when $P_i = p_i \in \mathbb{K}$.

We shall prove that *Hurwitz* stability radii problems are trivial if the leading coefficient p_k is zero. Assume e.g. that $p_k = p_{k-1} = \cdots = p_{l+1} = 0$ and $p_l \neq 0$. It then appears that the degree k polynomial $p(\lambda)$ has $k - l$ zeros at infinity (to see this, observe that the polynomial $x^k\, p(1/x)$ has a zero of multiplicity $k - l$ at $x = 0$). In such a situation, there exist arbitrarily small perturbations $\delta p_0, \delta p_1, \ldots, \delta p_k$ such that $p(\lambda) + \delta p(\lambda)$ has a zero in the unstable part $\mathbb{C} \setminus \mathbb{C}^-$ of the complex plane. For example, choose $\delta p_i = 0$ for all $i \neq l + 1$ and $\delta p_{l+1} = -\epsilon p_l$ with $\epsilon > 0$ but arbitrarily small; then the zeros $1/x$ of the perturbed polynomial appear to be given by the roots of the polynomial equation

$$x^{k-l-1}\big(- \epsilon p_l + p_l x + p_{l-1}x^2 + \cdots + p_1 x^l + p_0 x^{l+1}\big) = 0. \tag{17}$$

This polynomial has a zero of order $k - l - 1$ at $x = 0$, while its other zeros are the solutions of $-\epsilon p_l + p_l x + \cdots + p_0 x^{l+1} = 0$; in particular, they satisfy the relation

$$x = \epsilon - \frac{1}{p_l}\big(p_{l-1}x^2 + p_{l-2}x^3 + \cdots + p_1 x^l + p_0 x^{l+1}\big). \tag{18}$$

For ϵ sufficiently small and in view of the polynomial zero continuity theorem, the above equation has a solution of the form $x = \epsilon + \mathcal{O}(\epsilon^2)$, arbitrary close to $\epsilon > 0$ for $\epsilon \to 0$. Therefore, the perturbed polynomial is unstable.

To sum up, it appears that the *distance to instability of a polynomial with zero leading coefficient* is inherently zero. The same is true for matrix polynomials with singular leading coefficient matrix; indeed, the above argument can be extended to cover the matrix case without difficulty. It is left to the reader to verify that the various formulas for the stability radii presented in this paper are coherent with this property; for example, one finds $\delta P_i = 0$ in (39) as expected in case $P(\lambda)$ is singular at infinity. Let us finally note that the stability radii theory of polynomials can be retrieved as a particular case of the stability radii theory of polynomial matrices.

Thus one can rewrite (16) as

$$r_{\mathbb{K}}(P, \mathbb{C}_{\mathrm{g}}) = \inf_{\Delta} \left\{ \|\Delta\| : \exists\lambda \in \partial\mathbb{C}_{\mathrm{g}} \text{ s.t. } \det(P(\lambda) + \delta P(\lambda)) = 0 \right\}. \tag{19}$$

For any polynomial matrix P and for every $\lambda_0 \in \mathbb{C}$ introduce

$$\nu_{\mathbb{K}}(P, \lambda_0) := \inf_{\Delta} \left\{ \|\Delta\| : \det(P(\lambda_0) + \delta P(\lambda_0)) = 0 \right\}, \tag{20}$$

i.e. $\nu_{\mathbb{K}}(P, \lambda_0)$ is the norm of the smallest perturbation needed to make one eigenvalue of P equal to λ_0. From (19) and (20), one obtains

$$r_{\mathbb{K}}(P; \mathbb{C}_{\mathrm{g}}) = \inf_{\lambda \in \partial\mathbb{C}_{\mathrm{g}}} \nu_{\mathbb{K}}(P, \lambda). \tag{21}$$

Therefore, the computation of $\nu_{\mathbb{K}}(P, \lambda)$ appears to be the key issue in evaluating the stability radius of P. Moreover, $\nu_{\mathbb{K}}$ is involved as well in determining the real or complex *pseudospectra* of polynomial matrices (see [8,12]).

Let us consider the following perturbation structures:

$$\begin{aligned} \Delta_1 &= \begin{bmatrix} \delta P_0 & \cdots & \delta P_k \end{bmatrix}, \\ \Delta_2 &= \begin{bmatrix} \delta P_0 \\ \vdots \\ \delta P_k \end{bmatrix} \\ \Delta_3 &= \begin{bmatrix} \delta P_0 & & \\ & \ddots & \\ & & \delta P_k \end{bmatrix}. \end{aligned} \tag{22}$$

The corresponding polynomial matrix perturbation $\delta P(\lambda)$ can be expressed, respectively, as

$$\begin{aligned} \delta P(\lambda) = \Delta_1 \begin{bmatrix} I \\ \lambda I \\ \vdots \\ \lambda^k I \end{bmatrix} &= \begin{bmatrix} I & \lambda I & \cdots & \lambda^k I \end{bmatrix} \Delta_2 \\ &= \begin{bmatrix} I & \xi_1^{-1} I & \cdots & \xi_k^{-1} I \end{bmatrix} \Delta_3 \begin{bmatrix} I \\ \xi_1 \lambda I \\ \vdots \\ \xi_k \lambda^k I \end{bmatrix}, \end{aligned} \tag{23}$$

where the $\xi_i \in \mathbb{C}$ are arbitrary, $\xi_i \neq 0$. For any λ for which $P(\lambda)$ is invertible, introduce

$$\begin{aligned} M_1(\lambda) &:= \begin{bmatrix} I \\ \lambda I \\ \vdots \\ \lambda^k I \end{bmatrix} P^{-1}(\lambda), \\ M_2(\lambda) &:= P^{-1}(\lambda)\begin{bmatrix} I & \lambda I & \cdots & \lambda^k I \end{bmatrix}, \\ M_3(\lambda,\xi) &:= \begin{bmatrix} I \\ \xi_1\lambda I \\ \vdots \\ \xi_k\lambda^k I \end{bmatrix} P^{-1}(\lambda)\begin{bmatrix} I & \xi_1^{-1} I & \cdots & \xi_k^{-1} I \end{bmatrix}. \end{aligned} \tag{24}$$

By using the well-known equality $\det(I + AB) = \det(I + BA)$, one can deduce from (23) and (24) that

$$\begin{aligned} \det(P(\lambda) + \delta P(\lambda)) = 0 &\iff \det(I + \delta P(\lambda)\ P^{-1}(\lambda)) = 0 \\ &\iff \det(I + P^{-1}(\lambda)\ \delta P(\lambda)) = 0 \\ &\iff \det(I + \Delta_i M_i(\lambda)) = 0, \quad i = 1, 3. \end{aligned} \tag{25}$$

Let us check, for instance, (25) when $i = 2$. One has

$$\begin{aligned} \det(P(\lambda) + \delta P(\lambda)) = 0 &\iff \det(I + P^{-1}(\lambda)\ \delta P(\lambda)) = 0 \\ &\iff \det\left(I + P^{-1}(\lambda)[I \ \ \lambda I \ \ \cdots \ \ \lambda^k I]\Delta_2\right) = 0 \\ &\iff \det(I + M_2(\lambda)\Delta_2) = 0. \end{aligned}$$

Remark 7. The perturbation structures Δ_1 and Δ_2 are dual to each other, because solving the problem for Δ_1 yields automatically a solution for Δ_2^*, and hence for Δ_2. Henceforth, we shall restrict our discussion to Δ_1 and Δ_3.

The following preliminary result holds.

Lemma 8. *The complex and real stability radii* (19) *of the matrix polynomial* $P(\lambda)$ *with respect to the perturbation matrix* Δ_1 *are*, *respectively*, *given by*

$$r_{\mathbb{C}}(P, \mathbb{C}_{\mathrm{g}}; \Delta_1) = \inf_{\lambda\in\partial\mathbb{C}_{\mathrm{g}}} \|M_1(\lambda)\|^{-1} = \left[\sup_{\lambda\in\partial\mathbb{C}_{\mathrm{g}}} \|M_1(\lambda)\|\right]^{-1} \tag{26}$$

and

$$r_{\mathbb{R}}(P, \mathbb{C}_{\mathrm{g}}; \Delta_1) = \inf_{\lambda\in\partial\mathbb{C}_{\mathrm{g}}} \mu_{\mathbb{R}}^{-1}(M_1(\lambda)) = \left[\sup_{\lambda\in\partial\mathbb{C}_{\mathrm{g}}} \mu_{\mathbb{R}}(M_1(\lambda))\right]^{-1}. \tag{27}$$

Proof. Since $P(\lambda)$ is $\mathbb{C}_g$-stable, $P(\lambda)$ is invertible for any $\lambda \in \partial\mathbb{C}_g$, so $M_1(\lambda)$ is well defined. For the perturbation structures Δ_1, relation (20) reads

$$\nu_{\mathbb{K}}(P, \lambda; \Delta_1) = \inf_{\Delta_1} \left\{ \|\Delta_1\| : \det(P(\lambda) + \delta P(\lambda)) = 0 \right\}.$$

In view of the equivalences in (25), the above equality can be transformed into

$$\nu_{\mathbb{K}}(P, \lambda; \Delta_1) = \inf_{\Delta_1} \left\{ \|\Delta_1\| : \det(I + \Delta_1 M_1(\lambda)) = 0 \right\}, \tag{28}$$

for any λ for which $P(\lambda)$ is invertible. If $\mathbb{K} = \mathbb{C}$, Lemma 2 shows that

$$\nu_{\mathbb{C}}(P, \lambda; \Delta_1) = \|M_1(\lambda)\|^{-1}, \tag{29}$$

and (26) follows automatically from (28) and (21). Analogously, if $\mathbb{K} = \mathbb{R}$, it follows from definition (9) that

$$\nu_{\mathbb{R}}(P, \lambda; \Delta_1) = \mu_{\mathbb{R}}^{-1} (M_1(\lambda)) . \tag{30}$$

In view of (21), equality (27) holds as well. □

Remark 9.
1. Using a similar argument, one can also deal with *structured* perturbations. Assume e.g. that the coefficients of $\delta P(\lambda)$ are expressed as $\delta P_i = D\Delta E_i$, $i = 0{:}\,k$, where D and E_i are given, and Δ is the perturbation. Let $E(\lambda) := E_0 + E_1\lambda + \cdots + E_k\lambda^k$. It is not difficult to see that
$$\det(P(\lambda) + \delta P(\lambda)) = 0 \iff \det(I + \Delta E(\lambda) P^{-1}(\lambda) D) = 0.$$
Replacing now $M_1(\lambda)$ by $E(\lambda)P^{-1}(\lambda)D$ into (26) in Lemma 8, one retrieves precisely Theorem 2.2 in [16] or Lemma 2.5 in [8].
2. One can express $\mu_{\mathbb{R}}(M_1(\lambda))$ either via Theorem 3 (when considering Euclidean norms) or via Proposition 6 (when considering arbitrary norms).

The problem $\Delta = \Delta_3$ is a constrained problem which is much more difficult to solve due to the block diagonal structure of Δ_3. Further, some upper and lower bounds for the stability radius of $P(\lambda)$ will be given in the case when $\Delta = \Delta_3$ and when considering p-norms. These bounds are expressed in terms of the stability radius determined in Lemma 8, by using the available structure and by choosing appropriate scalars ξ_i.

Lemma 10. *Let* $\nu_{\mathbb{K}}(P, \lambda; \Delta_i)$, $i = 1, 3$, *be introduced as in* (30). *Then for all* λ *for which* $P(\lambda)$ *is invertible*, *the following inequalities hold*:

$$\begin{aligned}(k+1)^{-1/q}\nu_{\mathbb{K}}(P, \lambda; \Delta_1) &\leqslant \nu_{\mathbb{K}}(P, \lambda; \Delta_3) \\ &\leqslant \nu_{\mathbb{K}}(P, \lambda; \Delta_1), \quad 1/p + 1/q = 1.\end{aligned} \tag{31}$$

Proof. The proof is very simple and left to the reader; it is based on the definition of $\nu_{\mathbb{K}}$ combined with the following facts:
1. If $\Gamma = (\mathrm{diag}(\Gamma_i))_{i=1:k}$, then

$$\|\Gamma\|_p = \max_{i=1:k} \|\Gamma_i\|_p \,. \tag{32}$$

1. If $\Gamma = [\Gamma_1\ \Gamma_2\ \ldots\ \Gamma_k] \in \mathbb{K}^{m\times l}$, $\Gamma_i \in \mathbb{K}^{m\times l_i}$, $l = \sum_{i=1}^{k} l_i$, then

$$\max_{i=1:k} \|\Gamma_i\|_p \leqslant \|\Gamma\|_p \leqslant k^{1/q} \max_{i=1:k} \|\Gamma_i\|_p, \quad 1/p + 1/q = 1. \tag{33}$$

The inequalities are tight in the sense that they can be reached for particular Γ_i, $i \in 1{:}\,k$. □

The following result is a direct consequence of equality (21) and of the previous lemma.

Corollary 11. *The following inequalities hold*:

$$\begin{aligned}(k+1)^{-1/q} r_{\mathbb{K}}(P; \mathbb{C}_{\mathrm{g}}; \Delta_1) &\leqslant r_{\mathbb{K}}(P; \mathbb{C}_{\mathrm{g}}; \Delta_3) \\ &\leqslant r_{\mathbb{K}}(P; \mathbb{C}_{\mathrm{g}}; \Delta_1), \quad 1/p + 1/q = 1.\end{aligned} \tag{34}$$

4. Complex stability radii

The aim of this section is to obtain a computable version of the formula (26) when considering p-Hölder norms. In order to prove something about the p-norms for the perturbation structures (24), we first need the following lemma. For the proof, see [6].

Lemma 12.
1. *For every Hölder* (*or p*) *norm and vectors x and y*, *one has the multiplicative property*

$$\|x \otimes y\|_p = \|x\|_p \|y\|_p. \tag{35}$$

2. *The following identities hold true for the induced matrix p-norm*:

$$\begin{aligned}\|(x \otimes I) M (y^* \otimes I)\|_p &= \|x \otimes I\|_p \|M\|_p \|y^* \otimes I\|_p \\ &= \|x\|_p \|M\|_p \|y^*\|_p.\end{aligned} \tag{36}$$

We can now state the main result of this section.

Theorem 13. *For all* λ *for which* $P(\lambda)$ *is invertible*, *one has the relation*

$$\begin{aligned}\nu_{\mathbb{C}}(P, \lambda; \Delta_i) &= \inf_{\Delta_i} \left\{ \|\Delta_i\|_p : \det(P(\lambda) + \delta P(\lambda)) = 0 \right\} \\ &= \|d_i(\lambda) P^{-1}(\lambda)\|_p^{-1},\end{aligned} \tag{37}$$

where $d_i(\lambda)$ *for* Δ_i, $i = 1, 3$, *is respectively equal to*

$$d_1(\lambda) = \left(\sum_{i=0}^{k} |\lambda|^{ip}\right)^{1/p} \quad \textit{and} \quad d_3(\lambda) = \left(\sum_{i=0}^{k} |\lambda|^{i}\right). \tag{38}$$

Proof. Let us prove first (37) for $i = 1$. By rewriting the equality (29) updated to p-norms and by applying then Lemma 12 to the particular structure of $M_1(\lambda)$, one obtains

$$\begin{aligned} \nu_{\mathbb{C}}(P, \lambda; \varDelta_1) = \|M_1(\lambda)\|_p^{-1} &= \left(\|P^{-1}(\lambda)\|_p \left\| \begin{bmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^k \end{bmatrix} \right\|_p \right)^{-1} \\ &= \left(\|d_1(\lambda) P^{-1}(\lambda)\|_p\right)^{-1}. \end{aligned}$$

Let $\xi \in \mathbb{C}^k$. According to statement 2 in Lemma 12,

$$\begin{aligned} \|M_3(\lambda, \xi)\|_p^{-1} &= \left(\|P^{-1}(\lambda)\|_p \left\| \begin{bmatrix} 1 \\ \xi_1 \lambda \\ \vdots \\ \xi_k \lambda^k \end{bmatrix} \right\|_p \left\| \begin{bmatrix} 1 \\ \xi_1^{-*} \\ \vdots \\ \xi_k^{-*} \end{bmatrix} \right\|_q \right)^{-1} \\ &=: \left(\|P^{-1}(\lambda)\|_p \|x\|_p \|y\|_q\right)^{-1} \\ &\leqslant \left(\|P(\lambda)^{-1}\|_p \left(\sum_{i=0}^{k} |x_i|\,|y_i| \right)\right)^{-1} \\ &= \left(\|d_3(\lambda) P(\lambda)^{-1}\|_p\right)^{-1}, \quad 1/p + 1/q = 1. \end{aligned}$$

The above inequality is nothing else than the Hölder inequality, applied to the vectors x and y. Equality is reached when these vectors are *dual* to each other, which is the case for $|\xi_i| \doteq |\lambda|^{-i/q}$. Furthermore, it can be shown that the above lower bound is actually reached for $\varDelta_3$, although it is constrained to be block diagonal. To that aim, let us construct a particular perturbation for $\varDelta_3$ which establishes equality. Let u and v be two vectors of unit p-norm such that $P^{-1}(\lambda)\, u = \|P^{-1}(\lambda)\|_p v$ and let v_d be the dual of v. Hence $|v_d^* \, v| = 1$ with $\|v_d^*\|_p = 1$. The matrix entries δP_3, defined by

$$\varDelta_3 : \quad \delta P_i = -\left(\|d_3(\lambda) P^{-1}(\lambda)\|_p\right)^{-1} u v_d^* \left(\frac{|\lambda|^2}{\lambda}\right)^i, \tag{39}$$

yield equality in its lower bound and also satisfy $\delta P(\lambda) = -(\|P(\lambda)^{-1}\|_p)^{-1} u v_d^*$ so that $\delta P(\lambda) P^{-1}(\lambda) u = -u$ and $(P(\lambda) + \delta P(\lambda))v = 0$.

Analogously, one can verify that the "optimal" destabilizing perturbation $\varDelta_1$ is given by

$$\Delta_1: \quad \delta P_i = -\left(\|d_1(\lambda)^p P(\lambda)^{-1}\|_p\right)^{-1} u v_d^* \left(\frac{|\lambda|^p}{\lambda}\right)^i. \qquad \square \tag{40}$$

This is now used in the following characterization of the stability radius of polynomial matrices.

Theorem 14. *The smallest perturbation of a polynomial matrix $P(\lambda)$ causing a zero of $P(\lambda) + \Delta P(\lambda)$ to reach the boundary $\partial\mathbb{C}_g$ of the stability region $\mathbb{C}_g$ is given by*

$$\begin{aligned} r_{\mathbb{C}}(P, \mathbb{C}_g; \Delta_i) &= \inf_{\Delta_i, \lambda \in \mathbb{C}_g} \left\{\|\Delta_i\|_p : \det(P(\lambda) + \Delta P(\lambda)) = 0\right\} \\ &= \inf_{\lambda \in \partial\mathbb{C}_g} \|d_i(\lambda) P(\lambda)^{-1}\|_p^{-1} \\ &= \left\{\sup_{\lambda \in \partial\mathbb{C}_g} \|d_i(\lambda) P(\lambda)^{-1}\|_p\right\}^{-1}, \end{aligned} \tag{41}$$

where $d_i(\lambda)$ for Δ_i, $i = 1, 3$, are defined as in Theorem 13.

Remark 15.

1. The result of Theorem 13 can be generalized to matrices of the form $x \otimes M$, in the sense that

$$\begin{aligned} &\inf_{\Delta} \left\{\|\Delta\|_p : \det(I + \Delta(x \otimes M)) = 0\right\} \\ &\quad = \|x\|_p^{-1} \|M\|_p^{-1}, \text{ or,} \\ &\quad = \|x\|_1^{-1} \|M\|_p^{-1}, \text{for block diagonal perturbations.} \end{aligned} \tag{42}$$

   The proof follows closely the line of the proof of Theorem 13.
2. Theorem 14 is an extension of Corollary 2.4 in [16] to the *non-monic* case. For the sake of simplicity, we only considered *unstructured* stability radii. If the overall perturbation matrix (as Δ_i) can be represented in block row form, then Theorem 14 can be easily extended to the structured case as well, see [8, Lemma 2.5].

5. Real stability radii

The main purpose of this section is to derive a computable formula for $r_{\mathbb{R}}(P, \mathbb{C}_g; \Delta_1)$. In the second part of the section, we determine the minimum norm perturbations which are actually attaining the corresponding stability radii.

5.1. Closed formulas

First, we give appropriate state space realizations for the rational matrix function $M_1(\lambda)$ defined in (24). Introduce $E := I_{kn} \oplus 0_n$,

$$A_1 := \begin{bmatrix} 0_n & I_n & & \\ & \ddots & \ddots & \\ & & 0_n & I_n \\ -P_0 & \dots & -P_{k-1} & -P_k \end{bmatrix} \in \mathbb{R}^{(k+1)n\times(k+1)n}$$
$$B_1 := \begin{bmatrix} 0_n \\ 0_n \\ \vdots \\ I_n \end{bmatrix} \in \mathbb{R}^{(k+1)n\times n}. \tag{43}$$

Straightforward computations show that

$$M_1(\lambda) = \begin{bmatrix} I \\ \lambda I \\ \vdots \\ \lambda^k I \end{bmatrix} P^{-1}(\lambda) = (\lambda E - A_1)^{-1} B_1. \tag{44}$$

In accordance with definition (3) and relation (6), the formula (44) shows that the real stability radii problems of polynomial matrices are equivalent to real *structured* stability radii problems of "companion" pencils, like (E, A_1). Furthermore, the realization (44) enables us to express the real and imaginary parts of $M_1(\lambda)$ in terms of initial data. Let

$$\lambda := \lambda_x + \mathrm{j}\lambda_y \quad \text{and} \quad M_1(\lambda) := M_{1,x} + \mathrm{j}M_{1,y}.$$

Here $M_{1,x}$, $M_{1,y}$ are real matrices of the same dimension as M_1. Then one has

$$\begin{aligned} M_{1,x} &= \mathrm{Re}\big[(\lambda E - A_1)^{-1}\big]B_1 \\ &= \big[(\lambda_x E - A_1) + \lambda_y^2 E(\lambda_x E - A_1)^{-1}E\big]^{-1} B_1 \\ M_{1,y} &= \mathrm{Im}\big[(\lambda E - A_1)^{-1}\big]B_1 \\ &= -\lambda_y(\lambda_x E - A_1)^{-1}\big[(\lambda_x E - A_1) + \lambda_y^2 E(\lambda_x E - A_1)^{-1}E\big]^{-1} B_1 \end{aligned} \tag{45}$$

An explicit formula for $\mu_{\mathbb{R}}(M_1(\lambda))$ is given below.

Lemma 16. *Let $M_{1,x}$ and $M_{1,y}$ be given by* (46). *Then*

$$\mu_{\mathbb{R}}(M_1(\lambda)) = \inf_{\gamma\in(0,1]} \sigma_2\left(\begin{bmatrix} M_{1,x} & -\gamma M_{1,y} \\ \gamma^{-1}M_{1,y} & M_{1,x} \end{bmatrix}\right) \tag{46}$$

in the Euclidean norm case, and

$$\mu_{\mathbb{R}}(M_1(\lambda)) = \left[\min_{(u,v)\neq(0,0)} \delta(M_{1,x}u - M_{1,y}v, M_{1,x}v + M_{1,y}u\ ;\ u, v) \right]^{-1} \quad (47)$$

for arbitrary pairs of norms.

Proof. The proof of relations (46) and (47) is a direct consequence of Theorem 3 and Proposition 6, applied to $M_1(\lambda)$. □

By combining now Lemmas 8 and 16 we obtain the main result of the paper.

Theorem 17. *The real stability radius* (19) *of the polynomial matrix* $P(\lambda)$ *with respect to the perturbation matrix* Δ_1 *is given by*

$$r_{\mathbb{R}}(P, \mathbb{C}_{\mathrm{g}}; \Delta_1) = \left[\sup_{\lambda\in\partial\mathbb{C}_{\mathrm{g}}} \inf_{\gamma\in(0,1]} \sigma_2 \left(\begin{bmatrix} M_{1,x} & -\gamma M_{1,y} \\ \gamma^{-1} M_{1,y} & M_{1,x} \end{bmatrix} \right) \right]^{-1} \quad (48)$$

for Euclidean norms and

$$r_{\mathbb{R}}(P, \mathbb{C}_{\mathrm{g}}; \Delta_1) = \inf_{\substack{\lambda\in\partial\mathbb{C}_{\mathrm{g}} \\ (u,v)\neq(0,0)}} \delta\left(M_{1,x}u - M_{1,y}v, M_{1,x}v + M_{1,y}v;\ u, v\right) \quad (49)$$

for arbitrary pairs of norms.

Some additional comments are given below.

Remark 18.
1. The state-space realization (43) is not unique. One can consider realizations that are more convenient to a specific purpose. In this respect, alternative state-space realizations where $\widehat{E}$ is non-singular are used when computing the Hurwitz stability radius (see (78)).
2. Formulas (46) show that $M_{1,x}$, $M_{1,y}$ depend explicitly on the real and imaginary parts of λ, i.e. λ_x and λ_y, respectively. When considering Hurwitz or Schur stability, $M_{1,x}$ and $M_{1,y}$ will depend on a single real parameter, such as ω: $\omega = \lambda_y$ for Hurwitz stability or $\mathrm{e}^{\mathrm{j}\omega} = \lambda_x + \mathrm{j}\lambda_y$ for Schur stability.
3. Although they have at this moment only some theoretical relevance, equalities (47) and (49) might prove to be useful when p-Hölder norms are considered, provided that an efficient computation of $\mu_{\mathbb{R}}$ in (47) is available.

As already mentioned, we can derive for the third case lower and upper bounds in terms of $r_{\mathbb{R}}(P; \mathbb{C}_{\mathrm{g}}; \Delta_1)$. The result is a direct consequence of Corollary 11 for $p = 2$ and $\mathbb{K} = \mathbb{R}$.

Lemma 19. *The following inequalities hold*:

$$(k+1)^{-1/2} r_{\mathbb{R}}(P; \mathbb{C}_g; \Delta_1) \leqslant r_{\mathbb{R}}(P; \mathbb{C}_g; \Delta_3) \leqslant r_{\mathbb{R}}(P; \mathbb{C}_g; \Delta_1). \tag{50}$$

Special attention will be paid (when $p = 2$) to the particular structure of $M_1(\lambda)$. This structure will be fully exploited in the light of Theorem 3, in order to reduce the complexity of the minimization over γ when calculating $r_{\mathbb{R}}$ and to obtain simpler expressions for the smallest "destabilizing" perturbations.

According to formula (11) define

$$N_1(\lambda, \gamma) := N_{M_1(\lambda)}(\gamma) = \begin{bmatrix} M_{1,x} & -\gamma M_{1,y} \\ \gamma^{-1} M_{1,y} & M_{1,x} \end{bmatrix}. \tag{51}$$

If

$$\lambda^i P^{-1}(\lambda) := X_i(\lambda) + \mathrm{j} Y_i(\lambda), \quad X_i, Y_i \in \mathbb{R}^{n \times n}, \ \ i = 0{:}\,k,$$

then

$$M_{1,x} = \begin{bmatrix} X_0 \\ X_1 \\ \vdots \\ X_k \end{bmatrix}, \quad M_{1,y} = \begin{bmatrix} Y_0 \\ Y_1 \\ \vdots \\ Y_k \end{bmatrix}.$$

Let $\lambda = \lambda_x + \mathrm{j}\lambda_y = \rho(\cos\theta + \mathrm{j}\sin\theta) = \rho \mathrm{e}^{\mathrm{j}\theta}, \ \rho > 0, \ \theta \in [0, 2\pi)$. Associate to λ the matrix

$$\Lambda := \begin{bmatrix} \lambda_x & -\lambda_y \\ \lambda_y & \lambda_x \end{bmatrix} = \rho \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

and let

$$D_\gamma := \begin{bmatrix} \gamma & 0 \\ 0 & 1 \end{bmatrix},$$

$$\Lambda_\gamma := N_\lambda(\gamma) = D_\gamma \Lambda D_\gamma^{-1} = \begin{bmatrix} \lambda_x & -\gamma\lambda_y \\ \gamma^{-1}\lambda_y & \lambda_x \end{bmatrix} = \rho \begin{bmatrix} \cos\theta & -\gamma\sin\theta \\ \gamma^{-1}\sin\theta & \cos\theta \end{bmatrix}.$$

From the definition of X_i, Y_i and since $(\Lambda \otimes I_n)^i = \Lambda^i \otimes I_n$ one gets

$$\begin{bmatrix} X_i & -Y_i \\ Y_i & X_i \end{bmatrix} = (\Lambda^i \otimes I_n) \begin{bmatrix} X_0 & -Y_0 \\ Y_0 & X_0 \end{bmatrix} = \begin{bmatrix} X_0 & -Y_0 \\ Y_0 & X_0 \end{bmatrix} (\Lambda^i \otimes I_n).$$

As $(D_\gamma \otimes I_n)^{-1} = D_\gamma^{-1} \otimes I_n$ the above equalities imply that

$$\begin{bmatrix} X_i & -\gamma Y_i \\ \gamma^{-1} Y_i & X_i \end{bmatrix} = (\Lambda_\gamma^i \otimes I_n) \begin{bmatrix} X_0 & -\gamma Y_0 \\ \gamma^{-1} Y_0 & X_0 \end{bmatrix} = \begin{bmatrix} X_0 & -\gamma Y_0 \\ \gamma^{-1} Y_0 & X_0 \end{bmatrix} (\Lambda_\gamma^i \otimes I_n). \tag{52}$$

We have used the identity $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$.

Consider the permutation $\pi \in \mathscr{S}^{2k+2}$ defined by $\pi(1 : 2k + 2) = (1 : 2 : 2k + 1, 2 : 2 : 2k + 2)$ and introduce now the orthogonal matrix

$$\Pi := \left[e_{\pi(1)} e_{\pi(2)} \cdots e_{\pi(2k+2)} \quad\right] \in \mathbb{R}^{(2k+2)\times(2k+2)},$$

where $e_i \in \mathbb{R}^{(2k+2)}$ denotes the ith column of the identity matrix I_{2k+2}. Let

$$N_0(\lambda, \gamma) := N_{P^{-1}(\lambda)} = \begin{bmatrix} X_0 & -\gamma Y_0 \\ \gamma^{-1} Y_0 & X_0 \end{bmatrix} \in \mathbb{R}^{2n\times 2n} \quad \text{and}$$
$$A_1(\lambda, \gamma) := \begin{bmatrix} I_2 \\ \Lambda_\gamma \\ \vdots \\ \Lambda_\gamma^k \end{bmatrix} \in \mathbb{R}^{2(k+1)\times 2}. \tag{53}$$

Then

$$(\Pi \otimes I_n)\ N_1(\lambda, \gamma) = (A_1(\lambda, \gamma) \otimes I_n)\ \ N_0(\lambda, \gamma)$$

as (52) shows. Since Π is orthogonal we deduce that for every $\gamma \in (0, 1]$,

$$\sigma_2(N_1(\lambda, \gamma)) = \sigma_2((A_1(\lambda, \gamma) \otimes I_n) N_0(\lambda, \gamma)). \tag{54}$$

Since $A_1^{\mathrm{T}} A_1$ is positive definite, one can find a *real* spectral factor L_1 for A_1, that is, $A_1^{\mathrm{T}} A_1 = L_1^{\mathrm{T}} L_1$. For instance, a Cholesky factor can be always obtained, but it has no rational expression in terms of γ and λ. Thus relation (54) reads

$$\sigma_2(N_1(\lambda, \gamma)) = \sigma_2((L_1(\lambda, \gamma) \otimes I_n) N_0(\lambda, \gamma)), \tag{55}$$

where $L_1(\lambda, \gamma) \in \mathbb{R}^{2\times 2}$ is a Cholesky factor of $A_1^{\mathrm{T}}(\lambda, \gamma)\ A_1(\lambda, \gamma)$.

The following result is a direct consequence of Lemma 8 and Theorem 3, combined with relation (55).

Theorem 20. *The real stability radius* (19) *of the polynomial matrix* $P(\lambda)$ *with respect to the perturbation structure* Δ_1 *is given by*

$$r_{\mathbb{R}}(P; \mathbb{C}_{\mathrm{g}}; \Delta_1) = \left[\sup_{\lambda\in\partial\mathbb{C}_{\mathrm{g}}} \inf_{\gamma\in(0,1]} \sigma_2 \left((L_1(\lambda, \gamma) \otimes I_n) N_{P^{-1}(\lambda)}(\gamma) \right) \right]^{-1}. \tag{56}$$

5.2. Minimum norm perturbations

Subsequently we shall derive simpler expressions for the minimum norm perturbation Δ_1 attaining $\mu_{\mathbb{R}}(M_1(\lambda))$ for given λ for which $\det P(\lambda) \neq 0$.

Let Δ_1 be the minimum norm “destabilizing” perturbation that attains $\mu_{\mathbb{R}}(M_1(\lambda))$. Let

$$\begin{bmatrix} u_x \\ u_y \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} v_x \\ v_y \end{bmatrix}$$

be a pair of left and right singular vectors of the matrix $N_1(\lambda, \gamma)$ corresponding to the "optimal" σ_2. Then one has (see (11))

$$\begin{bmatrix} M_{1,x} & -\gamma M_{1,y} \\ \gamma^{-1} M_{1,y} & M_{1,x} \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix} = \sigma_2 \begin{bmatrix} u_x \\ u_y \end{bmatrix}. \tag{57}$$

By exploiting the structure of $M_{1,x}$ and $M_{1,y}$ one infers from (57) that for every $i \in 1{:}\,k$,

$$\begin{bmatrix} X_i & -\gamma Y_i \\ \gamma^{-1} Y_i & X_i \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix} = \sigma_2 \begin{bmatrix} u_{x,i} \\ u_{y,i} \end{bmatrix} = \left[\begin{array}{c} u_{x,0} \\ \vdots \\ u_{x,k} \\ \hline u_{y,0} \\ \vdots \\ u_{y,k} \end{array}\right] \tag{58}$$

and by replacing now relation (52) into (58) we obtain

$$(\varLambda_\gamma^i \otimes I_n) \begin{bmatrix} X_0 & -\gamma Y_0 \\ \gamma^{-1} Y_0 & X_0 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix} = \sigma_2 \begin{bmatrix} u_{x,i} \\ u_{y,i} \end{bmatrix}.$$

Writing now (58) for $i = 0$, one deduces from above that

$$(\varLambda_\gamma^i \otimes I_n) \begin{bmatrix} u_{x,0} \\ u_{y,0} \end{bmatrix} = \begin{bmatrix} u_{x,i} \\ u_{y,i} \end{bmatrix}, \tag{59}$$

or, in a more compact form, for every $i \in 0{:}\,k$,

$$\begin{bmatrix} u_{x,0} & u_{y,0} \end{bmatrix} \begin{bmatrix} \lambda_x^i & \gamma^{-1} \lambda_y^i \\ -\gamma \lambda_y^i & \lambda_x^i \end{bmatrix} = \begin{bmatrix} u_{x,i} & u_{y,i} \end{bmatrix} \iff \varLambda_\gamma^i U_0^{\mathrm{T}} = U_i^{\mathrm{T}}. \tag{60}$$

Here $U_i := [u_{x,i}\ u_{y,i}] \in \mathbb{R}^{n \times 2}$ and $\lambda_x^i := \rho^i \cos \mathrm{i}\theta,\ \lambda_y^i := \rho^i \sin \mathrm{i}\theta$, $i = 0{:}\,k$.

Since $\varDelta_1$ is the minimum norm "destabilizing" perturbation that attains $\mu_{\mathbb{R}}(M_1(\lambda))$, formula (13) reads

$$\begin{aligned} \varDelta_1 &= -\sigma_2^{-1} \begin{bmatrix} v_x & v_y \end{bmatrix} \left(\begin{bmatrix} u_x^{\mathrm{T}} \\ u_y^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} u_x & u_y \end{bmatrix} \right)^{-1} \begin{bmatrix} u_x^{\mathrm{T}} \\ u_y^{\mathrm{T}} \end{bmatrix} \\ &=: \begin{bmatrix} \varDelta_{1,0} & \varDelta_{1,1} & \dots & \varDelta_{1,k} \end{bmatrix}. \end{aligned} \tag{61}$$

By combining now (61) with (60) one can write

$$\begin{aligned} \varDelta_1 &= -\sigma_2^{-1} \begin{bmatrix} v_x & v_y \end{bmatrix} \left(\begin{bmatrix} u_x^{\mathrm{T}} \\ u_y^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} u_x & u_y \end{bmatrix} \right)^{-1} \begin{bmatrix} U_0^{\mathrm{T}} & U_1^{\mathrm{T}} & \dots & U_k^{\mathrm{T}} \end{bmatrix} \\ &= -\sigma_2^{-1} (V^{\dagger})^{\mathrm{T}} \begin{bmatrix} I_2 & \varLambda_\gamma & \dots & \varLambda_\gamma^k \end{bmatrix} \begin{bmatrix} U_0^{\mathrm{T}} & & & \\ & U_0^{\mathrm{T}} & & \\ & & \ddots & \\ & & & U_0^{\mathrm{T}} \end{bmatrix}. \end{aligned} \tag{62}$$

Here

$$(V^{\dagger})^{\mathrm{T}} := \begin{bmatrix} v_x & v_y \end{bmatrix} \left(\begin{bmatrix} v_x^{\mathrm{T}} \\ v_y^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} v_x & v_y \end{bmatrix} \right)^{-1} = \begin{bmatrix} v_x & v_y \end{bmatrix} \left(\begin{bmatrix} u_x^{\mathrm{T}} \\ u_y^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} u_x & u_y \end{bmatrix} \right)^{-1}$$

as Remark 5 states. Relation (62) also shows that for every $i \in 0\colon k$,

$$\Delta_{1,i} = -\sigma_2^{-1} \, (V^{\dagger})^{\mathrm{T}} \Lambda_{\gamma}^{i} U_0^{\mathrm{T}}. \tag{63}$$

Essentially relying on formula (63) the following result holds.

Theorem 21. *For every $\lambda \in \mathbb{C}$ which is not a root of P and for every $i \in 0\colon k$, we have*

$$\Delta_{1,i} = -\sigma_{2,\mathrm{opt}}^{-1}(N_1) \, (V^{\dagger})^{\mathrm{T}} \, \Lambda_{\gamma_{\mathrm{opt}}}^{i} \, U_0^{\mathrm{T}}. \tag{64}$$

Furthermore,

$$\delta P(\lambda) = -\sigma_{2,\mathrm{opt}}^{-1}(N_1) \, (V^{\dagger})^{\mathrm{T}} \left[\sum_{i=0}^{k} \lambda^i \Lambda_{\gamma_{\mathrm{opt}}}^{i} \right] U_0^{\mathrm{T}}. \tag{65}$$

6. Computational aspects

The aim of this section is to show how the real and complex stability radii can be computed efficiently in some important situations. In the Euclidean norm case ($p = 2$), the algorithm proposed in this paper is based on a crucial result, connecting the singular values of a rational transfer function matrix and the imaginary or unitary eigenvalues of a corresponding Hamiltonian or symplectic pencil.

A common representation of a general rational matrix $G \in \mathbb{C}^{p\times m}(\lambda)$ is

$$G(\lambda) = C(\lambda E - A)^{-1}B + D,$$

where $A, E \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{n\times m}$, $C \in \mathbb{C}^{p\times n}$ and $D \in \mathbb{C}^{p\times m}$. So as to consider an arbitrary rational matrix together with one of its realizations, let us use the notation

$$G(\lambda) \stackrel{\sim}{=} \left[\begin{array}{c|c} \lambda E - A & -B \\ \hline C & D \end{array}\right].$$

Note the sign convention used above; $G(\lambda)$ is in fact the Schur complement of $\lambda E - A$. Let us begin with the continuous-time case.

Proposition 22. *Let $G(s) = C(sE - A)^{-1}B + D$ and let $\xi > 0$ be such that $D_\xi := D^*D - \xi^2 I$ is non-singular. If $(sE - A)$ is a regular pencil and has no generalized* *eigenvalues on the imaginary axis, then, for all* $\omega \in \mathbb{R}$, ξ *is a singular value of* $G(\mathrm{j}\omega)$ *if and only if* $\mathrm{j}\omega$ *is a generalized eigenvalue of the Hamiltonian pencil*

$$\begin{aligned} &s\mathscr{L}(\xi, G) - \mathscr{H}(\xi, G) \\ &\quad = s\begin{bmatrix} E & 0 \\ 0 & E^* \end{bmatrix} - \begin{bmatrix} A - BD_\xi^{-1}D^*C & -BD_\xi^{-1}B^* \\ \xi^2 C^*(DD^* - \xi^2 I)^{-1}C & -A^* + C^*DD_\xi^{-1}B^* \end{bmatrix} \\ &\quad = \begin{bmatrix} sE - A & 0 \\ C^*C & sE^* + A^* \end{bmatrix} - \begin{bmatrix} -B \\ C^*D \end{bmatrix} (D^*D - \xi^2 I)^{-1} \begin{bmatrix} D^*C & B^* \end{bmatrix}. \end{aligned}$$

Proof. Let $\xi > 0$ and $\omega \in \mathbb{R}$. Let us prove that

$$\begin{aligned} &\det\left(G^*(\mathrm{j}\omega)G(\mathrm{j}\omega) - \xi^2 I\right) = 0 \\ &\quad \Longleftrightarrow \det\left(\mathrm{j}\omega\mathscr{L}(\xi, G) - \mathscr{H}(\xi, G)\right) = 0, \end{aligned} \tag{66}$$

where $G^*(s) := \overline{G}^{\mathrm{T}}(-\overline{s}) = B^*(-sE^* - A^*)^{-1}C^* + D^*$. To that aim, let us first note the relation

$$\begin{aligned} G^*(\mathrm{j}\omega)G(\mathrm{j}\omega) - \xi^2 I &= D^*D - \xi^2 I + D^*C(\mathrm{j}\omega E - A)^{-1}B \\ &\quad + B^*(-\mathrm{j}\omega E^* - A^*)^{-1}C^*D \\ &\quad + B^*(-\mathrm{j}\omega E^* - A^*)^{-1}C^*C(\mathrm{j}\omega E - A)^{-1}B \\ &\stackrel{\sim}{=} \left[\begin{array}{cc|c} \mathrm{j}\omega E - A & 0 & -B \\ C^*C & \mathrm{j}\omega E^* + A^* & C^*D \\ \hline D^*C & B^* & D^*D - \xi^2 I \end{array}\right] \\ &=: \mathscr{S}. \end{aligned}$$

As the Schur complement of the upper left corner in $\mathscr{S}$ is recognized in $G^*(\mathrm{j}\omega)G(\mathrm{j}\omega) - \xi^2 I$, one has

$$\det\left(G^*(\mathrm{j}\omega)G(\mathrm{j}\omega) - \xi^2 I\right) \det\left(\begin{bmatrix} \mathrm{j}\omega E - A & 0 \\ C^*C & \mathrm{j}\omega E^* + A^* \end{bmatrix}\right) = \det \mathscr{S}. \tag{67}$$

Furthermore, considering the Schur complement of $D_\xi = D^*D - \xi^2 I$ in $\mathscr{S}$ yields the relation

$$\begin{aligned} \det \mathscr{S} &= \det D_\xi \det\left(\begin{bmatrix} \mathrm{j}\omega E - A & 0 \\ C^*C & \mathrm{j}\omega E^* + A^* \end{bmatrix} - \begin{bmatrix} -B \\ C^*D \end{bmatrix} D_\xi^{-1} \begin{bmatrix} D^*C & B^* \end{bmatrix}\right) \\ &= \det D_\xi \det(\mathrm{j}\omega\mathscr{L} - \mathscr{H}). \end{aligned} \tag{68}$$

By combining now (67) and (68), it follows that

$$\begin{aligned} &\det(G^*(\mathrm{j}\omega)G(\mathrm{j}\omega) - \xi^2 I) \det(\mathrm{j}\omega E - A) \det(\mathrm{j}\omega E^* + A^*) \\ &\quad = \det D_\xi \det(\mathrm{j}\omega\mathscr{L} - \mathscr{H}). \end{aligned} \tag{69}$$

Since $\det(\mathrm{j}\omega E - A) \neq 0$, $\det(\mathrm{j}\omega E^* + A^*) \neq 0$ for every $\omega \in \mathbb{R}$, and as D_ξ is non-singular, it appears that (66) holds and this completes the proof. □

The discrete-time counterpart of Proposition 22 is stated below without proof. In this case, one can apply the same argument as above.

Proposition 23. *Let $G(z) = C(zE - A)^{-1}B + D$ and let $\xi > 0$ be such that $D_\xi := D^*D - \xi^2 I$ is non-singular. If the pencil $(zE - A)$ is regular and has no generalized eigenvalues on the unit circle, then, for every $\omega \in \mathbb{R}$, ξ is a singular value of $G(\mathrm{e}^{\mathrm{j}\omega})$ if and only if $\mathrm{e}^{\mathrm{j}\omega}$ is a generalized eigenvalue of the symplectic pencil*

$$\begin{aligned} &z\mathscr{G}(\xi, G) - \mathscr{F}(\xi, G) \\ &\quad = z\begin{bmatrix} E & BD_\xi^{-1}B^* \\ 0 & A^* - C^*DD_\xi^{-1}B^* \end{bmatrix} - \begin{bmatrix} A - BD_\xi^{-1}D^*C & 0 \\ \xi^2 C^*(DD^* - \xi^2 I)^{-1}C & E^* \end{bmatrix} \\ &\quad = \begin{bmatrix} zE - A & 0 \\ C^*C & zA^* - E^* \end{bmatrix} - \begin{bmatrix} -B \\ C^*D \end{bmatrix} (D^*D - \xi^2 I)^{-1} \begin{bmatrix} D^*C & zB^* \end{bmatrix}. \end{aligned}$$

6.1. *The complex case*

The complex stability radius can be computed efficiently in case $G(\lambda) = d(\lambda)P(\lambda)^{-1}$ is rational in λ for $\lambda \in \partial\mathbb{C}_\mathrm{g}$. This is obviously true for the unit circle since the $d(\lambda)$ functions are constant:

$$d_1(\lambda) = (k+1)^{1/p}, \quad d_3(\lambda) = (k+1).$$

For the $\mathrm{j}\omega$ axis, one can substitute for $d_3(\lambda)$ the following polynomials of the same amplitude:

$$\begin{aligned} d(\lambda) &= \sum_{i=0}^{k} (-\mathrm{j}\lambda)^i \quad \text{for } \omega \geqslant 0 \\ d(\lambda) &= \sum_{i=0}^{k} (\mathrm{j}\lambda)^i \quad \text{for } \omega \leqslant 0, \end{aligned} \tag{70}$$

so that two different rational functions have to be considered depending on whether ω is assumed to take positive or negative values. For $d_1(\lambda)$, one can only make this substitution for the special cases $p = 1, 2, \infty$. For $p = 2$, one finds

$$|d_1(\mathrm{j}\omega)|^2 = (1 + \omega^2 + \cdots + \omega^{2k}) = |d(\mathrm{j}\omega)|^2,$$

where $d(\lambda)$ is the (stable) spectral factor of $1 + \omega^2 + \cdots + \omega^{2k}$, equal to

$$d(\lambda) = \prod_{i=1}^{k} \left(\lambda + \sin\frac{i\pi}{k+1} - \mathrm{j}\cos\frac{i\pi}{k+1} \right). \tag{71}$$

For the case $p = 1$, $d_1(\lambda)$ reduces to $d(\lambda)$ as given by (70), while for $p = \infty$, $d_1(\lambda)$ simplifies into $\max\{1, |\omega|^k\}$, whence has the same amplitude as the rational functions

$$d(\lambda) = 1 \ \text{ for } \omega \leqslant 1, \quad d(\lambda) = \mathrm{j}\omega^k \ \text{ for } \omega > 1.$$

Note that in each of these cases, the constructed polynomial $d(\lambda)$ has degree k or less, i.e. $d(\lambda) = \sum_{i=1}^{k} d_i \lambda^i$. The transfer function matrix $d(\lambda)P(\lambda)^{-1}$ admits then a generalized state space realization of the form $C(\lambda E - A_1)^{-1} B_1 =: G(\lambda)$ where $C = [d_0 I_n \; d_1 I_n \; \cdots \; d_k I_n]$ and with E, A_1, B_1 given by (43).

For the 2-norm, the corresponding complex stability radius reduces to the H_∞-norm of the transfer function $G(\lambda)$:

$$\sigma_* = \sup_{\omega \in \mathbb{R}} \sigma_1 \left(G(f(\omega)) \right), \tag{72}$$

where $f(\omega)$ is the parameterization of $\partial\mathbb{C}_{\mathrm{g}}$ in terms of $\omega \in \mathbb{R}$, and $\sigma_1(M)$ is the largest singular value of the matrix M. This calculation can be carried out iteratively by a repeated computation of the real zeros ω_i of the matrix function

$$G^*(f(\omega))\, G(f(\omega)) - \sigma_o^2 I, \tag{73}$$

based on Proposition 22 or 23. These apply to the generalized state-space model $G(\lambda) = C(\lambda E - A_1)^{-1} B_1$ yielding the following Hamiltonian and symplectic pencils:

$$\begin{bmatrix} \mathrm{j}\omega E - A_1 & -\sigma_o^{-2} B_1 B_1^* \\ C^* C & \mathrm{j}\omega E^* + A_1^* \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \mathrm{e}^{\mathrm{j}\omega} E - A_1 & -\mathrm{e}^{\mathrm{j}\omega} \sigma_o^{-2} B_1 B_1^* \\ C^* C & \mathrm{e}^{\mathrm{j}\omega} A_1^* - E^* \end{bmatrix}, \tag{74}$$

respectively.

This procedure yields efficient algorithms to find the maximum of the scalar function $\sigma(\omega) = \sigma_1(G(f(\omega)))$ [1,5,10], in term of recursive eigenvalue computation of the associated Hamiltonian or symplectic pencil.

For $p = 1, \infty$ one still has a rational matrix to deal with, but the largest singular value calculation degenerates into the largest sum of absolute values of a column or row of $G(\lambda)$. This is a scalar piecewise rational function, which can be maximized using symbolic manipulation programs: each “piece” is rational and the “branching points” are the zeros of some polynomial.

In the special case of scalar polynomials, obviously $\|P(\lambda)^{-1}\|_p = |p(\lambda)^{-1}|$. If moreover, $p = 1, 2, \infty$, then $d(\lambda)$ can also be chosen polynomial, so that one has to find the maximum of the absolute value of a scalar rational function $G(\lambda) = d(\lambda)/p(\lambda)$ on $\partial\mathbb{C}_{\mathrm{g}}$. The zeros of $d'(\lambda)p(\lambda) - p'(\lambda)d(\lambda)$ are then the extrema of this function and it suffices to look for the largest of these $2(k-1)$ values. This can be obtained in $\mathrm{O}(k^2)$ flops using polynomial root finding algorithms. Note that other approaches have been proposed in the literature for complex and real stability radii of scalar polynomials [11,13].

6.2. The real case

The real stability radius can be computed efficiently for Hurwitz and Schur polynomial matrices in the Euclidean norm case, by updating the algorithm proposed by Sreedhar et al. in [18] to deal with generalized state-space models, like $M_1(\lambda)$ appearing in (44). In order to compute the real stability radius in the continuous-time and discrete-time cases, one evaluates (48) for $\mathbb{C}_\mathrm{g} = \mathbb{C}^-$ and for $\mathbb{C}_\mathrm{g} = \mathbb{D}$, respectively. As it is shown in [18], $r_\mathbb{R}$ is computed iteratively. For the sake of completeness we shall present the basic ideas behind the development in [18], by specifying, when necessary, the changes related to our specific situation.

The algorithm is based on the connection between the singular values of a transfer function matrix and the imaginary (or unitary) eigenvalues of a related Hamiltonian (or symplectic) pencil. Such a relationship has been described by Propositions 22 and 23, respectively.

6.2.1. Hurwitz stability radius

Assume that $\Lambda(P) \subset \mathbb{C}_\mathrm{g} = \mathbb{C}^-$. In this case, the boundary of the stability region is the imaginary axis. Then take $\lambda = \mathrm{j}\omega$ in (48) and rewrite it in accordance with (51) as

$$r_\mathbb{R}^{-1}(P, \mathbb{C}^-; \Delta_1) = \sup_{\omega\in\mathbb{R}} \mu_\mathbb{R}(M_1(\mathrm{j}\omega)) = \sup_{\omega\in\mathbb{R}} \inf_{\gamma\in(0,1]} \sigma_2\left(N_1(\mathrm{j}\omega,\gamma)\right). \tag{75}$$

Our first goal is to find some *rational* matrix function $\widetilde{G}_1(\gamma, M_1(\mathrm{j}\omega))$, which is unitarily equivalent to $N_1(\mathrm{j}\omega,\gamma)$. Then one can apply Proposition 22 in order to determine the singular values of $N_1(\mathrm{j}\omega,\gamma)$. To this aim introduce for any $\gamma \in (0, 1]$,

$$\widetilde{G}_1(\gamma, M_1(\mathrm{j}\omega)) := \begin{bmatrix} I & 0 \\ 0 & \mathrm{j}I \end{bmatrix} N_1(\mathrm{j}\omega,\gamma) \begin{bmatrix} I & 0 \\ 0 & -\mathrm{j}I \end{bmatrix}. \tag{76}$$

It follows from (76) that $\widetilde{G}_1$ and N_1 are unitarily equivalent, hence they share the same singular values and we can limit our attention to $\widetilde{G}_1(\gamma, M_1(\mathrm{j}\omega))$. Further (see relations (7) and (8) in [18])

$$\widetilde{G}_1(\gamma, M_1(\mathrm{j}\omega)) = \frac{1}{2}\begin{bmatrix} I & \gamma I \\ \gamma^{-1}I & -I \end{bmatrix}\begin{bmatrix} M_1(\mathrm{j}\omega) & 0 \\ 0 & \overline{M_1(\mathrm{j}\omega)} \end{bmatrix}\begin{bmatrix} I & \gamma I \\ \gamma^{-1}I & -I \end{bmatrix}. \tag{77}$$

Here $\overline{M_1(\mathrm{j}\omega)}$ stands for the complex conjugate of $M_1(\mathrm{j}\omega)$. Since M_1 is a *real* rational matrix function in s it follows that $\overline{M_1(\mathrm{j}\omega)} = M_1(-\mathrm{j}\omega)$ is a rational matrix function in $\mathrm{j}\omega$ as well. Hence $\widetilde{G}_1(\gamma, M_1(\mathrm{j}\omega))$ is rational in $\mathrm{j}\omega$. Below we derive appropriate state-space realizations for $\widetilde{G}_1$, in order to apply Proposition 22. For, consider the alternative state-space realization $M_1(s) = \widehat{C}_1(s\widehat{E} - \widehat{A}_1)^{-1}\widehat{B}_1 + \widehat{D}_1$, where

$$\widehat{E} := I_n \oplus P_k, \quad \widehat{A}_1 := \begin{bmatrix} 0_n & I_n & & \\ & \ddots & \ddots & \\ & & 0_n & I_n \\ -P_0 & \dots & -P_{k-2} & -P_{k-1} \end{bmatrix} \in \mathbb{R}^{kn \times kn},$$

$$\widehat{B}_1 := \begin{bmatrix} 0_n \\ 0_n \\ \vdots \\ I_n \end{bmatrix} \in \mathbb{R}^{kn \times n}, \quad \widehat{C}_1 := \begin{bmatrix} I_{kn} \\ \widehat{Q}_1 \end{bmatrix} \in \mathbb{R}^{(k+1)n \times kn}, \tag{78}$$

$$\widehat{D}_1 := \begin{bmatrix} 0_n \\ 0_n \\ \vdots \\ P_k^{-1} \end{bmatrix} \in \mathbb{R}^{(k+1)n \times n},$$

and $\widehat{Q}_1 := [-P_k^{-1}P_0 \;\; -P_k^{-1}P_1 \;\; \dots \;\; -P_k^{-1}P_{k-1}] \in \mathbb{R}^{n \times kn}$. Comparing (78) with (43), one notices that $s\widehat{E} - \widehat{A}_1$ has all its eigenvalues in $\mathbb{C}^-$, while $sE - A_1$ has at least n infinite eigenvalues. In order for $\widetilde{G}_1$ to verify the assumptions of Proposition 22 we consider here for technical reasons the alternative realization (78), even though the expressions (43) are simpler. Elementary algebraic manipulations show now that

$$\widetilde{G}_1(\gamma, M_1(\mathrm{j}\omega)) = \widetilde{C}_{1,\gamma}(\mathrm{j}\omega\widetilde{E} - \widetilde{A}_1)^{-1}\widetilde{B}_{1,\gamma} + \widetilde{D}_1, \tag{79}$$

i.e. $\widetilde{G}_1(\gamma, M_1(\mathrm{j}\omega))$ is rational in $\mathrm{j}\omega$. Here

$$\begin{aligned} &\widetilde{A}_1 = \begin{bmatrix} \widehat{A}_1 & 0 \\ 0 & -\widehat{A}_1 \end{bmatrix}, \quad \widetilde{E} = \begin{bmatrix} \widehat{E} & 0 \\ 0 & \widehat{E} \end{bmatrix}, \quad \widetilde{D}_1 = \begin{bmatrix} \widehat{D}_1 & 0 \\ 0 & \widehat{D}_1 \end{bmatrix} \\ &\widetilde{B}_{1,\gamma} = \tfrac{1}{\sqrt{2}} \begin{bmatrix} \widehat{B}_1 & \gamma\widehat{B}_1 \\ -\gamma^{-1}\widehat{B}_1 & \widehat{B}_1 \end{bmatrix}, \quad \widetilde{C}_{1,\gamma} = \tfrac{1}{\sqrt{2}} \begin{bmatrix} \widehat{C}_1 & \gamma\widehat{C}_1 \\ \gamma^{-1}\widehat{C}_1 & -\widehat{C}_1 \end{bmatrix}. \end{aligned} \tag{80}$$

The following result is in fact a reformulation of Proposition 22 updated for $\widetilde{G}_1(\gamma, M_1(\mathrm{j}\omega))$ given by (79) and (80).

Theorem 24. *Let $\gamma \in (0, 1]$ and $\xi > 0$ be given such that $\widetilde{D}_1^{\mathrm{T}}\widetilde{D}_1 - \xi^2 I$ is nonsingular. Then, for every $\omega \in \mathbb{R}$, ξ is a singular value of $\widetilde{G}_1(\gamma, M_1(\mathrm{j}\omega))$ if and only if $\mathrm{j}\omega$ is a generalized eigenvalue of the Hamiltonian pencil $s\mathscr{L}(\xi, \widetilde{G}_1(\gamma, M_1)) - \mathscr{H}(\xi, \widetilde{G}_1(\gamma, M_1))$.*

Due to Theorem 24 and relation (76), the computation of the singular values of $N_1(\mathrm{j}\omega, \gamma)$ in (75) reduces now to the computation of the generalized eigenvalues of the Hamiltonian pencil $s\mathscr{L}(\xi, \widetilde{G}_1(\gamma, M_1)) - \mathscr{H}(\xi, \widetilde{G}_1(\gamma, M_1))$.

6.2.2. Schur stability radius

The stability region is in this case the open unit disc, that is, $\Lambda(P) \subset \mathbb{C}_g = \mathbb{D}$. Consequently, the boundary i.e. the unit circle is parametrized by $\lambda = e^{j\omega}$. Then one infers from (48) and (51) that

$$\begin{aligned} r_{\mathbb{R}}^{-1}(P, \mathbb{D}; \Delta_1) &= \sup_{\omega\in[0,2\pi)} \mu_{\mathbb{R}}(M_1(e^{j\omega})) \\ &= \sup_{\omega\in[0,2\pi)} \inf_{\gamma\in(0,1]} \sigma_2\big(N_1(e^{j\omega}, \gamma)\big). \end{aligned} \tag{81}$$

Clearly, relations (76) and (77) hold for $M_1(e^{j\omega})$ as well. But $\overline{M_1(e^{j\omega})} = M_1(e^{-j\omega})$ is a rational matrix function in $e^{j\omega}$, hence $\widetilde{G}_1(\gamma, M_1(e^{j\omega}))$ is rational in $e^{j\omega}$. Straightforward computations show that

$$\widetilde{G}_1(\gamma, M_1(e^{j\omega})) = \check{C}_{1,\gamma}(e^{j\omega}\check{E}_1 - \check{A}_1)^{-1}\check{B}_{1,\gamma}, \tag{82}$$

where

$$\begin{aligned} &\check{A}_1 = \begin{bmatrix} -A_1 & 0 \\ 0 & E \end{bmatrix}, \quad \check{E}_1 = \begin{bmatrix} -E & 0 \\ 0 & A_1 \end{bmatrix}, \\ &\check{B}_{1,\gamma} = \frac{1}{\sqrt{2}}\begin{bmatrix} -B_1 & -\gamma B_1 \\ -\gamma^{-1}B_1 & B_1 \end{bmatrix}, \quad \check{C}_{1,\gamma} = \frac{1}{\sqrt{2}}\begin{bmatrix} I & \gamma e^{j\omega} I \\ \gamma^{-1} I & -e^{j\omega} I \end{bmatrix}. \end{aligned} \tag{83}$$

The analogue discrete-time result to Theorem 24 is stated as follows.

Theorem 25. *Let $\gamma \in (0, 1]$ and $\xi > 0$ be given. Then, for every $\omega \in \mathbb{R}$, ξ is a singular value of $\widetilde{G}_1(\gamma, M_1(e^{j\omega}))$ if and only if $e^{j\omega}$ is a generalized eigenvalue of the symplectic pencil $z\mathscr{G}(\xi, \widetilde{G}_1(\gamma, M_1)) - \mathscr{F}(\xi, \widetilde{G}_1(\gamma, M_1))$.*

The proof follows immediately by applying Proposition 23 to $\widetilde{G}_1(\gamma, M_1(e^{j\omega}))$ in (82) and (83).

6.2.3. Key ideas

Theorems 24 and 25 reduce the computation of the singular values of $N_1(\lambda, \gamma)$ at a given frequency $\lambda = j\omega$ or $\lambda = e^{j\omega}$ to the computation of the generalized eigenvalues of a corresponding Hamiltonian or symplectic pencil.

Denote by $f(\omega)$ either $j\omega$ or $e^{j\omega}$. Define

$$\begin{aligned} \widehat{\mu}(\omega) := \mu_{\mathbb{R}}(M_1(f(\omega))) &= \inf_{\gamma\in(0,1]} \sigma_2\big(\widetilde{G}_1(\gamma, M_1(f(\omega)))\big) \\ &= \inf_{\gamma\in(0,1]} \sigma_2\left(N_1(f(\omega), \gamma)\right) = \sigma_2\big(\widetilde{G}_1(\gamma_\omega^*, M_1(f(\omega)))\big). \end{aligned}$$

The goal of the algorithm is to maximize $\widehat{\mu}(\omega)$ over $\omega \in \mathbb{R}$ since

$$\widehat{\mu}^* := \sup_{\omega\in\mathbb{R}} \widehat{\mu}(\omega) = r_{\mathbb{R}}^{-1}(P; \mathbb{C}_g; \Delta_1),$$

as relation (48) shows.

Assume that such a unique maximizer exists and let

$$\omega^* := \arg \max_{\omega \in \mathbb{R}} \widehat{\mu}(\omega).$$

Suppose that at each iteration $k = 0, 1, \ldots \xi_{k-1}$ is the best known lower bound to $\widehat{\mu}^*$ so far and let ω_k be the current trial frequency. Suppose further that ω^* is known to lie in a certain "maximizing" open set Ω_k. At each iteration, one has to perform two basic steps (see Figs. 1 and 2).

First compute the optimal γ at the current frequency ω_k

$$\gamma_k^* = \arg \inf_{\gamma \in (0,1]} \sigma_2\big(N_1(f(\omega_k), \gamma)\big).$$

Despite the higher dimension of M_1 this only involves at each step the SVD of a $2n \times 2n$ matrix, as relation (54) shows. Moreover, by denoting $\lambda_k = f(\omega_k)$ one has that

$$\sigma_2\big((L_1(\lambda_k, \gamma) \otimes I_n)N_{P^{-1}(\lambda_k)}(\gamma)\big) = \sigma_{n-1}^{-1}\big(N_{P(\lambda_k)}(\gamma)(L_1(\lambda_k, \gamma)^{-1} \otimes I_n)\big).$$

Hence the computation of the real and imaginary parts of $P^{-1}(\lambda_k)$ is replaced by a simple inversion of a 2 by 2 upper or lower triangular matrix $L_1(\lambda_k, \gamma)$. Thus

$$\gamma_k^* = \arg \inf_{\gamma \in (0,1]} \sigma_{n-1}^{-1} \left(N_{P(\lambda_k)}(\gamma)\ (L_1(\lambda_k, \gamma)^{-1} \otimes I_n)\right). \tag{84}$$

The *second* step consists in finding an improved lower bound to $\widehat{\mu}^*$, as well as the next "maximizing" set Ω_{k+1} and within a new trial frequency point ω_{k+1}. If $\widehat{\mu}(\omega_k) > \xi_{k-1}$, take $\xi_k = \widehat{\mu}(\omega_k) = \sigma_2(\widetilde{G}_1(\gamma_k^*, M_1(f(\omega_k))))$ as the new estimate of μ^*, otherwise keep the old estimate, that is, $\xi_k = \xi_{k-1}$. Locate now the "level-set" of frequencies, say Ω'_{k+1}, defined as

$$\Omega'_{k+1} = \big\{\omega \in \mathbb{R} : \sigma_2\big(N_1(f(\omega), \gamma_k^*)\big) = \sigma_2\big(\widetilde{G}_1(\gamma_k^*, M_1(f(\omega)))\big) > \xi_k\big\}.$$

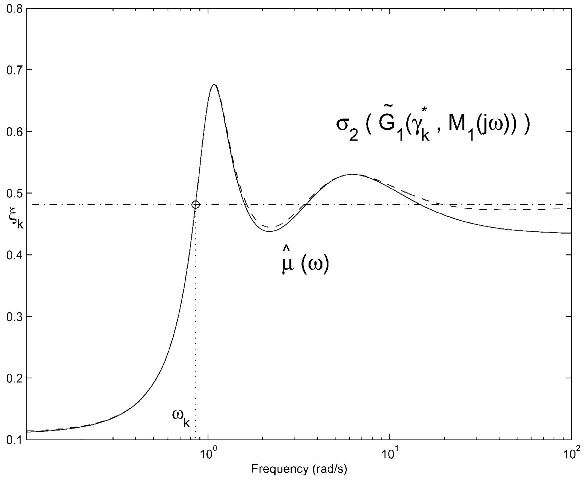

Fig. 1.

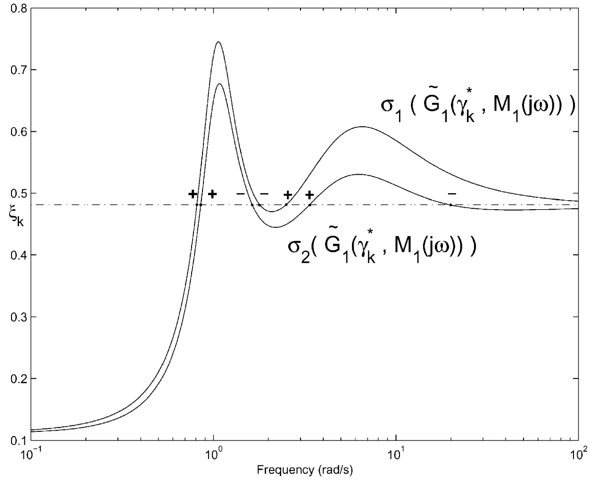

Fig. 2.

By Theorem 24 (or Theorem 25), the pure imaginary (or unitary) eigenvalues of the Hamiltonian (or symplectic) pencil $s\mathscr{L}(\xi_k, \widetilde{G}_1(\gamma_k^*, M_1)) - \mathscr{H}(\xi_k, \widetilde{G}_1(\gamma_k^*, M_1))$ (or $z\mathscr{G}(\xi_k, \widetilde{G}_1(\gamma_k^*, M_1)) - \mathscr{F}(\xi_k, \widetilde{G}_1(\gamma_k^*, M_1))$) are exactly those ω for which some singular value of $\widetilde{G}_1(\gamma_k^*, M_1(f(\omega)))$ equals ξ_k. The endpoints of the frequency intervals where $\sigma_2(\widetilde{G}_1(\gamma_k^*, M_1(f(\omega))))$ equals or exceeds ξ_k must be among these and can be identified using derivative information of the imaginary (or unitary) generalized eigenvalues.

Let $(\mathscr{A}, \mathscr{E}) \in \{(\mathscr{H}, \mathscr{L}), (\mathscr{F}, \mathscr{G})\}$. If λ_l is the lth generalized eigenvalue, assumed simple, of $\mathscr{A} - \lambda\mathscr{E}$, then

$$\frac{\partial\lambda_l}{\partial\xi}(\xi_k) = \frac{u_l^*\left(\frac{\partial\mathscr{A}}{\partial\xi}(\xi_k) - \lambda_l\frac{\partial\mathscr{E}}{\partial\xi}(\xi_k)\right)v_l}{u_l^*\mathscr{E}v_l}. \tag{85}$$

Here v_l and u_l are a pair of right and left eigenvectors associated to λ_l and are automatically obtained when computing the generalized eigenvalues of the pencil $\mathscr{A} - \lambda\mathscr{E}$.

By using formula (85) one can deduce that

$$\tilde{s}_l = \frac{\partial\xi}{\partial\omega}(\widetilde{\omega}_l) = \left(-\mathrm{j}\frac{\partial\lambda_l}{\partial\xi}(\xi_k)\right)^{-1} \quad \text{for } \lambda_l = \mathrm{j}\tilde{\omega}_l \tag{86}$$

and

$$\check{s}_l = \frac{\partial\xi}{\partial\omega}(\check{\omega}_l) = \left(-\mathrm{j}\mathrm{e}^{-\mathrm{j}\check{\omega}_l}\frac{\partial\lambda_l}{\partial\xi}(\xi_k)\right)^{-1} \quad \text{for } \lambda_l = \mathrm{e}^{\mathrm{j}\check{\omega}_l}. \tag{87}$$

By using equalities (86) and (87) in conjunction with (85), one can actually prove that $\tilde{s}_l$ and $\check{s}_l$ are both *real*. The trick of the proof is the relation between u_l and v_l.

Because of the Hamiltonian and symplectic structure of the considered pencils, one can show, respectively, that

$$\tilde{u}_l = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \tilde{v}_l \quad \text{and} \quad \check{u}_l = \begin{bmatrix} 0 & I \\ \mathrm{e}^{\mathrm{j}\omega} I & 0 \end{bmatrix} \check{v}_l.$$

The sign of $\tilde{s}_l$ (or $\check{s}_l$) at different crossing points $\tilde{\omega}_l$ (or $\check{\omega}_l$) can now be used to determine the σ_2-intervals of interest. For complete details see [15,18].

Since any frequency point ω *not* in Ω'_{k+1} satisfies

$$\widehat{\mu}(\omega) \leqslant \sigma_2\big(\widetilde{G}_1(\gamma_k^*, M_1(f(\omega)))\big) \leqslant \xi_k,$$

the global maximizer ω^* cannot lie outside Ω'_{k+1}, if $\Omega'_{k+1} \neq \emptyset$. Thus, by setting $\Omega_{k+1} = \Omega'_{k+1} \cap \Omega_k$, we can bracket ω^* at every iteration. Several possibilities to choose ω_{k+1} in Ω_{k+1} are proposed in [15,18]. For instance, set ω_{k+1} equal to the mid-point of the largest interval contained in Ω_{k+1}.

Algorithm.
Input: $P_0, P_1, \ldots, P_k$. Tolerance $\tau > 0$.
Output: $r_{\mathbb{R}}(P; \mathbb{C}_{\mathrm{g}}; \Delta_1)$, $\omega^* = \arg\max_{\omega\in\mathbb{R}} \widehat{\mu}(\omega)$.
Initialization: $k = 0$, pick ω_0, $\xi_0 = \mu_{\mathbb{R}}(M_1(f(\omega_0)))$, $\Omega_0 = (0, \infty)$.
1. Compute γ_k^*, ξ_k.
2. Compute $\Omega_{k+1} = \Omega'_{k+1} \cap \Omega_k$.
3. Compute ω_{k+1}.
4. $k \longleftarrow k + 1$. If an appropriate stopping criterion (in terms of τ) is satisfied STOP. Otherwise GOTO 1.

Step 1 involves a golden section search over γ. At each iteration one has to compute a SVD of an $n \times n$ matrix (see (84)). If r is the number of steps required by the search over γ, then the complexity of Step 1 is $\mathrm{O}(n^3 r)$. For instance, in order to obtain a four-digit accuracy on γ_{opt}, one needs about $r = 20$ iterations on the golden section search. The complexity of Step 2 is that of a Hamiltonian or symplectic eigenvalue problem of dimension $2nk$ (see (80)) or $2n(k+1)$ (see (83)), that is, $\mathrm{O}((2nk)^3)$ or $\mathrm{O}((2n(k+1))^3)$.

Numerical tests suggest that the rate of convergence is quadratic; conditions under which this can be proved are under investigation.

7. Conclusions

In this paper, an efficient computational scheme to compute the real (unstructured) stability radius of non-monic polynomial matrices has been presented. We adapted the numerical algorithm proposed in [18] to deal with generalized state-space realizations. This enables us to consider both non-monic polynomials and polynomial matrices. The proposed approach can be extended immediately to deal with structured stability radius computation as well.

Several problems are clearly left open. A first important goal would consist in extending the result of Theorem 3 to arbitrary *p*-norms. Secondly, one should improve the optimization scheme over γ as it shows up in relation (12).

Obtaining closed formulae for the real stability radius in the Δ_3 case is known to be a difficult problem in the μ literature. Nevertheless it is hoped that an appropriate design of efficient optimization schemes could be of significant help in that respect.

Acknowledgements

We thank to the anonymous reviewer for pointing out the generalization of Theorem 13 mentioned in Remark 15.

This paper presents research supported by the Belgian Programme on Inter-university Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. The scientific responsibility rests with its authors.

This work was supported by the National Science Foundation under Grant No. CCR 99-12415.

The second author has been partially supported by NATO.

References

[1] S. Boyd, V. Balakrishnan, A regularity result for the singular values of a transfer matrix and a quadratically convergent algorithm for computing its L_∞ norm, Systems Control Lett. 15 (1990) 1–7.

[2] R. Byers, A bisection method for measuring the distance of a stable matrix to the unstable matrices, SIAM J. Sci. Statist. Comput. 9 (1988) 875–881.

[3] H. Chapellat, S.P. Bhattacharyya, A generalization of Kharitonov's theorem: robust stability of interval plants, IEEE Trans. Automat. Control 34 (1997) 306–311.

[4] J.C. Doyle, G. Stein, Multivariable feedback design: concepts for a classical/modern synthesis, IEEE Trans. Automat. Control 26 (1981) 4–16.

[5] Y. Genin, P. Van Dooren, V. Vermaut, Convergence of the calculation of $\mathscr{H}_\infty$-norms and related questions, in: Proceedings MTNS-98, Padua, July 1998.

[6] Y. Genin, P. Van Dooren, Stability radii of polynomial matrices, in: Proceedings IEEE CACSD Conference, Hawai, Paper MM1-4, August 1999.

[7] A. Halanay, Vl. Rasvan, Stability radii for some propagation models, IMA J. Math. Control Inform. 14 (1997) 95–107.

[8] F. Tisseur, N.J. Higham, Structured pseudospectra for polynomial eigenvalue problems, with applications, SIAM J. Matrix Anal. Appl. 23 (2001) 187–208.

[9] D. Hinrichsen, A.J. Pritchard, Stability radii of linear systems, Systems Control Lett. 7 (1986) 1–10.

[10] D. Hinrichsen, N.K. Son, The complex stability radius of discrete-time systems and symplectic pencils, in: Proceedings of the IEEE 28th Conference on Decision and Control, Tampa, FL, 1989, pp. 2265–2270.

[11] D. Hinrichsen, A.J. Pritchard, Robustness measures for linear systems with application to stability radii of Hurwitz and Schur polynomials, Int. J. Control 55 (4) (1992) 809–844.

[12] D. Hinrichsen, B. Kelb, Spectral value sets: a graphical tool for robustness analysis, Systems Control Lett. 21 (1993) 127–136.

[13] L. Qiu, E.J. Davison, A simple procedure for the exact stability robustness computation of polynomials with affine coefficient perturbations, Systems Control Lett. 13 (1989) 413–420.

[14] L. Qiu, B. Bernhardsson, A. Rantzer, E.J. Davison, P.M. Young, J.C. Doyle, A formula for the computation of the real stability radius, Automatica 31 (6) (1995) 879–890.

[15] C.T. Lawrence, P. Van Dooren, A.L. Tits, A fast algorithm for the computation of an upper bound on the μ-norm, Automatica 36 (2000) 449–456.

[16] G. Pappas, D. Hinrichsen, Robust stability of linear systems described by higher order dynamic equations, IEEE Trans. Automat. Control 38 (9) (1993) 1430–1435.

[17] A.J. Pritchard, S. Townley, Robustness of linear systems, J. Differential Equations 77 (1989) 254-286.

[18] J. Sreedhar, P. Van Dooren, A.L. Tits, A fast algorithm to compute the real structured stability radius, Internat. Ser. Numer. Math. 121 (1996) 219–230.

[19] L.N. Trefethen, Pseudospectra of linear operators, SIAM Rev. 39 (3) (1997) 383–406.

[20] Ch. Van Loan, How near is a stable matrix to an unstable matrix? Contemp. Math., AMS Contemp. Math. Series 47 (1985) 465–478.

[21] F. Wirth, Robust stability of discrete-time systems under time-varying perturbations, Dissertation, Bremen, Germany, 1995.