Vietnam Journal of Mathematics **35**:4(2007) 429–462

Vietnam Journal
of
MATHEMATICS

Robust Stability and Transient Behaviour of Positive Linear Systems

D. Hinrichsen and E. Plischke

Zentrum für Technomathematik, Universität Bremen,
D-28334 Bremen, Germany

Dedicated to Professor Hoang Tuy on the occasion of his 80th-birthday

Received August 22, 2007

Abstract. After a brief review of available results the main focus of the paper is on the transient behaviour of positive systems and their stability radii with respect to highly structured perturbations. Simple upper bounds for the transient gain of positive systems are obtained by means of linear Lyapunov functions on the positive orthant. The minimization of these bounds is discussed and algorithms for computing optimal Lyapunov vectors are presented. By means of linear Lyapunov functions we get new formulae for the stability radii of positive linear systems with respect to structured and time-varying perturbations of Gershgorin-Brualdi type. With every time-invariant linear system we associate a corresponding positive system and this correspondence allows to transfer some of the results to non-positive linear systems.

2000 Mathematicals Subject Classification: 37C65, 47A55, 34D20
Keywords: Positive systems, transient behaviour, stability radii, Lyapunov norms, structured uncertainties, time-varying perturbations.

1. Introduction

Stability and stabilization are fundamental concepts in linear systems theory and in most design problems, exponential stability is the minimal requirement that has to be met. However, it is often not enough to achieve this property. Suppose that the stabilized feedback system is described by a time-invariant linear model

$$\dot{x}(t) = Ax(t), \quad t \in \mathbb{R}, \tag{1}$$

where $A \in \mathbb{K}^{n \times n}$, $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. This model will in general not yield an accurate description of the behaviour of the regulated real plant. The system parameters may be uncertain so that the real system is better described by a family of perturbed systems $\dot{x}(t) = A(\Delta)x(t)$, where the perturbation parameter Δ is bounded by a given uncertainty level $\|\Delta\| < \delta$. Or the model (1) is obtained by linearization of a nonlinear system around a given operating point so that a more realistic model would be $\dot{x}(t) = Ax(t) + N(x(t))$ where $N : \mathbb{K}^n \to \mathbb{K}^n$ is a nonlinear perturbation. In both cases it will not be sufficient that the nominal system (1) be exponentially stable but this property must be guaranteed for the whole set of perturbed systems where the perturbations are bounded by some realistic uncertainty level. In this case the system (1) is said to be robustly stable (for the given perturbation class and perturbation bound).

There is another reason why exponential stability may not be sufficient. By definition, exponential stability requires that there exist two constants $\beta < 0$, $\gamma \geq 0$ such that $\|x(t)\| \leqslant \gamma e^{\beta t}\|x(0)\|$ for all solutions of (1). Thus exponential stability guarantees that initial deviations from the equilibrium $x = 0$ die away exponentially *in the long run*. But it does not guarantee a satisfactory *transient* behaviour (if γ is large). Trajectories of an exponentially stable linear system may temporarily move a long way from the origin before approaching it as $t \to \infty$. From a practical point of view, if the "state excursions" are very large, the stable system actually behaves like an unstable one. Moreover, if the system is obtained by linearization of a nonlinear system around an equilibrium point the large transients of the linear part may incite the nonlinearities to drive the system permanently far away from the equilibrium. In such cases the practical instability of the equilibrium point is reflected by an extreme thinness of its domain of attraction. The interaction between large transient motions of the linearization and existing nonlinearities in the system has been put forward as an explanation for the stark contrast between some experimental results and theoretical predictions in fluid dynamics, see for example [21].

The aim of this paper is to study these two intimately related problems, *robust stability* and *transient behaviour*, for *positive* exponentially stable systems of the form (1). By definition, a system (1) is called *positive* if it leaves the positive orthant $\mathbb{R}^n_+$ invariant, i.e., e^{At} is a nonnegative matrix for all $t \geq 0$.

Robust stability and robust stabilization have been dominant themes in systems and control theory over the past three decades. Different approaches have been developed and a large number of papers and books are available on this subject, we only mention H^∞-optimal control theory [22], the parametric approach based on Kharitonov's result on interval polynomials [3], the theory of stability radii and μ-analysis [7]. For positive systems, robustness issues have been discussed to a much lesser extent, but there are a number of satisfactory results available on stability radii of positive systems with respect to "full block" and "block-diagonal" perturbations, see [9, 20]. In this paper we consider *Gershgorin–Brualdi* perturbations of positive systems where the matrix entries at an arbitrarily prescribed set of positions are independently perturbed. Recent results in [12] are extended from diagonal to general positive nominal systems. We also determine stability radii with respect to *time-varying* parameter perturbations. This is in general a difficult problem, but for positive systems computable formulae can be obtained. Up till now similar results have only been obtained for diagonal systems, see [8].

The transient gain of linear systems has recently attracted some attention, mainly in the stability analysis of fluid dynamics [21] and in numerical analysis [5]. In control theory, the transient behaviour should play an important role since it is related to such classical criteria as "overshoot" of systems responses and to problems of state constraints and saturation effects. Nevertheless, the study of the transient behaviour of linear systems has just only begun in the context of control theory [7] and we do not know of any paper on the *transient behaviour of positive linear systems*. In this paper we present a number of estimates for the transient gain of positive systems and show how upper bounds for the transient behaviour of an arbitrary linear system can be obtained from bounds for the corresponding Metzler system.

The paper is organized as follows. In the next section we present some auxiliary results on the properties of Metzler matrices. Most of these results are well known and follow from the Perron–Frobenius theory of nonnegative matrices, but some of them are not found in the literature and those results will be proved. In Sec. 3 we briefly review available results about stability radii for full block perturbations and point out the dramatic simplification of the results if positive systems are considered. In Sec. 4, Sec. 5 and Sec. 6 new results are presented concerning the transient behaviour of positive systems and their stability radii with respect to structured perturbations. In Sec. 4 the concepts of *Lyapunov norm* and *Lyapunov vector* are introduced and it is shown how they can be used to derive upper bounds for the transient gain of positive systems. The problem of minimizing these upper bounds is discussed and two algorithms for optimizing the estimates are presented. In Sec. 5 we determine stability radii and transient bounds for positive systems with respect to Gershgorin–Brualdi perturbations. As a preparation for the study of *time-varying* linear or nonlinear perturbations we derive necessary and sufficient criteria for the existence of joint linear Lyapunov functions for sets of positive linear systems. Finally we study positive differential inclusions and determine the stability radius of positive systems with respect to *time-varying* Gershgorin–Brualdi perturbations.

2. Preliminaries

A matrix $P \in \mathbb{R}^{m\times n}$ is said to be nonnegative ($P \geq 0$) if all its entries are nonnegative. The set of all nonnegative $m \times n$-matrices is denoted by $\mathbb{R}_+^{m\times n}$. For $A, B \in \mathbb{R}^{m\times n}$ we write $A \geq B$ if $A - B \geq 0$. Similar notations will be used for vectors in $\mathbb{R}^n$. So inequalities between real matrices (resp. vectors) will be understood componentwise. If $x = (x_i) \in \mathbb{C}^n$ and $A = (a_{ij}) \in \mathbb{C}^{m\times n}$ then $|x| \in \mathbb{R}_+^n$ and $|A| \in \mathbb{R}_+^{m\times n}$ are defined by $|x| = (|x_i|)$, $|A| = (|a_{ij}|)$. Clearly,

$$|A+B| \leqslant |A| + |B|,\ A, B \in \mathbb{C}^{m\times n} \ \text{ and }\ |AB| \leqslant |A||B|,\ A \in \mathbb{C}^{m\times n},\ B \in \mathbb{C}^{n\times p}.$$

For any matrix $A \in \mathbb{C}^{n\times n}$ the *spectral radius* (resp. *spectral abscissa*) of A are defined by

$$\varrho(A) = \max\{|\lambda|; \lambda \in \sigma(A)\} \quad (\text{resp. } \alpha(A) = \max\{\mathrm{Re}\lambda; \lambda \in \sigma(A)\}),$$

where $\sigma(A) \subset \mathbb{C}$ is the spectrum of A. The spectral radius has the following monoticity property [10, Subsec. 8.1],

$$\forall A \in \mathbb{C}^{n\times n}, B \in \mathbb{R}_+^{n\times n}: \ |A| \le B \ \Rightarrow \ \varrho(A) \le \varrho(|A|) \le \varrho(B). \tag{2}$$

The linear system $\dot{x}(t) = Ax(t)$, $A \in \mathbb{R}^{n\times n}$ is said to be *positive* if the positive orthant $\mathbb{R}_+^n$ is invariant for the corresponding flow, i.e., $e^{At} \ge 0$ for all $t \ge 0$. In this case, the matrix A is called a *Metzler matrix.* $A \in \mathbb{R}^{n\times n}$ is a Metzler matrix if and only if all the off-diagonal entries of A are nonnegative, i.e., $tI + A \ge 0$ for some $t \ge 0$. As a consequence, the spectral abscissa of a Metzler matrix has analogous properties to the spectral radius of a nonnegative matrix. The convex cone of $n \times n$ Metzler matrices is denoted by $\mathbb{R}_{\mathrm{M}}^{n\times n}$. The following results follow from the Perron–Frobenius theory of nonnegative matrices, see [10, Ch. 8], [2], [20].

Theorem 1. *Suppose that $A \in \mathbb{R}_{\mathrm{M}}^{n\times n}$ is a Metzler matrix. Then*

(i) *$\alpha(A)$ is an eigenvalue of A and there exists a nonnegative eigenvector $x \ge 0$, $x \ne 0$ (called Perron vector of A) such that $Ax = \alpha(A)x$. If $A \ge 0$ then $\alpha(A) = \varrho(A) \ge 0$. If A is irreducible then there exists $x > 0$ such that $Ax = \alpha(A)x$.*

(ii) *If $\lambda \ne \alpha(A)$ is any other eigenvalue of A then* $\mathrm{Re}\lambda < \alpha(A)$.

(iii) *Given $\beta \in \mathbb{R}$, there exists a nonzero vector $x \ge 0$ such that $Ax \ge \beta x$ if and only if $\alpha(A) \ge \beta$.*

(iv) *$(tI - A)^{-1}$ exists and is nonnegative if and only if $t > \alpha(A)$. Moreover,*

$$\alpha(A) < t_1 \leqslant t_2 \quad \Rightarrow \quad 0 \leqslant (t_2 I - A)^{-1} \leqslant (t_1 I - A)^{-1}.$$

If A is irreducible then $(tI - A)^{-1} > 0$ for all $t > \alpha(A)$.

With every matrix $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ we associate the Metzler matrix

$$M(A) = \mathrm{Re}D(A) + |A - D(A)| \tag{3}$$

called the *Metzler part* of A where $D(A) = \mathrm{diag}(a_{11}, \dots, a_{nn})$. An elementary proof shows the following

Lemma 1. *Let $A \in \mathbb{C}^{n\times n}$, then*

(i) *The function $r \mapsto \varrho(A + rI_n) - r$ is monotonically decreasing on $\mathbb{R}_+$ and*

$$\alpha(A) = \lim_{r\to\infty} (\varrho(A + rI_n) - r). \tag{4}$$

(ii) *The map $r \mapsto M_r(A) := |A + rI_n| - rI_n$ is componentwise decreasing on $\mathbb{R}_+$ and*

$$M(A) = \lim_{r\to\infty} |A + rI_n| - rI_n. \tag{5}$$

Proof. (i) For every $\lambda \in \mathbb{C}$ we have

$$0 \leqslant r_1 \leqslant r_2 \;\Rightarrow\; |\lambda + r_2| - r_2 = |\lambda + r_1 + (r_2 - r_1)| - r_2 \leqslant |\lambda + r_1| - r_1, \quad (6)$$

and since $(1+t)^{1/2} \leqslant 1 + t/2$ for $t \in \mathbb{R}, |t| < 1$ we get for $r > 0$ sufficiently large

$$|\lambda + r| - r = r\left(\sqrt{1 + \frac{2\mathrm{Re}\lambda}{r} + \frac{|\lambda|^2}{r^2}} - 1\right) \leq r\left(\frac{\mathrm{Re}\lambda}{r} + \frac{|\lambda|^2}{2r^2}\right),$$

whence

$$\lim_{r\to\infty} (|\lambda + r| - r) = \mathrm{Re}\lambda, \quad \lambda \in \mathbb{C}. \quad (7)$$

Now $\varrho(A + rI_n) - r = \max\{|\lambda + r| - r; \lambda \in \sigma(A)\}$ and so the monoticity property of $r \mapsto \varrho(A + rI_n) - r$ follows directly from (6), whilst (4) follows from (7).
(ii) Applying (6) to the diagonal entries of $M_r(A)$ we see that these entries decrease monotonically with increasing $r \geq 0$ towards the limits $\mathrm{Re}\, a_{ii}$ whereas the off-diagonal entries $|a_{ij}|$ of $M_r(A)$ remain constant. This proves (ii) by definition of $M(A)$ in (3). ∎

As a consequence we obtain the following monoticity property which is a counterpart to (2) for the spectral abscissa,

$$\forall A \in \mathbb{C}^{n\times n},\ B \in \mathbb{R}^{n\times n}_{\mathrm{M}} :\ M(A) \leq B \;\Rightarrow\; \alpha(A) \leq \alpha(M(A)) \leqslant \alpha(B). \quad (8)$$

In fact, we have by (2), Theorem 1 (i), the previous lemma and the continuity of the spectral abscissa that

$$\begin{aligned}\alpha(A) &= \lim_{r\to\infty} \varrho(A + rI_n) - r \leqslant \lim_{r\to\infty} \varrho(|A + rI_n|) - r \\ &= \lim_{r\to\infty} \alpha(|A + rI_n| - rI_n) = \alpha(M(A))\end{aligned}$$

which proves the first inequality in the conclusion of (8). The second inequality follows directly from (2) since by Theorem 1 (i) we have for any Metzler matrix $M \in \mathbb{R}^{n\times n}_{\mathrm{M}}$

$$\alpha(M) = \alpha(M + rI_n) - r = \varrho(M + rI_n) - r, \quad r \in \{t \geq 0; M + tI_n \geq 0\}.$$

A norm $\|\cdot\|$ on $\mathbb{K}^n$ is said to be *monotone* if it satisfies

$$|x| \leq |y| \;\Rightarrow\; \|x\| \leq \|y\|, \quad x, y \in \mathbb{C}^n.$$

For instance, every p-norm $\|\cdot\|_p$ on $\mathbb{C}^n$, $1 \leq p \leq \infty$ is monotone. Unfortunately, the operator norm $\|\cdot\|$ associated with a pair of monotone vector norms need not be monotone. However, we have the following properties, see [20].

Lemma 2. *Suppose that $\mathbb{C}^m, \mathbb{C}^n$ are provided with monotone norms and $\|\cdot\|$ denotes the corresponding operator norm on $\mathbb{C}^{m\times n}$. Then*

(i) *For every $P \in \mathbb{R}^{m\times n}_+$ there exists $u \in \mathbb{R}^n_+$, $\|u\|_{\mathbb{C}^n} = 1$ such that $\|Pu\|_{\mathbb{C}^m} = \|P\|$.*
(ii) *If $P \in \mathbb{C}^{m\times n}$, $Q \in \mathbb{R}^{m\times n}_+$ and $|P| \leq Q$, then $\|P\| \leq \|\,|P|\,\| \leq \|Q\|$.*
(iii) *If $P \in \mathbb{C}^{m\times n}$ is of rank one then $\|P\| = \|\,|P|\,\|$.*

3. Robust Stability of Positive Systems

In this section we give a brief survey of available results concerning robust stability of positive systems. The results illustrate the dramatic simplification of robustness analysis obtained by positivity assumptions.

A linear system $\dot{x} = Ax$, $A \in \mathbb{C}^{n\times n}$ is exponentially stable if and only if the system matrix A is Hurwitz stable, i.e., $\sigma(A) \subset \mathbb{C}_- = \{s \in \mathbb{C}; \mathrm{Re}s < 0\}$ or, equivalently, $\alpha(A) < 0$. There are various ways to verify this property of A in the general case. One consists in computing the eigenvalues of A and checking whether they all belong to $\mathbb{C}_-$. Another one is to solve the Lyapunov equation $A^*X + XA = -I_n$ and to check that the solution X is positively definite. Still another possibility is to determine the characteristic polynomial of A and apply the classical algebraic stability criteria (Hermite, Routh or Hurwitz tests) to this polynomial. All these methods are computationally demanding and may pose serious numerical problems. Under the condition of positivity much simpler stability criteria are available. The next theorem collects some well known stability criteria for positive systems of the form (1). They follow easily from the results in the previous section, see also [14].

Theorem 2. *For a Metzler matrix $A \in \mathbb{R}^{n\times n}_{\mathrm{M}}$ the following conditions are equivalent:*

(i) $\sigma(A) \subset \mathbb{C}_-$.
(ii) *A is invertible and $-A^{-1} \geq 0$.*
(iii) *For every $b \in \mathbb{R}^n_+$ there exists $\tilde{x} \in \mathbb{R}^n_+$ such that $A\tilde{x} + b = 0$.*
(iv) *There exists a vector $b \in \mathbb{R}^n_+$ with strictly positive coordinates and $\tilde{x} \in \mathbb{R}^n_+$ such that $A\tilde{x} + b = 0$.*

The last stability test is particularly simple. It shows that there exists a nonnegative equilibrium point $\tilde{x}(b)$ of the positive system $\dot{x} = Ax + b$ for some (any) $b > 0$ if and only if A is Hurwitz stable. So if the equilibrium point $\tilde{x}(b) = -A^{-1}b$ corresponding to $b > 0$ exists and is nonnegative then the system is automatically exponentially stable.

To state the next lemma for later use we employ the following notation. If $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ and I, J are two non-empty subsets of $\underline{n} := \{1, \ldots, n\}$ we denote by $A(I, J)$ the submatrix of A consisting of the entries a_{ij} with $i \in I, j \in J$, i.e., $A(I, J) = (a_{ij})_{i\in I, j\in J}$. If $J \subset \underline{n}$ is any nonempty index set we write $A(J)$ for the principal submatrix $A(J, J)$ and denote by $|J|$ the number of elements of J, by J' the complementary index set $J' = \underline{n} \setminus J$.

Lemma 3. *Suppose that the Metzler matrix $A \in \mathbb{R}^{n\times n}_{\mathrm{M}}$ is Hurwitz stable. Then its principal submatrices $A(J)$ where $J \subset \underline{n}$, $0 < |J| < n$, and their Schur complements in A, $A(J') - A(J', J)A(J)^{-1}A(J, J')$, are Hurwitz stable Metzler matrices, too. The principal submatrices of $-A^{-1} \geq 0$ are in bijective correspondence with the negative inverses of these Schur complements via the formula*

$$\left[-A^{-1}\right](J') = -\left[A(J') - A(J', J)A(J)^{-1}A(J, J')\right]^{-1}, \quad J \subset \underline{n},\ 0 < |J| < n. \tag{9}$$

Proof. That the principal submatrices of A are Hurwitz stable Metzler matrices follows from the corresponding result for M-matrices, see [11, §2.5]. To prove the second statement we may assume without restriction of generality that the principal submatrix sits in the upper left corner of A. Then A can be partitioned into $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ where $A_{11} \in \mathbb{R}_{\mathrm{M}}^{k \times k}$ is the principal submatrix under consideration and both A_{11} and $A_{22} \in \mathbb{R}_{\mathrm{M}}^{(n-k)\times(n-k)}$ are Hurwitz stable and hence regular. An easy calculation shows that A can be factorized as follows.

$$A = \begin{bmatrix} A_{11} & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} -I & 0 \\ A_{21} & -I \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \begin{bmatrix} -I & A_{12} \\ 0 & -I \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & -I \end{bmatrix}. \tag{10}$$

Since A is invertible, so is the Schur complement $\tilde{A} = A_{22} - A_{21}A_{11}^{-1}A_{12}$ of A_{11} in A. Now $-A_{11}^{-1} \geq 0$ by Theorem 2, hence $-A_{21}A_{11}^{-1}A_{12} \geq 0$ and so $\tilde{A} \in \mathbb{R}_{\mathrm{M}}^{(n-k)\times(n-k)}$. Let $\tilde{z}$ be a Perron vector of the Metzler matrix $\tilde{A}$. Then we have $z := \begin{bmatrix} -A_{11}^{-1}A_{12}\tilde{z} \\ \tilde{z} \end{bmatrix} \geq 0$, and by (10),

$$Az = \begin{bmatrix} -A_{11} & 0 \\ -A_{21} & I \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & \tilde{A} \end{bmatrix} \begin{bmatrix} 0 \\ \tilde{z} \end{bmatrix} = \alpha(\tilde{A}) \begin{bmatrix} 0 \\ \tilde{z} \end{bmatrix}.$$

But this implies $\alpha(\tilde{A}) < 0$, since otherwise $Az \geq 0$ and this would imply $\alpha(A) \geq 0$ by Theorem 1 (iii). Therefore $\tilde{A}$ is a Hurwitz stable Metzler matrix. Finally, $-A^{-1} \geq 0$ follows by Theorem 2 and the formula (9) can be found in [10, §0.7.3]. ∎

We will now assume that the positive system $\dot{x} = Ax$ is exponentially stable and investigate the robustness of its stability with respect to perturbations of the output feedback type

$$A \rightsquigarrow A + B\Delta C, \quad \Delta \in \mathbb{K}^{l \times q}, \tag{11}$$

where $B \in \mathbb{R}_+^{n \times l}$ and $C \in \mathbb{R}_+^{q \times n}$ are arbitrary given nonnegative matrices and Δ is an unknown disturbance matrix in $\mathbb{K}^{l \times q}$, either real ($\mathbb{K} = \mathbb{R}$) or complex ($\mathbb{K} = \mathbb{C}$). This is the so-called *full block case* where Δ is varying in the whole matrix space $\mathbb{K}^{l \times q}$. In Sec. 5 we will consider more structured perturbations where Δ is constrained to some linear subspace $\boldsymbol{\Delta} \subset \mathbb{K}^{l \times q}$.

In the following we assume that the size of Δ is measured by its operator norm $\|\Delta\|$ with respect to a given pair of *monotone* norms on $\mathbb{K}^q, \mathbb{K}^l$. According to whether complex or real disturbances are considered, two distinct *stability radii* of (1) with respect to perturbations of the form (11) are introduced:

$$r_{\mathbb{K}} = r_{\mathbb{K}}(A; B, C) = \inf\{\|\Delta\|; \Delta \in \mathbb{K}^{l \times q}, \alpha(A + B\Delta C) \geq 0\}, \quad \mathbb{K} = \mathbb{R}, \mathbb{C}. \tag{12}$$

We call a perturbation Δ *destabilizing* if $\alpha(A+B\Delta C) \geq 0$. In case there does not exist a destabilizing (complex or real) disturbance, the corresponding stability radius is defined to be ∞.[1] Otherwise, by continuity of the spectrum, there exists a minimum norm destabilizing disturbance in $\mathbb{C}^{l\times q}$ or $\mathbb{R}^{l\times q}$, respectively. Clearly, we always have

$$0 < r_{\mathbb{C}}(A;B,C) \leqslant r_{\mathbb{R}}(A;B,C).$$

In the general case where $A \in \mathbb{R}^{n\times n}$, $\sigma(A) \subset \mathbb{C}_-$, $B \in \mathbb{R}^{n\times l}$ and $C \in \mathbb{R}^{q\times n}$, the quotient $r_{\mathbb{R}}/r_{\mathbb{C}}$ may be arbitrarily large. The complex stability radius can be expressed via the transfer matrix $G(s)$ associated with the triplet (A,B,C)

$$r_{\mathbb{C}}(A;B,C) = \left[\max_{\omega\in\mathbb{R}} \|G(\imath\omega)\|\right]^{-1}, \quad G(s) = C(sI_n - A)^{-1}B. \tag{13}$$

A formula for the real stability radius exists only for the special case where $\mathbb{R}^l$ and $\mathbb{R}^q$ are provided with their Euclidean norms and accordingly $\|\Delta\|$ is the spectral norm. The formula for $r_{\mathbb{R}}$ available in this case is considerably more complicated than (13) and requires the solution of a maxmin problem with two parameters, see [7]. The situation is completely different if the nominal system (1) and the structure matrices B, C are nonnegative. In this case the real and the complex stability radii coincide and can be expressed by a simple computable formula which does not require any optimization [20].

Theorem 3. *Suppose that $A \in \mathbb{R}^{n\times n}$ is a Hurwitz stable Metzler matrix, $(B,C) \in \mathbb{R}_+^{n\times l} \times \mathbb{R}_+^{q\times n}$ are given nonnegative structure matrices and $\mathbb{C}^l$, $\mathbb{C}^q$ are provided with monotone norms. Then, with respect to the induced operator norms $\|\cdot\|$ on $\mathbb{K}^{l\times q}$ and $\mathbb{K}^{q\times l}$,*

$$r_{\mathbb{C}}(A;B,C) = r_{\mathbb{R}}(A;B,C) = \|CA^{-1}B\|^{-1}.$$

There is, in general, an important difference between the real and the complex stability radius if time-varying and/or nonlinear perturbations are considered. We explain this for the case where $\mathbb{K}^l$, $\mathbb{K}^q$ are both provided with their Euclidean norms and $\|\cdot\|$ denotes the corresponding operator norm on $\mathbb{K}^{l\times q}$. There are many examples where, e.g., *time-varying real* perturbations $\Delta(t)$ of norm $\sup_{t\geq 0}\|\Delta(t)\| < r_{\mathbb{R}}$ can destabilize a system whereas it is known that exponential stability will be preserved for all time-varying complex perturbations of norm $\sup_{t\geq 0}\|\Delta(t)\| < r_{\mathbb{C}}$. This latter result is based on the fact that in this case $r_{\mathbb{C}}$ can be characterized by a parameterized algebraic Riccati equation of the form

$$A^*P + PA - \rho^2 C^*C - PBB^*P = 0, \tag{14}$$

where $\rho > 0$. It was shown in [6] that (14) admits a Hermitian solution P_ρ satisfying $\sigma(A - BB^*P_\rho) \subset \mathbb{C}_-$ if and only if $\rho < r_{\mathbb{C}}$. If additionally the pair (A,C) is observable, $V(x) = -\langle x, P_\rho x\rangle$ yields a joint Lyapunov function for all

[1] Throughout the paper we set $\inf\emptyset = \infty$.

perturbed systems $\dot{x} = (A + B\Delta C)x$ where $\Delta \in \mathbb{C}^{l\times q}$, $\|\Delta\| < \rho$. This Lyapunov function works also for time-varying and nonlinear perturbations bounded by ρ.

To make this more precise, let (A, B, C) be as above and consider a perturbed system of the form

$$\dot{x}(t) = Ax(t) + BN(Cx(t), t), \quad t \geq 0, \tag{15}$$

where the perturbation $N : \mathbb{K}^q \times \mathbb{R}_+ \to \mathbb{K}^l$ is an unknown time-varying nonlinearity such that $N(y, t)$ is continuous on $\mathbb{K}^q \times \mathbb{R}_+$ and continuously differentiable in y. We assume that N is *of finite gain*, i.e., there exists $\gamma \geq 0$ such that

$$\|N(y, t)\| \leq \gamma\|y\|, \quad y \in \mathbb{K}^q, \ t \geq 0.$$

The size of such a nonlinear perturbation $N(\cdot, \cdot)$ is measured by the norm

$$\|N\| = \inf\{\gamma \in \mathbb{R}_+; \forall y \in \mathbb{K}^q \ \forall t \in \mathbb{R}_+ : \|N(y, t)\| \leq \gamma\|y\|\}.$$

By $x^N(\cdot; t_0, x_0)$ we denote the solution of (15) on $[t_0, \infty)$ starting at $x(t_0) = x_0 \in \mathbb{C}^n$. Then the following result can be proved, see [20].

Corollary 1. *Suppose that (A, B, C) is as in Theorem 3, $\mathbb{C}^l, \mathbb{C}^q$ are provided with their standard Euclidean norms and $\|\cdot\|$ denotes the induced operator norms. If $\rho < \|CA^{-1}B\|^{-1}$ then there exists a joint quadratic Lyapunov function for all the systems (15) with $\|N\| \leqslant \rho$ and the joint equilibrium $\overline{x} = 0$ is globally exponentially stable for all these systems.*

Another way of accounting for parameter uncertainties in systems of the form (1) is to view the system matrix as member of a *matrix interval*. Given two arbitrary real matrices $A, B \in \mathbb{R}^{n\times n}$ such that $A \geq B$, then the matrix interval

$$[[B, A]] := \{X \in \mathbb{R}^{n\times n};\ B \leq X \leq A\}$$

is called *Hurwitz stable* if $\sigma(X) \subset \mathbb{C}_-$ for each matrix $X \in [[B, A]]$. The problem then is to find a small set of test matrices in $[[B, A]]$ whose Hurwitz stability suffices to ensure the Hurwitz stability of the whole given matrix interval. A famous theorem of Kharitonov [13] says that the Hurwitz stability of an interval of monic polynomials can be verified by examining only four of its vertex polynomials. It is well-known that there does not exist a counterpart of this result for interval matrices. The stability of all corner matrices of the polytope $[[B, A]]$ do not imply stability of the whole interval. This changes dramatically if the interval consists of Metzler matrices. Then it suffices to check just the upper corner A of the interval. We conclude the section with this simple but illustrative result.

Theorem 4. *If A is a Hurwitz stable Metzler matrix then every matrix in the set*

$$\mathcal{A} = \{X \in \mathbb{R}^{n\times n} : M(X) \leqslant A\}$$

is Hurwitz stable. In particular if $A, B \in \mathbb{R}^{n\times n}_{\mathrm{M}}$ are arbitrary Metzler matrices and $A \geq B$, then $[[B, A]]$ is Hurwitz stable if and only if A is Hurwitz stable.

Proof. By (8) we have $\alpha(M(X)) \leqslant \alpha(A) < 0$ for all $X \in \mathcal{A}$ if A is Hurwitz stable. ■

4. Transient Behaviour of Positive Systems

As mentioned in the introduction, an exponentially stable linear system may temporarily exhibit an unstable behaviour. In this section we will derive upper bounds to the transient behaviour of positive systems and we will see that positivity not only greatly simplifies the analysis of robust stability (as shown in the previous section) but also the analysis of transient behaviour. Lyapunov functions and, in particular, *Lyapunov norms* provide a key tool for bounding the transient behaviour of a system, see [7, 16, 17]. For positive systems simple transient bounds can be obtained by means of *linear* Lyapunov functions.

4.1. (γ, β)-Stability and Contractions

We begin by introducing a stability concept which imposes not only conditions on the long-term behaviour, but also on the transient behaviour of the system. *Throughout this section we suppose that* $\mathbb{C}^n$ *is endowed with an arbitrary vector norm* $\|\cdot\|$ *and the matrix space* $\mathbb{C}^{n\times n}$ *is provided with the associated operator norm which is also denoted by* $\|\cdot\|$.

Definition 1. *Let* $\beta < 0$ *and* $\gamma \geq 1$ *be given. A linear system* $\dot{x} = Ax$ *is said to be* (γ, β)*-stable with respect to the norm* $\|\cdot\|$ *if*

$$\|e^{At}\| \leq \gamma e^{\beta t}, \quad t \geq 0. \tag{16}$$

In contrast to exponential stability, where only the *existence* of some constants $\beta < 0$, $\gamma \geq 1$ satisfying (16) is required, the *a priori* prescription of both γ and β imposes not only conditions on the long-term behaviour, but also on the transient behaviour of the system.

A lower bound for the exponent β is given by $\alpha(A)$ and therefore the possible values of β merely depend on the spectrum of A while, for a given value of β, the values of γ depend heavily on the used norm. As a result, the concept of (γ, β)-stability – unlike the concept of exponential stability – is norm-dependent. If A is stable then the minimum value of γ such that A is $(\gamma, 0)$-stable is called the *transient gain* of A,

$$\gamma(A) = \max_{t\geq 0} \|e^{At}\|. \tag{17}$$

The optimally achievable transient gain $\gamma = 1$ is obtained if the system matrix A of (1) generates a *contraction semigroup*, i.e., $\|e^{At}\| \leqslant 1$ holds for all $t \geq 0$.

Definition 2. *The initial growth rate*[2] *of* $A \in \mathbb{C}^{n\times n}$ *with respect to* $|\cdot|$ *is defined*

[2] $\mu(A)$ is also called *logarithmic derivative*, *logarithmic norm* or *matrix measure* in the literature.

by

$$\mu(A) = \mu_{\|\cdot\|}(A) = \min\left\{\mu \in \mathbb{R};\, \forall t \geq 0 :\, \|e^{At}\| \leq e^{\mu t}\right\}. \tag{18}$$

The name "initial growth rate" is due to the following characterization which shows that $\mu(A)$ only depends upon e^{At} for $t \in [0, \varepsilon]$, $\varepsilon > 0$ arbitrarily small, see [7. Prop. 5.5.8]

$$\begin{aligned}\mu(A) &= \frac{d^+}{dt}\|e^{At}\|\,\Big|_{t=0} = \lim_{t\downarrow 0}\frac{1}{t}\log\|e^{At}\| \\ &= \lim_{h\downarrow 0} h^{-1}\left(\|I + hA\| - 1\right) = \lim_{t\to\infty}\|A + tI\| - t.\end{aligned}$$

The initial growth rate enjoys the following properties, see [7]. For all matrices $A, B \in \mathbb{C}^{n\times n}$ and all $\alpha > 0$, $z \in \mathbb{C}$

$$|\mu(A)| \leq \|A\|, \quad \mu(\alpha A) = \alpha\mu(A),$$
$$\mu(A + zI) = \mu(A) + \operatorname{Re} z, \quad \mu(A + B) \leq \mu(A) + \mu(B).$$

Definition 3. *A norm $\|\cdot\|$ on $\mathbb{C}^n$ is called a (strict) Lyapunov norm for the system (1) or the matrix $A \in \mathbb{C}^{n\times n}$ if the initial growth rate of A with respect to the norm $\|\cdot\|$ satisfies $\mu(A) \leq 0$ (resp. $\mu(A) < 0$).*

The following characterization of the contraction property is easily proved.

Lemma 4. *For every $A \in \mathbb{C}^{n\times n}$ the following statements are equivalent:*

(i) *The matrix A generates a contraction semigroup with respect to $\|\cdot\|$.*
(ii) *For every $x^0 \in \mathbb{C}^n$ and $t \geq 0 :\, \|x(t; x^0)\| \leq \|x^0\|$ where $x(t; x^0) = e^{At}x^0$.*
(iii) *The unit ball $\mathbb{B} = \{x \in \mathbb{C}^n;\, \|x\| \leq 1\}$ is forward invariant under the flow of $\dot{x} = Ax$.*
(iv) *The norm $\|\cdot\|$ is a Lyapunov norm for $\dot{x} = Ax$.*

Computable formulae for the initial growth rates associated with some standard norms are readily available. Given $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ the initial growth rates of A with respect to the norms $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$ are given by [7, Lemma 5.5.11]

$$\mu_1(A) = \max_j\left(\operatorname{Re} a_{jj} + \sum_{i\neq j}|a_{ij}|\right), \quad \mu_\infty(A) = \max_i\left(\operatorname{Re} a_{ii} + \sum_{j\neq i}|a_{ij}|\right),$$
$$\mu_2(A) = \frac{1}{2}\max\{\lambda;\, \lambda \in \sigma(A + A^*)\} = \tfrac{1}{2}\,\alpha(A + A^*).$$

We see that the initial growth rates with respect to the 1- or ∞-norms are determined with particular ease, and for these norms the initial growth rates do not change if the matrix $A \in \mathbb{C}^{n\times n}$ is replaced by its Metzler part $M(A)$ (see (3)):

$$\begin{aligned}\mu_1(A) &= \mu_1(M(A)) = \max_j\,(\mathbf{1}_n^\top M(A))_j, \\ \mu_\infty(A) &= \mu_\infty(M(A)) = \max_i\,(M(A)\mathbf{1}_n)_i,\end{aligned} \tag{19}$$

where $\mathbf{1}_n = (1,\dots,1)^\top \in \mathbb{R}^n_+$. In particular, $A \in \mathbb{C}^{n\times n}$ generates a contraction semigroup with respect to $\|\cdot\|_1$ or $\|\cdot\|_\infty$ if and only if $M(A)$ generates a contraction semigroup with respect to these norms, and this happens if and only if $\mathbf{1}_n^\top M(A) \leqslant 0$ or $M(A)\mathbf{1}_n \leqslant 0$, respectively. Stable Metzler matrices A with these properties are *column* or *row diagonal dominant*, respectively. The 1-norm (resp. ∞-norm) is a joint Lyapunov function for all the systems (1) with a column (resp. row) diagonally dominant system matrix A.

We will now relate the transient gain of a matrix to the transient gain of its Metzler part. Note that if A is a Metzler matrix then $e^{At} \geq 0$ for all $t \geq 0$ and so the operator norm of e^{At} with respect to the 1- or ∞-norms are given by (see [10, 5.6.4, 5.6.5])

$$\|e^{At}\|_1 = \max_j(\mathbf{1}_n^\top e^{At})_j \quad \text{and} \quad \|e^{At}\|_\infty = \max_i(e^{At}\mathbf{1}_n)_i.$$

Thus, in order to determine a transient bound for a Metzler matrix A with respect to the 1- or ∞-norms, it suffices to determine the solution of $\dot{x} = A^\top x$ (resp. $\dot{x} = Ax$) starting at $x^0 = \mathbf{1}_n$.

Theorem 5. (i) *If $A, B \in \mathbb{R}^{n\times n}_M$ and $A \leqslant B$ then $e^A \leqslant e^B$ (componentwise).*
(ii) *For every $A \in \mathbb{C}^{n\times n}$ the following inequalities hold (componentwise),*

$$|e^{At}| \leq e^{M(A)t} = \lim_{r\to\infty} e^{(|A+rI|-rI)t} \leqslant e^{(|A+rI|-rI)t}, \;\; t \geq 0, \; r \geq 0. \tag{20}$$

Proof. (i) Choose $r > 0$ such that $0 \leqslant A + rI \leqslant B + rI$. Then (i) follows from

$$e^r e^A = e^{(A+rI)} = \sum_{k=0}^{\infty} \frac{(A+rI)^k}{k!} \leqslant \sum_{k=0}^{\infty} \frac{(B+rI)^k}{k!} = e^r e^B.$$

(ii) For all $t \geq 0$ and $r \in \mathbb{R}$ we obtain

$$e^{rt}|e^{At}| = |e^{(A+rI)t}| \leq \sum_{k=0}^{\infty} \frac{|(A+rI)t|^k}{k!} = e^{|A+rI|t}.$$

The continuity of the matrix exponential and Lemma 1 yield

$$|e^{At}| \leq \lim_{r\to\infty} e^{(|A+rI|-rI)t} = e^{M(A)t}, \;\; t \geq 0.$$

Moreover, as the map $r \mapsto M_r(A) := |A+rI_n|-rI_n$ is componentwise decreasing on $\mathbb{R}_+$ by Lemma 1, it follows from (i) that $e^{M(A)t} = \lim_{r\to\infty} e^{(|A+rI|-rI)t} \leqslant e^{(|A+rI|-rI)t}$ for every $r \geq 0$. ∎

We conclude this subsection relating the (γ,β)-stability of an arbitrary complex matrix A with that of $M(A)$.

Corollary 2. *Suppose $\|\cdot\|$ is a monotone vector norm on $\mathbb{C}^n$, $A \in \mathbb{C}^{n\times n}$ and $M(A)$ is (γ,β)-stable with respect to $\|\cdot\|$, $\gamma \geq 1, \beta \leqslant 0$. Then $A \in \mathbb{C}^{n\times n}$*

is (γ, β)*-stable and the associated initial growth rates and transient gains* (*with respect to the norm* $\|\cdot\|$) *satisfy*

$$\mu(A) \leq \mu(M(A)) \quad \textit{and} \quad \gamma(A) \leqslant \gamma(M(A)). \tag{21}$$

Proof. Applying the corresponding operator norm $\|\cdot\|$ to the results of Theorem 5 and using Lemma 2 we get

$$\|e^{At}\| \leq \||e^{At}|\| \leq \|e^{M(A)t}\| \leq \gamma e^{\beta t}, \quad t \geq 0, \tag{22}$$

which shows that A is also (γ, β)-stable and $\gamma(A) \leqslant \gamma(M(A))$. To prove the first inequality in (21) note that, by definition, $\|e^{M(A)\,t}\| \leqslant e^{\mu(M(A))\,t}$ for all $t \geq 0$. Hence (22) implies $\|e^{At}\| \leq e^{\mu(M(A))\,t}$, $t \geq 0$, and so $\mu(A) \leq \mu(M(A))$ by (18). ∎

4.2. Weighted Norms and Eccentricity

If the system $\dot{x} = Ax$ does not generate a contraction semigroup with respect to the given norm $\|\cdot\|$ then we may possibly introduce a suitable weighted norm $\|x\|_W = \|Wx\|$ which provides a Lyapunov norm for the system under investigation. By the equivalence of norms on $\mathbb{C}^n$, there exist constants which relate the weighted norm to the original norm. Making use of these constants one can derive estimates for the transient gain (17) of the system with respect to the original norm.

Definition 4. *Suppose* $\nu(\cdot)$ *and* $\|\cdot\|$ *are norms on* $\mathbb{C}^n$*. Then the eccentricity of* $\nu(\cdot)$ *with respect to* $\|\cdot\|$ *is defined by*

$$\mathrm{ecc}(\nu) = \mathrm{ecc}(\nu, \|\cdot\|) = \frac{\max_{\|x\|=1} \nu(x)}{\min_{\|x\|=1} \nu(x)}.$$

Knowing the eccentricity of $\nu(\cdot)$ with respect to $\|\cdot\|$ and the initial growth rate of a matrix A with respect to the norm $\nu(\cdot)$ one obtains an exponential estimate for $\|e^{At}\|$, $t \geq 0$ by the following result.

Theorem 6. *If* $\nu(\cdot)$ *is any norm on* $\mathbb{C}^n$*,* $A \in \mathbb{C}^{n\times n}$*, and* $\mu_\nu(A)$ *denotes the initial growth rate of* A *with respect to the norm* $\nu(\cdot)$*, then*

$$\|e^{At}\| \leq ecc(\nu)\, e^{\,\mu_\nu(A)t}, \quad t \geq 0.$$

In particular, if $\mu_\nu(A) \leqslant 0$ *then* (1) *is* $(ecc(\nu), \mu_\nu(A))$*-stable.*

Proof. From (18) we obtain the exponential estimate $\nu(e^{At}) \leq e^{\mu_\nu(A)t}$ where $\nu(e^{At})$ denotes the operator norm of e^{At} with respect to the auxiliary norm $\nu(\cdot)$. Moreover,

$$\min_{\|z\|=1} \nu(z) \leq \nu\left(\frac{y}{\|y\|}\right) \leq \max_{\|z\|=1} \nu(z), \quad y \in \mathbb{C}^n,\ y \neq 0.$$

This implies $\|y\| \min_{\|z\|=1} \nu(z) \leqslant \nu(y) \leqslant \|y\| \max_{\|z\|=1} \nu(z)$ for all $y \in \mathbb{C}^n,\ y \neq 0$ and hence for the associated operator norms of any $T \in \mathbb{C}^{n\times n}$

$$\|T\| = \sup_{x\neq 0} \frac{\|Tx\|}{\|x\|} \leq \sup_{x\neq 0} \left(\frac{\nu(Tx)}{\min_{\|z\|=1} \nu(z)}\right) \left(\frac{\nu(x)}{\max_{\|z\|=1} \nu(z)}\right)^{-1} = \mathrm{ecc}\nu \cdot \nu(T).$$

Setting $T = e^{At}$ gives the desired result. ∎

In the following we consider the case that ν is a weighted version of $\|\cdot\|$. For any invertible matrix $W \in \mathbb{C}^{n\times n}$ the weighted vector norm $\|x\|_W = \|Wx\|$ on $\mathbb{C}^n$ induces the operator norm $\|A\|_W = \|WAW^{-1}\|$ on $\mathbb{C}^{n\times n}$. From this fact and the definitions we obtain

Lemma 5. *Given $W \in Gl_n(\mathbb{C})$, the eccentricity of the weighted norm $\nu(x) = \|Wx\|$ on $\mathbb{C}^n$ with respect to the norm $\|\cdot\|$ equals the condition number of W with respect to $\|\cdot\|$,*

$$ecc(\nu, \|\cdot\|) = \kappa(W) = \|W\|\,\|W^{-1}\|.$$

The initial growth rate of $A \in \mathbb{C}^{n\times n}$ with respect to the weighted norm ν is given by

$$\mu_{\|\cdot\|,W}(A) := \mu_\nu(A) = \mu_{\|\cdot\|}(WAW^{-1}).$$

In particular, if $w \in \mathbb{R}^n_+, w > 0$ and $W = \mathrm{diag}(w_i)$ is the corresponding scaling matrix, the eccentricity of the scaled norm $\|\cdot\|_w := \|\cdot\|_W$ with respect to $\|\cdot\|$ is given by $\kappa(W) = \frac{\max_i w_i}{\min_i w_i}$. Note that the scaled norm $\|\cdot\|_w$ on $\mathbb{C}^n$ is monotone if $\|\cdot\|$ is monotone.

In the following theorem we apply the above technique to positive systems by introducing a scaled norm $\|\cdot\|_W$ where the diagonal of $W = \mathrm{diag}(w_i)$ is a positive Perron vector of A. For any $w > 0$ we call $\kappa(w) = (\max_i w_i)/(\min_i w_i)$ the *condition number of w*.

Theorem 7. *Suppose $A \in \mathbb{R}^{n\times n}$ is a Metzler matrix.*

(i) *If A has a strictly positive left Perron vector $w > 0$ then*

$$\|e^{At}\|_1 \leq \kappa(w) e^{\alpha(A)t}, \quad t \geq 0.$$

(ii) *If A has a strictly positive right Perron vector $v > 0$ then*

$$\|e^{At}\|_\infty \leq \kappa(v) e^{\alpha(A)t}, \quad t \geq 0.$$

(iii) *If A has strictly positive left and right Perron vectors $w > 0$ and $v > 0$ then*

$$\|e^{At}\|_2 \leq \left(\kappa((\tfrac{w_i}{v_i})_i)\right)^{\frac{1}{2}} e^{\alpha(A)t}, \quad t \geq 0.$$

Proof. (i) Suppose that $w > 0$ is a left Perron vector of A, i.e., $w^\top A = \alpha(A) w^\top$. Setting $W = \mathrm{diag}(w_i)$ gives $\mathbf{1}_n^\top WAW^{-1} = w^\top AW^{-1} = \alpha(A) w^\top \mathrm{diag}(w_i^{-1}) = \alpha(A)\mathbf{1}_n^\top$, hence the initial growth rate of A with respect to the scaled norm

$\|x\|_{1,W} = \|Wx\|_1$ is $\mu_{1,W}(A) = \mu_1(WAW^{-1}) = \alpha(A)$ (by Lemma 5 and (19)). Since the condition number of W is given by $\kappa(w)$, (i) follows from Theorem 7.

(ii) Analogously, if $v > 0$ is a right Perron vector of A then $W = \mathrm{diag}(v_i^{-1})$ gives $\mu_{\infty,W}(A) = \mu_\infty(WAW^{-1}) = \alpha(A)$ with condition number $\kappa((v_i^{-1})_i) = \kappa(v)$.

(iii) For the spectral norm, suppose that $w > 0$ (resp. $v > 0$) are left (resp. right) Perron vectors of A and set $D = \mathrm{diag}(\frac{w_i}{v_i})$. Then $DA + A^{\mathbb{T}}D - 2\alpha(A)D$ is a singular Hermitian Metzler matrix with eigenvector $v > 0$

$$(DA + A^\top D - 2\alpha(A)D)v = (\alpha(A)I + A^\top - 2\alpha(A)I)Dv = (A^\top - \alpha(A)I)w = 0.$$

It follows from [2, Cor. 2.1.12] that v is a Perron vector of $DA + A^\top D - 2\alpha(A)D$ corresponding to the eigenvalue $\alpha(DA + A^\top D - 2\alpha(A)D) = 0$. Hence

$$DA + A^\top D - 2\alpha(A)D \preceq 0,$$

where $\preceq$ denotes the order relation between Hermitian matrices ($B \preceq A$ if $A - B$ is symmetric and positive semidefinite). Applying [7, Prop. 5.5.33] we obtain (iii). ∎

Remark 1. The Metzler matrix A may have neither left nor right Perron vectors which are strictly positive. However, if A is irreducible then A has strictly positive left and right Perron vectors and these are uniquely determined modulo multiplication by a positive number, see [2].

The choice of Perron vectors as weights provides exponential estimates with optimal decay rate $\alpha(A)$ but the estimates for the transient gain obtained in this way may be far from optimal. Fortunately, the weights can be chosen from a much larger set.

Definition 5. *For a given Metzler matrix $A \in \mathbb{R}^{n\times n}_{\mathrm{M}}$ a strictly positive vector $w \in \mathbb{R}^n_+$ is called a left (or right) Lyapunov vector of A if $w^\top A \leq 0$ or $Aw \leq 0$, respectively. If strict inequalities hold, $w^\top A < 0$ or $Aw < 0$, then the Lyapunov vector w is called strict.*

We have the following geometric interpretation. The vector $w > 0$ is a left Lyapunov vector for $\dot{x} = Ax$ if and only if the unit ball of $\|\cdot\|_{1,w}$ restricted to the positive orthant, $\mathbb{B}^+_w = \{z \in \mathbb{R}^n_+ \,;\, w^\top z \leqslant 1\}$, is invariant under the flow of $\dot{x} = Ax$. The vector $w > 0$ is a strict left Lyapunov vector, if for all $x \in \partial\mathbb{B}^+_w := \{z \in \mathbb{R}^n_+ \;;\; w^\top z = 1\}$ the direction $\dot{x} = Ax$ at x points into the interior of $\mathbb{B}^+_w$ (relative to $\mathbb{R}^n_+$).

The next theorem shows that every left (resp. right) Lyapunov vector for A defines a weighted Lyapunov norm for $\dot{x} = Ax$.

Theorem 8. *Let $A \in \mathbb{R}^{n\times n}_{\mathrm{M}}$ be an arbitrary Metzler matrix. Then*

(i) *If A is exponentially stable then for every vector $b \in \mathbb{R}^n_+$ ($b > 0$) there exists a vector $w \in \mathbb{R}^n_+$ (resp. $w > 0$) such that $Aw = -b$.*

(ii) *If there exists a (strict)Lyapunov vector $w > 0$ of A then A is (exponentially) stable.*

(iii) *If $w > 0$ is a left Lyapunov vector of A then the scaled 1-norm $\|x\|_{1,w} = \sum_i w_i|x_i|$ is a Lyapunov norm for $\dot{x} = Ax$. Its eccentricity with respect to the 1-norm is given by $\kappa(w)$, the corresponding initial growth rate is $\mu_{1,w}(A) = \max_j \dfrac{(w^\top A)_j}{w_j}$ and we have the exponential estimate*

$$\|e^{At}\|_1 \le \kappa(w)\, e^{\max_j \frac{(w^\top A)_j}{w_j} t}, \quad t \ge 0.$$

(iv) *If $w > 0$ is a right Lyapunov vector of A and $w^{-1} = (w_i^{-1})$, then the scaled ∞-norm $\|x\|_{\infty,w^{-1}} = \max_i \dfrac{x_i}{w_i}$ is a Lyapunov norm for $\dot{x} = Ax$. Its eccentricity with respect to the ∞-norm is given by $\kappa(w)$, the corresponding initial growth rate is $\mu_{\infty,w^{-1}}(A) = \max_j \dfrac{(Aw)_j}{w_j}$ and we have the exponential estimate*

$$\|e^{At}\|_\infty \le \kappa(w)\, e^{\max_j \frac{(Aw)_j}{w_j} t}, \quad t \ge 0.$$

Proof. (i) Suppose that A is exponentially stable. Then $-A^{-1} \in \mathbb{R}_+^{n\times n}$ by Theorem 2 (ii). Hence $w = -A^{-1}b \in \mathbb{R}_+^{n\times n}$ and $w = -A^{-1}b > 0$ if $b > 0$.
(iv) Suppose that $w > 0$ and $b = -Aw \ge 0$. Setting $W = \mathrm{diag}(w_i^{-1})$ we obtain $WAW^{-1}\mathbf{1}_n = WAw = -Wb \le 0$. Hence the initial growth rate of A with respect to the weighted norm $\|x\|_{\infty,w^{-1}} = \max_j |x_j|/w_j$ is by Lemma 5 and (19)

$$\mu_{\infty,w^{-1}}(A) = \mu_{\infty,W}(A) = \mu_\infty(WAW^{-1}) = \max_j (WAW^{-1}\mathbf{1}_n)_j = \max_j \{-\tfrac{b_j}{w_j}\} \le 0.$$

This concludes the proof of (iv) by Theorem 6.

Now, (iii) is obtained by applying (iv) to $A^\top$, and (ii) follows directly from (iii), (iv). ∎

Remark 2. Note that $\|x\|_{1,w}(x) = w^\top x$ for $x \in \mathbb{R}_+^n$ and so $x \mapsto w^\top x$ can be viewed as a linear Lyapunov function for the positive system $\dot{x} = Ax$ on the positive orthant.

4.3. Optimizing Transient Bounds

In the previous subsection we have seen that there is a broad range of Lyapunov norms available for positive systems. Theorem 8 (iii) (iv) shows how to obtain a transient estimate for a given Lyapunov vector w. In order to obtain an optimal estimate of the transient gain we minimize the condition number $\kappa(w)$ on the convex cone of right Lyapunov vectors, $\mathbf{W} = \{w \in \mathbb{R}_+^n\,;\, w > 0, Aw \leqslant 0\}$. Throughout this subsection we assume that $A \in \mathbb{R}_{\mathrm{M}}^{n\times n}$ is a given exponentially stable Metzler matrix. Consider the following minimax problem

$$\text{Minimize}\, \kappa(w) = \frac{\max_i w_i}{\min_i w_i} \quad \text{subject to } \; w > 0,\; Aw \leqslant 0. \tag{23}$$

As $\kappa(w)$ is invariant under multiplication with positive scalars, the optimization problem may be restricted to a basis of the convex cone $\mathbf{W}$. For this we choose

$$\mathbf{W}_1 = \{w \in \mathbb{R}_+^n\,;\, \min_i w_i = 1, Aw \leqslant 0\}.$$

For any two vectors $x, y \in \mathbb{R}^n$, define $z = \min\{x, y\}$ by $z_i = \min\{x_i, y_i\}$, $i \in \underline{n}$.

Lemma 6. *The basis* $\mathbf{W}_1$ *is closed under the operation* min*,*

$$v, w \in \mathbf{W}_1 \quad \Rightarrow \quad \min\{v, w\} \in \mathbf{W}_1.$$

Proof. Suppose $v, w \in \mathbf{W}_1$ and $u = \min\{v, w\}$. Then $\min_i u_i = 1$. Since $A = (a_{ij})$ is a Hurwitz stable Metzler matrix, the diagonal entries a_{ii} are strictly negative and the off-diagonal entries a_{ij} $(i \neq j)$ are nonnegative. By assumption we have

$$\sum_{j \neq i} a_{ij} v_j \leqslant -a_{ii} v_i \quad \text{and} \quad \sum_{j \neq i} a_{ij} w_j \leqslant -a_{ii} w_i, \quad i \in \underline{n}.$$

Given any $i \in \underline{n}$ suppose that, for instance, $u_i = w_i$ then

$$\sum_{j \neq i} a_{ij} u_j = \sum_{j \neq i} a_{ij} \min\{v_j, w_j\} \leqslant \sum_{j \neq i} a_{ij} w_j \leqslant -a_{ii} w_i = -a_{ii} u_i.$$

Similarly, $\sum_{j \neq i} a_{ij} u_j \leqslant -a_{ii} u_i$ if $u_i = v_i$, and so we conclude that $\sum_{j \neq i} a_{ij} u_j \leqslant -a_{ii} u_i$ for all $i = 1, \ldots, n$, i.e., $Au \leqslant 0$. This proves $u \in \mathbf{W}_1$. ∎

Now consider the following linear program

$$\text{Minimize}\, \mathbf{1}_n^\top w = \sum_{i=1}^{n} w_i \ \text{ subject to } \ w_i \geq 1,\ (Aw)_i \leq 0,\ \ i = 1, \ldots, n. \tag{24}$$

Theorem 9. *Suppose that* $A \in \mathbb{R}_{\mathrm{M}}^{n \times n}$ *is a Hurwitz stable Metzler matrix.*

(i) *The linear program* (24) *has a unique optimal solution* $\hat{w}$ *and this optimal solution of* (24) *is an optimal solution of* (23) *satisfying* $\hat{w} \in \mathbf{W}_1$ *and* $\kappa(\hat{w}) = \min_{w \in \mathbf{W}_1} \kappa(w)$.

(ii) $\hat{w} \in \mathbf{W}_1$ *is an optimal solution of* (24) *if and only if* $H(\hat{w}) \cup J(\hat{w}) = \underline{n}$ *where*

$$H(w) = \{h \in \underline{n}\,;\, (Aw)_h = 0\}, \quad J(w) = \{i \in \underline{n}\,;\, w_i = 1\}, \quad w \in \mathbf{W}_1. \tag{25}$$

Proof. (i) By Theorem 8 (i) the set of admissible solutions of the linear program (24) is non-empty, and the objective functional $\mathbf{1}_n^\top w$ is bounded below on this set. Hence there exists an optimal solution of (24). Its uniqueness follows from Lemma 6.

Now let $\hat{w}$ be the optimal solution of (24). By optimality $\min_i \hat{w}_i = 1$ whence $\hat{w} \in \mathbf{W}_1$. Suppose that there is a vector $w \in \mathbf{W}_1$ such that $\kappa(w) = \max_i w_i < \max_i \hat{w}_i = \kappa(\hat{w})$. Then $u = \min\{w, \hat{w}\} \in \mathbf{W}_1$ by Lemma 6 and so u is an admissible solution of the linear program (24) satisfying $u \leqslant \hat{w}$ and $u \neq \hat{w}$. But this implies $\mathbf{1}_n^\top u < \mathbf{1}_n^\top \hat{w}$, which is a contradiction. Hence $\kappa(w) \geq \kappa(\hat{w})$ for all $w \in \mathbf{W}_1$, i.e., $\hat{w}$ is optimal for problem (23).

(ii) Suppose that $w \in \mathbf{W}_1$, $H(w) \cup J(w) \neq \underline{n}$ and let $i \in \underline{n} \setminus (H(w) \cup J(w))$. Then $w_i > 1$ and $\sum_{j \neq i} a_{ij} w_j < -a_{ii} w_i$. Since A is a Hurwitz stable Metzler matrix we have $a_{ij} \geq 0$ for $j \neq i$ and $-a_{ii} > 0$. Define $w' \in \mathbb{R}^n_+$ by $w'_j = w_j$ for $j \neq i$ and $w'_i = \max\{1, (-a_{ii})^{-1} \sum_{j \neq i} a_{ij} w_j\} < w_i$. Then $\min_k w'_k = 1$, $Aw' \leqslant 0$, i.e., $w' \in \mathbf{W}_1$, and $\mathbf{1}_n^\top w' < \mathbf{1}_n^\top w$. Therefore w is not an optimal solution of (24).

Conversely, let $\hat{w} \in \mathbf{W}_1$ and $H(\hat{w}) \cup J(\hat{w}) = \underline{n}$. Suppose that $w \in \mathbf{W}_1$ is the optimal solution of (24). Then $\mathbf{1}_n^\top w \leqslant \mathbf{1}_n^\top \hat{w}$ and it follows from Lemma 6 that $w \leqslant \hat{w}$ and therefore $J(w) \supset J(\hat{w})$. Permuting the coordinates of $w, \hat{w}$ and the rows and columns of A accordingly we may assume that $H(\hat{w}) = \{1, \dots, k\}$, $J(\hat{w}) \supset \{k+1, \dots, n\}$. Partitioning A, $w, \hat{w}$ we have

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \ \ A_{11} \in \mathbb{R}^{k \times k}_{\mathrm{M}},$$

$$\hat{w} = \begin{bmatrix} \hat{w}^1 \\ \mathbf{1}_{n-k} \end{bmatrix}, \ \ \hat{w}^1 \in \mathbb{R}^k_+, \ \ w = \begin{bmatrix} w^1 \\ \mathbf{1}_{n-k} \end{bmatrix}, \ \ w^1 \in \mathbb{R}^k_+.$$

Since $A_{11}\hat{w}^1 + A_{12}\mathbf{1}_{n-k} = 0$ and $A_{11} w^1 + A_{12}\mathbf{1}_{n-k} \leqslant 0$, it follows that $A_{11}(\hat{w}^1 - w^1) \geq 0$. The principal submatrix A_{11} is a Hurwitz stable Metzler matrix by Lemma 3 and so $-A_{11}^{-1} \geq 0$. Thus $A_{11}(\hat{w}^1 - w^1) \geq 0$ implies $\hat{w}^1 - w^1 \leqslant 0$. On the other hand we have $w \leqslant \hat{w}$ and so $w^1 \leqslant \hat{w}^1$. Therefore $w = \hat{w}$ and hence $\hat{w}$ is the optimal solution of (24). ∎

Remark 3. (i) Note that the converse of (i) in Theorem 9 is not true. There are, in general, optimal solutions of the minimax problem (23) in $\mathbf{W}_1$ which do not solve the linear program (24). In particular, (23) may have more than one optimal solution in $\mathbf{W}_1$.

(ii) If $H(\hat{w})$ is non-empty – and this is always the case if A is not row diagonally dominant – then $\mu_{\infty, \hat{w}^{-1}}(A) = \max_j \frac{(A\hat{w})_j}{\hat{w}_j} = 0$ and the exponential estimate provided by Theorem 8 reduces to $\|e^{At}\|_\infty \leq \kappa(\hat{w})$ for $t \geq 0$.

It follows from Theorem 9 (i) that the simplex method can be used to compute a right Lyapunov vector $\hat{w}$ with a minimal condition number $\kappa(\hat{w})$. The following algorithm presents a different procedure towards solving the minimax problem (23) and is based on Theorem 9 (ii). We make use of the function findweight($\cdot$) which is defined recursively and given by the following pseudo-Matlab code.

Algorithm 1. Let $S \in \mathbb{R}^{m \times m}_+$ be an invertible matrix such that $M := -S^{-1}$ is a Hurwitz stable Metzler matrix.

```
function y=findweight(S)
 m = number_of_rows(S); # m,k,x,y,z,J,S_J are local variables
 x = solve(S,ones(m,1));  # Solve S x=one_m
 J = find(x>0);           # get indices of positive entries
 k = length(J);           # number of positive entries
 if(k<m)                  # there are nonpositive entries in x
  S_J = S(J,J);           # pass to principal submatrix
  z = findweight(S_J);    # recursive call
  y = zeros(m,1);
  y(J) = z;               # fill entries indexed by J
 else
  y = x;                  # use positive solution
 end
 return y;                # y is nonnegative, at most k positive entries
end
```

We employ this recursive formulation in order to reduce the index bookkeeping operations to a minimum. Usage and properties of this algorithm are discussed in the following corollary.

Corollary 3. *Let* $A \in \mathbb{R}_{\mathrm{M}}^{n\times n}$ *be a Hurwitz stable Metzler matrix and set* $R = -A^{-1}$. *Then* $\hat{w} = Ry$ *is an optimal solution for problems* (23) *and* (24) *where* $y = \mathrm{findweight}(R)$ *is computed by Algorithm 1.*

Proof. In order to guarantee that each step of Algorithm 1 works correctly, we show that if the argument $S \in \mathbb{R}_+^{m\times m}$ of findweight(S) is a regular nonnegative matrix such that $M := -S^{-1} \in \mathbb{R}_{\mathrm{M}}^{m\times m}$ is a Hurwitz stable Metzler matrix, the index set J is nonempty, and the submatrix $S(J)$ of S is also a regular nonnegative matrix such that $M^J = (-S(J))^{-1} \in \mathbb{R}_{\mathrm{M}}^{k\times k}$ is Hurwitz stable. By assumption, S is regular, hence $Sx = \mathbf{1}_m$ has a unique solution x. This vector x contains at least one positive entry as $S \geq 0, \mathbf{1}_m > 0$ and therefore $J \neq \emptyset$. In the case that this index set has a size $k = |J|$ less than m we have to examine the submatrix $S(J) \in \mathbb{R}_+^{k\times k}$ of S. By construction, it is a principal submatrix of $S = -M^{-1}$, hence given by the negative inverse of a Schur complement $\tilde{M}$ in M, see Lemma 3. Also by Lemma 3, $\tilde{M} \in \mathbb{R}_{\mathrm{M}}^{k\times k}$ is a Hurwitz stable Metzler matrix. As we start the algorithm with $S = R = -A^{-1}$, each recursive call of findweight$(\cdot)$ has an argument S which is regular, nonnegative, and has the property that $-S^{-1}$ is a Hurwitz stable Metzler matrix.

Clearly, the size of S strictly decreases as the recursion proceeds. The recursion is terminated when the linear system $Sx = \mathbf{1}_m$ has a positive solution $x > 0$. This is always so if $m = 1$ as then $S \in \mathbb{R}_+$ is some diagonal entry of R which is invertible by Lemma 3 and therefore positive, i.e., $x = S^{-1} > 0$. The algorithm therefore terminates the recursion after at most n steps.

If we encounter a positive solution x at the end of the recursion, then this vector is on return augmented with zeros to give a final return value $y \in \mathbb{R}_+^n$. By construction this y satifies $(Ry)_i = 1$ for all $i \in \underline{n}$ such that $y_i \neq 0$. Hence, when setting $\hat{w} = Ry$ we have $J \cup H = \underline{n}$ where $J = \{i \in \underline{n}\,;\, \hat{w}_i = 1\} = \{i \in$

$\underline{n}$; $(Ry)_i = 1\}$ and $H = \{i \in \underline{n}\,;\, (A\hat{w})_i = 0\} = \{i \in \underline{n}\,;\, y_i = 0\}$. To prove that $\hat{w} = Ry$ is an optimal solution of (24) it remains to show that $\hat{w} \in \mathbf{W}_1$, see Theorem 9 (ii). By construction, we have $A\hat{w} = -y \le 0$. To prove $\hat{w} = Ry \ge \mathbf{1}_n$, let us consider $z = -A\mathbf{1}_n$ which solves $Rz = \mathbf{1}_n$. Note that $J = \{i \in \underline{n}\,;\, z_i > 0\}$ is the set of positive indices computed in the first step of the algorithm, whence $\{i \in \underline{n}\,;\, y_i > 0\} \subset J$. We assume that after a suitable permutation $J = \{i \in \underline{n}\,;\, z_i > 0\} = \{1, \dots, m\}$, and that R, z, y are partitioned accordingly into

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix},\ z = \begin{bmatrix} z^1 \\ z^2 \end{bmatrix},\ y = \begin{bmatrix} y^1 \\ 0 \end{bmatrix}.$$

Then we have $z^1 > 0$, $z^2 \le 0$, $y^1 \ge 0$, $R_{11}z^1 + R_{12}z^2 = \mathbf{1}_m = R_{11}y^1$, and $R_{21}z^1 + R_{22}z^2 = \mathbf{1}_{n-m}$. The principal submatrices R_{11} and R_{22} of R are regular by Lemma 3, moreover $\tilde{R} = R_{22} - R_{21}R_{11}^{-1}R_{12}$ is a Schur complement in R and therefore the inverse of a principal submatrix of $R^{-1} = -A$ by [10, §0.7.3]. Hence $\tilde{R} \ge 0$ by Lemma 3 and Theorem 2 (ii). We conclude that $R_{22} \ge R_{21}R_{11}^{-1}R_{12}$ and

$$w^2 := R_{21}y^1 = R_{21}\left(z^1 + R_{11}^{-1}R_{12}z^2\right) \ge R_{21}z^1 + R_{22}z^2 = \mathbf{1}_{n-m}$$

(because of $z^2 \le 0$). This shows that $\hat{w} = \begin{bmatrix} \mathbf{1}_m \\ w^2 \end{bmatrix} = Ry \ge \mathbf{1}_n$ and therefore $\hat{w} \in \mathbf{W}_1$. ■

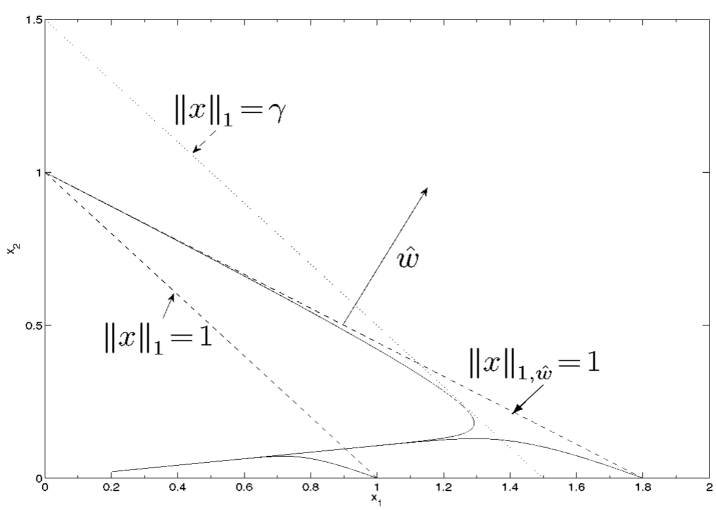

Fig. 1. Linear Lyapunov function.

Example 1. Let

$$A = \begin{bmatrix} -5 & 36 \\ 2 & -20 \end{bmatrix}, \quad R = -A^{-\top} = \frac{1}{28}\begin{bmatrix} 20 & 2 \\ 36 & 5 \end{bmatrix}.$$

Figure 1 shows some trajectories of the system $\dot{x} = Ax$ and the lines $\{x \in \mathbb{R}^2_+; \|x\|_1 = 1\}$, $\{x \in \mathbb{R}^2_+; \|x\|_1 = \gamma\}$, $\{x \in \mathbb{R}^2_+; \|x\|_{1,\hat{w}} = 1\}$, where $\hat{w} = (1, 1.8)^{\mathbb{T}}$

is a left Lyapunov vector of A with minimal condition number $\kappa(\hat{w})$. To calculate this optimal *left* Lyapunov vector we start Algorithm 1 with $R = -A^{-\top}$ and proceed as follows: For $S = R$, the equation $Sx = \mathbf{1}_2$ has the solution $x = -A^\top \mathbf{1}_2 = (3, -16)^\top$ which is not positive. Hence we consider the submatrix of $S = R$ which matches positive entries of x, i.e., $S(J) = [r_{11}] = [\frac{5}{7}]$ corresponding to $J = \{1\}$. In the next step of this recursion, where $m = 1$ and $S = S(J)$, the equation $Sx = \mathbf{1}_1$ is solved by $x = 1/r_{11} = \frac{7}{5} = 1.4$ which is positive. Hence the recursion terminates and the algorithm returns $y = (1.4, 0)^\top$. Applying Corollary 3 we obtain the optimal left Lyapunov vector $\hat{w} = Ry = (1, 1.8)^\top$ of A. Since $\hat{w}^\top A = (-1.4, 0)$ we obtain from Theorem 8 the exponential estimate

$$\|e^{At}\|_1 \le \kappa(\hat{w})\, e^{\max_{j=1,2} \frac{(\hat{w}^\top A)_j}{\hat{w}_j}\, t} = \kappa(\hat{w}) = 1.8, \quad t \ge 0.$$

(The actual transient gain of $\dot{x} = Ax$ is $\gamma(A) = \sup_{t\ge 0} \|e^{At}\|_1 \approx 1.5$). In comparison, any left Perron vector $\tilde{w}$ of A has the condition number $\kappa(\tilde{w}) = 9.41$ and the estimate provided by Theorem 7 is $\|e^{At}\|_1 \le 9.41e^{-1.18t}$.

Figure 1 shows that the line $\{x \in \mathbb{R}^2_+; \|x\|_{1,\hat{w}} = 1\}$ is tangential to the trajectory of $\dot{x} = Ax$ starting at the second unit vector $e^2 = (0, 1)^\top$, i.e., we have $\hat{w}^\top A e^2 = 0$. This is due to the general fact that, if $e^1, \ldots, e^n$ are the standard unit vectors in $\mathbb{R}^n$ and $\hat{w}$ is the optimal solution of (24) with A replaced by $A^\top$, then $\langle \hat{w}, Ae^h \rangle = (\hat{w}^\top A)_h = 0$ for all $h \in H$ where H is defined by (25) (with A replaced by $A^\top$).

5. Stability Radii and Transient Behaviour of Positive Systems with Structured Uncertainty

In Sec. 3 we have determined stability radii and joint quadratic Lyapunov functions of positive systems with full block uncertainty, i.e., subject to perturbations of the form (11). In this section we will consider highly structured perturbations. Without the assumption of positivity this is a difficult topic as can be seen from the problems of determining structured singular values in μ-analysis, see [7, §4.4].

We will first analyze the effect of entrywise perturbations of the nominal system matrix A at an arbitrary but fixed set of positions. Such perturbations have been analyzed in the context of Linear Algebra by Brualdi [10] in order to obtain spectral inclusion theorems which yields sharper estimates than Gershgorin's classical theorem. More precisely, Brualdi considered an arbitrary $n \times n$ matrix A as an off-diagonal perturbation of the diagonal matrix with the same diagonal as A and used the zero pattern of the matrix A to obtain tighter inclusion regions for its spectrum. In the next subsection we will analyze the stability radius and the transient behaviour of positive systems perturbed by such *Gershgorin–Brualdi* perturbations. In Subsec. 5.2 we will study the problem of constructing joint linear or quadratic Lyapunov functions for more general *sets* of time-invariant linear systems and use the results to obtain stability results and estimates of the transient behaviour for various types of uncertain positive systems. Finally, in Subsec. 5.3 we will study differential inclusions and determine the stability radius of positive systems with respect to *time-varying structured* perturbations.

5.1. Gershgorin-type Perturbation Classes

Throughout this subsection we suppose that $P \in \mathbb{R}_+^{n\times n}$ is a given nonnegative matrix. We consider the following set of complex perturbation matrices which preserve the zero positions in P,

$$\boldsymbol{\Delta} = \boldsymbol{\Delta}_P = \left\{\Delta \in \mathbb{C}^{n\times n}\,;\, \Delta_{ij} = 0 \text{ if } p_{ij} = 0\right\}, \tag{26}$$

and endow the vector space $\boldsymbol{\Delta}_P$ with the weighted ∞-norm

$$\|\Delta\|_P := \max_{(i,j)\in\mathcal{I}(P)} p_{ij}^{-1}|\Delta_{ij}|,\ \ \Delta \in \boldsymbol{\Delta}_P \quad \text{where } \mathcal{I}(P) = \left\{(i,j) \in \underline{n}^2\,;\, p_{ij} > 0\right\} \tag{27}$$

which is a monotone norm on $\boldsymbol{\Delta}_P$. Such *Gershgorin-Brualdi* perturbations have been studied before in [12] and [8]. Note that

$$\Delta \in \boldsymbol{\Delta}_P \quad \Rightarrow \quad M(\Delta) \in \boldsymbol{\Delta}_P.$$

The norm $\|\cdot\|_P$ has the following properties

$$\begin{aligned} &\|M(\Delta)\|_P \leqslant \|\Delta\|_P = \||\Delta|\|_P, && \Delta \in \boldsymbol{\Delta}_P,\\ &\|M(\Delta)\|_P = \|\Delta\|_P && \text{if the diagonal of } \Delta \in \boldsymbol{\Delta}_P \text{ is real,} \end{aligned} \tag{28}$$

$$\|\rho P\|_P = \rho,\ \ \rho > 0, \quad \text{and} \quad \forall \Delta \in \boldsymbol{\Delta}_P :\ \boldsymbol{\Delta}_P \leqslant \rho \Longleftrightarrow |\Delta| \leqslant \rho P.$$

We consider additive perturbations of the form

$$A \rightsquigarrow A + \Delta, \quad \Delta \in \boldsymbol{\Delta}.$$

The *stability radius* of a matrix $A \in \mathbb{C}^{n\times n}$ with respect to the perturbation structure $(\boldsymbol{\Delta}_P, \|\cdot\|_P)$ is defined by

$$r_{\boldsymbol{\Delta}_P}(A) = \inf\{\|\Delta\|_P\,;\, \Delta \in \boldsymbol{\Delta}_P, \alpha(A+\Delta) \geq 0\} \tag{29}$$

(compare (12)). If there exists a destabilizing perturbation $\Delta \in \boldsymbol{\Delta}_P$ such that $\alpha(A+\Delta) \geq 0$ then the inf in (29) may be replaced by a min, i.e., there exists a minimum norm destabilizing perturbation $\Delta_0 \in \boldsymbol{\Delta}_P$ satisfying $\|\Delta_0\|_P = r_{\boldsymbol{\Delta}_P}(A)$ and $\alpha(A+\Delta_0) = 0$. In [12, Coroll. 44], spectral value sets and stability radii with respect to the perturbation structure $(\boldsymbol{\Delta}_P, \|\cdot\|_P)$ have been determined for *diagonal* matrices A. In [8], the problem of constructing joint quadratic Lyapunov functions for systems of the form $\dot{x} = (A+\Delta)x,\ \Delta \in \boldsymbol{\Delta}_P, \|\Delta\| \leqslant \rho$ (where $\|\cdot\|$ is a different norm) has been studied, again for *diagonal* nominal matrices A. For general matrices A, no precise formulae are available, only upper or lower bounds. In the following we will see that for Metzler matrices A explicit formulae can be derived.

Theorem 10. *The stability radius of a Hurwitz stable Metzler matrix* $A \in \mathbb{R}_{\mathrm{M}}^{n\times n}$ *with respect to the perturbation structure* $(\mathbf{\Delta}_P, \|\cdot\|_P)$ *is given by*

$$r_{\mathbf{\Delta}_P}(A) = 1/\varrho(-PA^{-1}). \tag{30}$$

Proof. Suppose that $r_{\mathbf{\Delta}_P}(A) < \infty$ and let $\Delta \in \mathbf{\Delta}_P$, $\delta = \|\Delta\|_P$ be any destabilizing perturbation such that $\alpha(A+\Delta) \geq 0$ (whence $\delta > 0$). As $M(A+\Delta) \leqslant A+|\Delta|$ we have by (8) and (28)

$$0 \leqslant \alpha(A+\Delta) \leqslant \alpha(M(A+\Delta)) \leqslant \alpha(A+|\Delta|) \leqslant \alpha(A+\delta P).$$

Since $A+\delta P$ is a Metzler matrix, there exists a nonzero vector $v \geq 0$ such that $(A+\delta P)v = \alpha v \geq 0$ where $\alpha := \alpha(A+\delta P)$. Multiplying by $-A^{-1} \geq 0$ we obtain $-(I_n + \delta A^{-1}P)v \geq 0$ and therefore $\delta^{-1}v \leqslant -A^{-1}Pv$. Since $-A^{-1}P$ is a nonnegative matrix, Theorem 1 (iii) implies that $\delta^{-1} \leqslant \alpha(-A^{-1}P) = \varrho(-A^{-1}P)$, i.e., $\varrho(-A^{-1}P) > 0$ and $\delta \geq \varrho(-A^{-1}P)^{-1}$. By definition of $r_{\mathbf{\Delta}_P}(A)$ we conclude that $r_{\mathbf{\Delta}_P}(A) \geq \varrho(-A^{-1}P)^{-1} = \varrho(-PA^{-1})^{-1}$. Moreover, the above argument shows that there does not exist a $\Delta \in \mathbf{\Delta}_P$ such that $\alpha(A+\Delta) \geq 0$ if $\varrho(-A^{-1}P) = 0$. Hence $r_{\mathbf{\Delta}_P}(A) = \infty$ and (30) holds if $\varrho(-A^{-1}P) = 0$.

It remains to prove $r_{\mathbf{\Delta}_P}(A) \leqslant \varrho(-A^{-1}P)^{-1}$ if $\varrho(-A^{-1}P) > 0$. Let $\delta_0 = \varrho(-PA^{-1})^{-1}$, $\Delta_0 = \delta_0 P \in \mathbf{\Delta}_P$ and let w be a Perron vector of the nonnegative matrix $-PA^{-1}$. Then $z = -A^{-1}w \geq 0$ satisfies

$$(A+\Delta_0)z = -(A+\delta_0 P)A^{-1}w = -(w - \delta_0\varrho(-PA^{-1})w = 0.$$

Hence $\alpha(A+\Delta_0) \geq 0$ and so $\varrho(-PA^{-1})^{-1} = \|\Delta_0\|_P \geq r_{\mathbf{\Delta}_P}(A)$. ■

Corollary 4. *Suppose* $A, B \in \mathbb{R}_{\mathrm{M}}^{n\times n}$ *are Metzler matrices,* $A \leqslant B$*, and* B *is Hurwitz stable. Then* A *is Hurwitz stable and*

$$r_{\mathbf{\Delta}_P}(B) \leqslant r_{\mathbf{\Delta}_P}(A) \leqslant r_{\mathbf{\Delta}_P}(D(A)) \quad \textit{where} \quad D(A) = \textit{diag}(a_{11}, \dots, a_{nn}). \tag{31}$$

Proof. Since $D(A) \leqslant A \leqslant B$ we have $\alpha(D(A)) \leqslant \alpha(A) \leqslant \alpha(B)$ by (8) and it suffices to prove the first inequality in (31). From the previous proof we know that $\alpha(A + r_{\mathbf{\Delta}_P}(A)P) = 0$, hence $\alpha(B + r_{\mathbf{\Delta}_P}(A)P) \geq 0$. Therefore $r_{\mathbf{\Delta}_P}(A) = \|r_{\mathbf{\Delta}_P}(A)P\| \geq r_{\mathbf{\Delta}_P}(B)$ by (28). ■

Example 2. Let A and P be given by

$$A = \begin{bmatrix} -1 & 4 & 2 \\ 0 & -2 & 1 \\ 0 & 0 & -3 \end{bmatrix}, \qquad P = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Applying (30) to compute the stability radii of A and $D(A)$ with respect to the perturbation structure $(\mathbf{\Delta}_P, \|\cdot\|_P)$ we obtain

$$r_{\mathbf{\Delta}_P}(A) = 0.396 \quad \text{and} \quad r_{\mathbf{\Delta}_P}(D(A)) = 1.195.$$

The next theorem shows that all perturbed matrices $A+\Delta$, $\Delta \in \mathbf{\Delta}_P$, $\|\Delta\|_P \leqslant r_{\mathbf{\Delta}_P(A)}$ satisfy a common transient bound.

Theorem 11. *Suppose that $A \in \mathbb{R}^{n\times n}_{\mathrm{M}}$ is a Hurwitz stable Metzler matrix, $0 < \delta \leqslant r_{\mathbf{\Delta}_P}(A)$ and $w > 0$ is a left Lyapunov vector of $A + \delta P$. Then*

$$\|e^{(A+\Delta)t}\|_1 \leq \kappa(w) e^{\max_j \frac{(w^\top A)_j}{w_j} t}, \quad t \geq 0, \quad \Delta \in \mathbf{\Delta}_P,\ \|\Delta\|_P \leqslant \delta.$$

Proof. Let $\Delta \in \mathbf{\Delta}_P$, $\|\Delta\|_P \leqslant \delta$. Then $|\Delta| \leqslant \delta P$ and we obtain from Theorem 5

$$|e^{(A+\Delta)t}| \leq e^{M(A+\Delta)t} \leq e^{(A+|\Delta|)t} \leq e^{(A+\delta P)t}.$$

Now the operator norm $\|\cdot\|_1$ is monotone on $\mathbb{R}^{n\times n}$, hence Theorem 8 implies that

$$\|e^{(A+\Delta)t}\|_1 \leq \|e^{(A+\delta P)t}\|_1 \leq \kappa(w) e^{\max_j \frac{(w^\top A)_j}{w_j} t}, \quad t \geq 0.$$

This proves the theorem. ∎

The previous proof is based on the fact that any (left or right) Lyapunov vector of $A + \delta P$ is a joint Lyapunov vector for all perturbed Metzler matrices $M(A + \Delta)$ where $\Delta \in \mathbf{\Delta}_P$ and $\|\Delta\|_P \leqslant \delta$. In the next subsection we will study the problem under which conditions there exist joint Lyapunov vectors for a given set of Metzler matrices.

5.2. Joint Lyapunov Vectors for Polytopes of Positive Systems

In this subsection we introduce the concept of *linear stability* and derive necessary and sufficient linear stability criteria for polytopes of positive systems.

Over the last two decades the concept of *quadratic stability* has attracted some interest in the literature because it provides a useful tool for dealing with time-varying and/or nonlinear perturbations of uncertain systems. A set of matrices $\mathcal{A} \subset \mathbb{C}^{n\times n}$ is said to be *quadratically stable* if there exists a positively definite Hermitian matrix $P \in \mathbb{C}^{n\times n}$ such that $PA + A^*P$ is negatively definite for all $A \in \mathcal{A}$. In this case, P provides a joint strict quadratic Lyapunov function for all the systems $\dot{x} = Ax$, $A \in \mathcal{A}$. For sets of positive systems quadratic Lyapunov functions may be replaced by linear ones. This leads to the following definition of *linear stability*.

Definition 6. *A set $\mathcal{A} \subset \mathbb{R}^{n\times n}_{\mathrm{M}}$ of Metzler matrices is said to be linearly stable if there exists a joint strict left Lyapunov vector w for $\mathcal{A}$, i.e., a strictly positive vector w satisfying $w^\top A < 0$ for all $A \in \mathcal{A}$.*

Remark 3. Every joint (strict) Lyapunov vector for $\mathcal{A}$ is automatically a joint (strict) Lyapunov vector for the convex hull of $\mathcal{A}$, denoted by $\mathrm{conv}(\mathcal{A})$. Hence by Theorem 8 a necessary condition for the linear stability of $\mathcal{A}$ is that every matrix $A \in \mathrm{conv}(\mathcal{A})$ be Hurwitz stable.

Since every strict left Lyapunov vector of a Metzler matrix A is a strict left Lyapunov vector of every Metzler matrix $B \leqslant A$, linear stability extends automatically from any set $\mathcal{A} \subset \mathbb{R}^{n\times n}_{\mathrm{M}}$ to its *lower hull* $\mathcal{A}_{\leqslant} = \{M \in \mathbb{R}^{n\times n}_{\mathrm{M}}\,;\, \exists A \in \mathcal{A} : M \leqslant A\}$.

We conclude from Remark 3 that for sets of Metzler matrices, linear stability is a stronger property than exponential stability. A single Metzler matrix is exponentially stable if and only if it has a strict Lyapunov vector. However, this is no longer true for *sets of Metzler matrices*, as is illustrated by the next example.

Example 3. Consider the Metzler matrices

$$A_1 = \begin{bmatrix} -1 & 10 \\ 0 & -1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -1 & 0 \\ 10 & -1 \end{bmatrix}.$$

They are both Hurwitz stable but the Metzler matrix $(A_1+A_2)/2 \in \mathrm{conv}\{A_1, A_2\}$ is unstable. Hence the set $\mathcal{A} = \{A_1, A_2\}$ is not linearly stable.

It is well-known that even if $\mathrm{conv}(\mathcal{A})$ consists only of Hurwitz stable matrices, $\mathcal{A}$ need not be quadratically stable. The following example shows that an analogous statement holds true for linear stability. Moreover it illustrates that the quadratic stability of a set $\mathcal{A}$ of positive systems does not imply the linear stability of $\mathcal{A}$.

Example 4. There are pairs of Metzler matrices which do not have a common linear Lyapunov function, but a quadratic one. Consider the Metzler matrices

$$A_1 = \begin{bmatrix} -10 & 5 \\ 5 & -3 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} -10 & 2 \\ 8 & -3 \end{bmatrix}.$$

By an easy calculation one verifies that $P = \mathrm{diag}(5, 3)$ is a positive definite matrix with $PA_i + A_i^{\mathsf{T}} P \prec 0$ for $i = 1, 2$. Hence the segment of matrices

$$\mathcal{A} = [A_1, A_2] = \{\tau A_1 + (1-\tau)A_2\,;\, \tau \in [0, 1]\}$$

is quadratically stable. On the other hand, we will see later in Example 5 that there does not exist a joint strict Lyapunov vector for the two matrices A_1, A_2, and consequently the segment $[A_1, A_2]$ is not linearly stable.

We will now derive necessary and sufficient linear stability criteria for polytopes of positive systems. We make use of the following lemma which is a consequence of [18, Thm 22.2].

Lemma 7. *Suppose $A \in \mathbb{R}^{n\times N}$ where $n, N \geq 1$. Then one and only one of the following alternatives holds:*
(a) *There exists a vector $w \in \mathbb{R}^n_+, w > 0$ such that $w^\top A < 0$.*
(b) *There exists a vector $x \in \mathbb{R}^N_+$, $x \neq 0$ such that $Ax \geq 0$.*

Theorem 12. *If $A_1, \dots, A_m \in \mathbb{R}^{n\times n}_{\mathrm{M}}$, the following conditions are equivalent.*

(i) *The polytope $\mathcal{A} = conv\{A_1, \dots, A_m\}$ is linearly stable.*
(ii) *There exists $w > 0$ such that $w^\top A_j < 0$ for $j = 1, \dots, m$.*
(iii) *A_1 is Hurwitz stable and there exists $y > 0$ such that $y^\top A_1^{-1} A_j > 0$ for $j = 2, \dots, m$.*
(iv) *There do not exist $x^j \in \mathbb{R}^n_+$, $j \in \underline{m}$ not all zero such that $\sum_j A_j x^j \geq 0$.*

Proof. (i) ⇔ (ii) follows directly from Definition 6 and Remark 3.
(ii) ⇔ (iii): If $w > 0$ is a joint strict left Lyapunov vector for $\{A_1, \dots, A_m\}$ then A_1 is Hurwitz stable and

$$w^\top [A_1 \dots A_m] = w^\top A_1 [I \; A_1^{-1} A_2 \dots A_1^{-1} A_m] < 0.$$

Hence, setting $y = -A_1^\top w$ we obtain $y > 0$ and $y^\top A_1^{-1} A_j > 0$ for $j = 2, \dots, m$. Conversely, if A_1 is Hurwitz stable and $y > 0$ satisfies $y^\top A_1^{-1} A_j > 0$ for $j = 2, \dots, m$ then $w = -(A_1^{-1})^\top y$ defines a joint strict left Lyapunov vector for $\{A_1, A_2, \dots, A_m\}$.
(ii) ⇔ (iv) follows directly from Lemma 7 by setting $A = [A_1 \, A_2 \, \dots \, A_m] \in \mathbb{R}^{n\times(mn)}$. ■

It follows from this characterization that if A_1 is either not Hurwitz stable or one of the matrices $A_1^{-1} A_j$, $j = 2, \dots, m$ has a column consisting only of non-positive entries, then the set $\{A_1, \dots, A_m\}$ is not linearly stable.

Example 5. Consider again the two matrices A_1, A_2 defined in Example 4. Then A_1 is Hurwitz stable and $A_1^{-1} A_2 = \begin{bmatrix} -2 & 1.8 \\ -6 & 4 \end{bmatrix}$ has a column of strictly negative entries. Thus there does not exist a joint strict left Lyapunov vector for A_1, A_2 and so $[A_1, A_2]$ is not linearly stable although it is quadratically stable as was shown in Example 4.

The above results deal with joint left Lyapunov vectors. Analogous statements can be derived for joint right Lyapunov vectors applying the previous results to sets of transposed matrices. In this way it can be shown that there does not exist a joint strict right Lyapunov vector for the two matrices A_1, A_2 in the previous example because $A_2 A_1^{-1} = \begin{bmatrix} 4 & 6 \\ -1.8 & -2 \end{bmatrix}$ has a strictly negative row.

The question arises whether or not – as in the single matrix case – a set of Metzler matrices has a joint left Lyapunov vector if and only if it has a joint right Lyapunov vector. The next example shows that this is not so.

Example 6. Consider the two Metzler matrices

$$A_1 = \begin{bmatrix} -5 & 39 \\ 0 & -3 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -1 & 7 \\ 3 & -25 \end{bmatrix}.$$

These two matrices have the joint right Lyapunov vector $(8, 1)^\top$, but no joint left Lyapunov vector, as $A_1^{-1} A_2$ has a column with strictly negative entries, see

Theorem 12. Thus we have here an example of a set of two matrices $\{A_1, A_2\}$ which is not linearly stable whereas the set of transposes $\{A_1^\top, A_2^\top\}$ is linearly stable. A similar example cannot be found for quadratic stability. If a matrix set $\mathcal{A} \subset \mathbb{C}^{n\times n}$ is quadratically stable then the set $\mathcal{A}^\top = \{A^\top; A \in \mathcal{A}\}$ is also quadratically stable: If there exists $P \succ 0$ such that $PA + A^\top P \prec 0$ for all $A \in \mathcal{A}$, then $P^{-1} \succ 0$ and $P^{-1}A^\top + AP^{-1} \prec 0$ holds for all $A \in \mathcal{A}$.

We have seen above that there exists a set of two matrices $\{A_1, A_2\}$ which is quadratically stable but both $\{A_1, A_2\}$ and $\{A_1^\top, A_2^\top\}$ are not linearly stable. The question arises if the converse is also possible, i.e., that $\{A_1, A_2\}$ and $\{A_1^\top, A_2^\top\}$ are both linearly stable but $\{A_1, A_2\}$ is not quadratically stable. The answer is "no", since one can construct joint *diagonal* quadratic Lyapunov functions from a pair of joint strict left (resp. right) Lyapunov vectors.

Corollary 5. *Suppose that $A_1, A_2 \in \mathbb{R}^{n\times n}_{\mathrm{M}}$ are Metzler matrices. If $w > 0$ (resp. $v > 0$) are joint strict left (resp. right) Lyapunov vectors for A_1, A_2 then the diagonal matrix $P = diag(\frac{w_i}{v_i}) \succ 0$ defines a joint quadratic Lyapunov function $V(x) = \langle x, Px\rangle$ for $\{A_1, A_2\}$.*

Proof. We show that $R_i := PA_i + A_i^\top P$, $i = 1, 2$ are negative definite. By construction, we have $A_i v < 0$ and $A_i^\top w < 0$ for $i = 1, 2$. Therefore $R_i v = PA_i v + A_i^\top P v = PA_i v + A_i^\top w < 0$ for $i = 1, 2$, i.e. v is a strict joint right Lyapunov vector for the Metzler matrices R_i, $i = 1, 2$. From Theorem 8 we conclude that the R_i are Hurwitz stable, hence by symmetry they are negative definite. Therefore $V(x) = \langle x, Px\rangle$ is a joint quadratic Lyapunov function for $\{A_1, A_2\}$. ∎

5.3. Differential Inclusions of Positive Systems

A useful class of models for linear systems with time-varying parameter uncertainties is given by *linear differential inclusions* (LDI) of the form

$$\dot{x}(t) \in \mathcal{A}x(t), \ \ t \geq 0, \quad \mathcal{A} \subset \mathbb{C}^{n\times n} \text{ compact}, \tag{32}$$

where $\mathcal{A}x = \{Ax; A \in \mathcal{A}\}$ for $x \in \mathbb{R}^n$. An absolutely continuous function $x(\cdot) : \mathbb{R}_+ \to \mathbb{R}^n$ is said to be a *solution* of (32) if $\dot{x}(t) \in \mathcal{A}x(t)$ holds for almost all $t \in \mathbb{R}_+$. By a theorem of Filippov, $x(\cdot) : \mathbb{R}_+ \to \mathbb{R}^n$ is a solution of (32) if and only if there exists a measurable selection $A(\cdot) : \mathbb{R}_+ \to \mathcal{A}$ such that $x(\cdot)$ is solution of the time-varying linear differential equation $\dot{x}(t) = A(t)x(t)$ on $\mathbb{R}_+$, see [19, Thm. 2.3].

Definition 7. *The LDI* (32) *is said to be stable if, given any $\varepsilon > 0$, there exists $\delta > 0$ such that for every $x^0 \in \mathbb{R}^n$, $\|x^0\|_2 < \delta$ each solution $x(\cdot)$ of* (32) *with $x(0) = x^0$ satisfies $\|x(t)\|_2 < \varepsilon$ for all $t \geq 0$. It is said to be asymptotically stable (in the large) if additionally the origin is globally attractive, i.e., $\lim_{t\to\infty} x(t) = 0$ for all solutions $x(\cdot)$ of* (32).

It can be shown [10, §8.2] that (32) is asymptotically stable if and only if it is *exponentially stable* in the following sense: There exist constants $\beta < 0$, $\gamma \geq 1$

such that

$$\|x(t)\|_2 \leqslant \gamma e^{\beta t}, \quad t \geq 0 \tag{33}$$

for every solution $x(\cdot)$ of (32). By the following well known lemma, a necessary condition for the asymptotic stability of (32) is that the convex hull $\mathrm{conv}(\mathcal{A})$ consists of Hurwitz stable matrices, see [15].

Lemma 8. *The linear differential inclusion* (32) *is asymptotically stable if and only if its convexification* $\dot{x} \in \mathrm{conv}(\mathcal{A})$ *is asymptotically stable.*

The following theorem is apparently due to Molchanov and Pyatnitskij, see [15]. A detailed proof can be found in [19]. The theorem exhibits the close relationship between the asymptotic stability of linear differential inclusions and the existence of joint Lyapunov functions.

Theorem 13. *The linear differential inclusion* (32) *is asymptotically stable if and only if there exists a number* $\theta > 0$ *and a piecewise linear Lyapunov function of the form*

$$V(x) = \max\{|\langle l_i, x\rangle|, i \in \underline{m}\}, \quad x \in \mathbb{R}^n,$$

where $l_1, \dots, l_m \in \mathbb{R}^n$ *span* $\mathbb{R}^n$*, such that*

$$DV(x)(v) \leqslant -\theta V(x), \quad x \in \mathbb{R}^n,\ v \in \mathcal{A}x. \tag{34}$$

Here $DV(x)(v)$ denotes the directional derivative of V at x in the direction of v,

$$DV(x)(v) = \lim_{h \to 0+} h^{-1}\left[V(x+hv) - V(x)\right].$$

The limit exists since $V : \mathbb{R}^n \to \mathbb{R}_+$ is convex. By (34), V is a joint strict Lyapunov norm (in the sense of Definition 3) for all the systems $\dot{x} = Ax$, $A \in \mathcal{A}$.

Theorem 13 shows a fundamental difference between robustness properties with respect to time-varying and time-invariant perturbations. In general, the asymptotic stability of a set of time-invariant systems $\dot{x} = Ax$, $A \in \mathcal{A}$, does not guarantee the existence of a joint strict Lyapunov function for all these systems. In contrast, if all the *time-varying* systems $\dot{x}(t) = A(t)x(t)$, $A(\cdot) : \mathbb{R}_+ \to \mathcal{A}$ being measurable, are uniformly asymptotically stable then, by the above theorem, such a joint Lyapunov function always exists.

The above results only serve as a background for the following development. We will not use them directly but give independent proofs. With every LDI (32) we associate the positive LDI

$$\dot{x}(t) \in M(\mathcal{A})x(t), \ \text{ where } M(\mathcal{A}) = \{M(A); A \in \mathcal{A}\} \subset \mathbb{R}_{\mathrm{M}}^{n\times n}. \tag{35}$$

Note that $M(\mathcal{A})$ is compact since $A \mapsto M(A)$ defined by (3) is continuous and $\mathcal{A}$ is compact.

Theorem 14. *Suppose* $\mathcal{A} \subset \mathbb{C}^{n\times n}$ *is compact and* $w > 0$ *is a vector satisfying* $w^\top M(A) \leqslant 0$ (*resp.* $w^\top M(A) < 0$) *for all* $A \in \mathcal{A}$*. Then* (32) *is stable* (*resp.*

asymptotically stable) and every solution $x(\cdot)$ of (32) *on $\mathbb{R}_+$ satisfies*

$$\|x(t)\|_1 \leqslant \kappa(w)e^{\beta t}\,\|x(0)\|_1,\ t \geq 0 \ \ \textit{where}\ \beta = \max_{A\in\mathcal{A}}\max_j \frac{(w^\top M(A))_j}{w_j} \leqslant 0. \quad (36)$$

In particular, if the positive LDI (35) *is linearly stable then the LDI* (32) *is asymptotically stable.*

Proof. Suppose that $w > 0$ and $w^\top M(A) \leqslant 0$ for all $A \in \mathcal{A}$. Let $x(\cdot) : \mathbb{R}_+ \to \mathbb{C}^n$ be any solution of (32) and let $A(\cdot) : \mathbb{R}_+ \to \mathcal{A}$ be a measurable selection such that $x(\cdot)$ solves $\dot{x} = A(t)x$. Let $\nu(\cdot)$ be an arbitrary norm on $\mathbb{C}^n$ and $\mu_\nu(\cdot)$ the corresponding initial growth rate. Then Theorem II. 8.27 of [4] shows that

$$\nu\,(x(t)) \leq e^{\int_0^t \mu_\nu(A(\tau))\,d\tau}\nu(x(0)) \leq e^{\int_0^t \mu_\nu(M(A(\tau)))\,d\tau}\nu(x(0))$$

(where the second inequality is obtained by (21)). Now we choose the norm $\nu(x) = \|x\|_{1,w}$ which equals $w^\top x$ on $\mathbb{R}^n_+$. By Theorem 8, $\mu_\nu(M(A(\tau))) = \max_j \frac{(w^\top M(A(\tau)))_j}{w_j} \leqslant \beta \leqslant 0$ for all $\tau \geq 0$ and so all the solutions of (32) satisfy

$$(\min_i w_i)\|x(t)\|_1 \leqslant \|x(t)\|_{1,w} \leqslant e^{\int_0^t \beta\,d\tau}\,\|x(0)\|_{1,w} \leqslant e^{\beta t}(\max_i w_i)\|x(0)\|_1,\ \ t \geq 0.$$

This proves (36) and that (32) is stable if $w^\top M(A) \leqslant 0$ for all $A \in \mathcal{A}$. Finally, compactness arguments show that $\beta < 0$ if $w^\top M(A) < 0$ for all $A \in \mathcal{A}$. ∎

Corollary 6. *Suppose that $A, B \in \mathbb{R}^{n\times n}_M$ are Metzler matrices, $B \leqslant A$ and $\alpha(A) < 0$. Then the interval LDI $\dot{x}(t) \in \mathcal{A}x(t)$ where $\mathcal{A} = [[B, A]]$ is asymptotically stable.*

Proof. By Theorem 8 there exists $w > 0$ such that $w^\top A < 0$. Since the interval $[[B, A]]$ is a compact set of Metzler matrices and $w^\top C \leqslant w^\top A < 0$ for all $C \in [[B, A]]$, the corollary follows from Theorem 14. ∎

The following corollary is a direct consequence of Theorem 4 and Lemma 8.

Corollary 7. *Suppose that $A_1, \ldots, A_m \in \mathbb{R}^{n\times n}_M$ are Metzler matrices, $w > 0$ satisfies $w^\top A_j < 0$ (resp. $w^\top A_j \leqslant 0$) for all $j \in \underline{m}$. Then the LDI*

$$\dot{x}(t) \in \mathcal{A}x(t) \quad \text{where} \quad \mathcal{A} = \text{conv}\{A_1, \ldots, A_m\} \quad (37)$$

is asymptotically stable (resp. stable) and for every solution $x(\cdot)$ of (37),

$$\|x(t)\|_1 \leqslant \kappa(w)e^{\beta t}\,\|x(0)\|_1,\ \ t \geq 0 \quad \text{where}\ \ \beta = \max_{i\in\underline{m}}\max_j \frac{(w^\top A_i)_j}{w_j}. \quad (38)$$

The condition number $\kappa(w)$ in (38) can be minimized in a similar manner as for a single matrix, see §4.3. Let $A = [A_1\ A_2\ \cdots\ A_m]$ and

$$\mathbf{W}_1 = \{w \in \mathbb{R}^n_+\,;\ \min_i w_i = 1, w^\top A \leqslant 0\}.$$

As in the single matrix case one can prove

$$v, w \in \mathbf{W}_1 \quad \Rightarrow \quad \min\{v, w\} \in \mathbf{W}_1.$$

Now consider the optimization problem

$$\text{Minimize } \kappa(w) = \frac{\max_i w_i}{\min_i w_i} \quad \text{subject to } \; w > 0,\; w^\top A \leqslant 0 \tag{39}$$

and the associated linear program

$$\text{Minimize } \mathbf{1}_n^\top w = \sum_{i=1}^{n} w_i \quad \text{subject to } \; w \geq \mathbf{1}_n \quad \text{and} \quad w^\top A \leq 0. \tag{40}$$

Then the following result can be proved in a similar way as in the single matrix case, compare the proof of Theorem 9.

Theorem 15. *Suppose that $A_1, \ldots, A_m \in \mathbb{R}_M^{n\times n}$ are Metzler matrices and $w > 0$ satisfies $w^\top A_j \leqslant 0$ for $j \in \underline{m}$. Then the linear program* (40) *has a unique optimal solution $\hat{w}$. $\hat{w}$ is an optimal solution of* (39) *and $\hat{w} \in \mathbf{W}_1$. Every solution $x(\cdot)$ of the LDI* (37) *satisfies*

$$\|x(t)\|_1 \leqslant \kappa(\hat{w})\, \|x(0)\|_1, \quad t \geq 0.$$

Proof. By assumption, the set of admissible solution of the linear program (40) is non-empty and the objective functional $\mathbf{1}^\top w$ is bounded below on this set. Hence there exists an optimal solution $\hat{w} \in \mathbf{W}_1$ of (40). As in the single matrix case, one can show that $\hat{w}$ is uniquely determined and is an optimal solution of (39). The last statement of the theorem follows from (38) and $\hat{w}^\top A_i \leqslant 0$, $i \in \underline{n}$. ∎

As a consequence of Theorem 15, an optimal solution of (39) and thus a weight vector $w > 0$ of minimal condition number $\kappa(w)$ can be determined by applying the simplex method to the linear program (40).

We conclude the paper by applying the previous results to *time-varying linear* and/or *nonlinear* parameter perturbations of Gershgorin–Brualdi type. Let $A \in \mathbb{R}_M^{n\times n}$ be a Hurwitz stable Metzler matrix and $P = (p_{ij}) \in \mathbb{R}_+^{n\times n}$. We consider nonlinear time-varying perturbations of $\dot{x} = Ax$ of the form

$$\dot{x} = Ax + \Delta(x,t)x \quad \text{where } \; \Delta(\cdot,\cdot) \in \mathbf{\Delta}_{nt}. \tag{41}$$

Here $\mathbf{\Delta}_{nt}$ is the vector space of all *bounded* $\Delta(\cdot,\cdot) : \mathbb{C}^n \times \mathbb{R}_+ \to \mathbb{C}^{n\times n}$ of structure P, i.e., $\Delta(x,t) = (\delta_{ij}(x,t)) \in \mathbb{C}^{n\times n}$, $\delta_{ij}(x,t) = 0$ for all $x \in \mathbb{C}^n$ and $t \geq 0$ if $p_{ij} = 0$, such that $\Delta(x,\cdot) : \mathbb{R}_+ \to \mathbb{C}^{n\times n}$ is measurable for each $x \in \mathbb{C}^n$, $\Delta(\cdot,t) : \mathbb{C}^n \to \mathbb{C}^{n\times n}$ is continuous for each fixed $t \in \mathbb{R}_+$, and for each compact subset $K \times I \subset \mathbb{C}^n \times \mathbb{R}_+$, there exists an integrable $k(\cdot) : I \to \mathbb{R}_+$ such that

$$\|\Delta(x,t)x - \Delta(\hat{x},t)\hat{x}\|_2 \leq k(t)\|x - \hat{x}\|_2, \;\; (x,t), (\hat{x},t) \in K \times I.$$

The norm on $\mathbf{\Delta}_{nt}$ is taken to be

$$\Delta(\cdot,\cdot)\|_{\mathbf{\Delta}_{nt}} = \sup_{x\in\mathbb{C}^n, t\geq 0} \max_{(i,j)\in\mathcal{I}(P)} p_{ij}^{-1}|\delta_{ij}(x,t)|, \;\; \Delta = (\delta_{ij}) \in \mathbf{\Delta}_{nt}, \tag{42}$$

where $\mathcal{I}(P) = \{(i,j) \in \{1,\dots,n\}^2\,;\, p_{ij} > 0\}$.

By Carathéodory's Theorem, there exists, for every $(t_0, x^0) \in \mathbb{R}_+ \times \mathbb{C}^n$, a unique solution $x(t) = x(t; t_0, x^0)$ of (41) on $[t_0, \infty)$ with $x(t_0) = x^0$. We say that the nonlinear system (41) is *asymptotically stable* if the origin is a globally asymptotically stable equilibrium position of (41).

We also consider the following time-varying linear system

$$\dot{x}(t) = (A + \Delta(t))x(t), \quad \text{where } \ \Delta(\cdot) \in \mathbf{\Delta}_{tv}\,. \tag{43}$$

Here $\mathbf{\Delta}_{tv}$ is the vector space of all bounded measurable matrix functions $\Delta(\cdot)\colon \mathbb{R}_+ \to \mathbb{C}^{n\times n}$ of structure P, i.e.,

$$\Delta(t) = (\delta_{ij}(t)) \in \mathbb{C}^{n\times n}, \quad \delta_{ij}(t) = 0 \ \text{ for all } t \geq 0 \ \text{ if } \ p_{ij} = 0\,.$$

$\mathbf{\Delta}_{tv}$ is endowed with the norm

$$\|\Delta(\cdot)\|_{\mathbf{\Delta}_{tv}} = \sup_{t\geq 0} \max_{(i,j)\in\mathcal{I}(P)} p_{ij}^{-1}|\delta_{ij}(t)|, \quad \Delta(\cdot) = (\delta_{ij}(\cdot)) \in \mathbf{\Delta}_{tv}. \tag{44}$$

Note that with the obvious embeddings $\mathbf{\Delta}_P \subset \mathbf{\Delta}_{tv} \subset \mathbf{\Delta}_{nt}$, the norm $\|\cdot\|_{\mathbf{\Delta}_{tv}}$ is the restriction of the norm $\|\cdot\|_{\mathbf{\Delta}_{nt}}$ to $\mathbf{\Delta}_{tv}$ and the norm $\|\cdot\|_{\mathbf{\Delta}} = \|\cdot\|_P$ is the restriction of the norm $\|\cdot\|_{\mathbf{\Delta}_{tv}}$ to $\mathbf{\Delta}_P$. Here $\mathbf{\Delta}_P$ is defined by (26) and the norm $\|\cdot\|_P$ by (27).

For simplicity we say that the nonlinear system (41) is asymptotically stable if the origin $\bar{x} = 0$ is a globally asymptotically stable equilibrium position of (41).

Definition 8. *Given $A \in \mathbb{C}^{n\times n}$, the stability radius of A with respect to complex time-varying linear (resp.) perturbations $\Delta(\cdot) \in \mathbf{\Delta}_{tv}$ (respectively $\Delta(\cdot,\cdot) \in \mathbf{\Delta}_{nt}$) are defined by*

$$r_{\mathbf{\Delta}_{tv}}(A) = \inf\{\|\Delta(\cdot)\|_{\mathbf{\Delta}_{tv}}\,;\, \Delta(\cdot) \in \mathbf{\Delta}_{tv} \text{ and (43) is not asymptotically stable}\},$$
$$r_{\mathbf{\Delta}_{nt}}(A) = \inf\{\|\Delta(\cdot,\cdot)\|_{\mathbf{\Delta}_{nt}}\,;\, \Delta(\cdot,\cdot) \in \mathbf{\Delta}_{nt} \text{ and (41) is not asymptotically stable}\}.$$

As a consequence of the above results we obtain

Corollary 8. *Let $A \in \mathbb{R}^{n\times n}_M$ be a Hurwitz stable Metzler matrix and $P = (p_{ij}) \in \mathbb{R}^{n\times n}_+$. Then*

$$r_{\mathbf{\Delta}_{nt}}(A) = r_{\mathbf{\Delta}_{tv}}(A) = r_{\mathbf{\Delta}_P}(A) = \varrho\left(-PA^{-1}\right)^{-1}. \tag{45}$$

Moreover, if $\delta_0 < r_{\mathbf{\Delta}_P}(A)$ and $x(\cdot)$ is any solution of the LDI

$$\dot{x}(t) \in \mathcal{A}_{\delta_0}x(t), \quad \text{where } \ \mathcal{A}_{\delta_0} = \{A + \Delta\,;\, \Delta \in \mathbf{\Delta}_P, \|\Delta\|_P \leqslant \delta_0\} \tag{46}$$

then

$$|x(t)| \leqslant e^{(A+\delta_0 P)t}|x(0)|, \ \ t \geq 0 \quad (\text{componentwise}) \tag{47}$$

and, if $w > 0$ is any left Lyapunov vector of the Hurwitz stable matrix $A + \delta_0 P$,

$$\|x(t)\|_1 \leqslant \kappa(w)e^{\beta t}\,\|x(0)\|_1, \ \ t \geq 0, \ \ \text{where } \beta = \max_j \frac{(w^\top(A+\delta_0 P))_j}{w_j} \leqslant 0. \tag{48}$$

Proof. We first prove (47). Let $x(\cdot)$ be a solution of the LDI (46). By Filippov's Theorem, there exists a perturbation $\Delta(\cdot) \in \mathbf{\Delta}_{tv}$, $\|\Delta(t)\|_P \leqslant \delta_0$ for all $t \geq 0$ such that $x(\cdot)$ is a solution of $\dot{x}(t) = (A + \Delta(t))x(t)$. Then $|\Delta(t)| \leqslant \delta_0 P$ for all $t \geq 0$ (componentwise) and hence by the variation-of-constants formula, the function $u(\cdot) : \mathbb{R}_+ \to \mathbb{R}^n$ defined by $u(t) = |x(t)|$, $t \geq 0$ satisfies the integral inequality

$$\begin{aligned} u(t) &= \left| e^{At}x(0) + \int_0^t e^{A(t-s)}\Delta(s)x(s)ds \right| \\ &\leqslant e^{At}|x(0)| + \int_0^t e^{A(t-s)}\delta_0 P u(s)ds, \quad t \geq 0. \end{aligned}$$

On the other hand, $v(t) = e^{(A+\delta_0 P)t}|x(0)|$ satisfies the integral equation

$$v(t) = e^{(A+\delta_0 P)t}|x(0)| + \int_0^t e^{A(t-s)}\delta_0 P v(s)ds, \quad t \geq 0.$$

Since the matrix $e^{A(t-s)}\delta_0 P$ is nonnegative, we may apply Theorem 18.3 of [1] with $K(t, s, u) = e^{A(t-s)}\delta_0 P u$, $0 \leqslant s \leqslant t$, $u \in \mathbb{R}^n$, and obtain $u(t) \leqslant v(t)$ for all $t \geq 0$, i.e. (47) holds. (48) now follows from (47) and Theorem 8.

Finally, we prove (45). It follows from the definitions and the isometric embeddings $\mathbf{\Delta}_P \subset \mathbf{\Delta}_{tv} \subset \mathbf{\Delta}_{nt}$ that $r_{\mathbf{\Delta}_{nt}}(A) \leqslant r_{\mathbf{\Delta}_{tv}}(A) \leqslant r_{\mathbf{\Delta}_P}(A)$. Suppose that $\Delta(\cdot,\cdot) \in \mathbf{\Delta}_{nt}$ and $\delta_0 := \|\Delta(\cdot,\cdot)\|_{\mathbf{\Delta}_{nt}} < r_{\mathbf{\Delta}_P}(A)$. Let $x(\cdot)$ be any solution of the time-varying nonlinear system (41). Since $A + \Delta(x(t), t) \in \mathcal{A}_{\delta_0}$ for all $t \geq 0$, we conclude that $x(\cdot)$ is a solution of the LDI (46). Choosing for $w > 0$ a strict left Lyapunov vector of $A + \delta_0 P$ and applying (48) we see that (41) is exponentially stable for every $\Delta(\cdot,\cdot) \in \mathbf{\Delta}_{nt}$ satisfying $\|\Delta(\cdot,\cdot)\|_{\mathbf{\Delta}_{nt}} < r_{\mathbf{\Delta}_P}(A)$. Hence $r_{\mathbf{\Delta}_{nt}}(A) \geq r_{\mathbf{\Delta}_P}(A)$ and this proves the first two equalities in (45). The last equality in (45) follows from (30). ■

We conclude the paper with an example illustrating the above results.

Example 7. Consider the matrices A and P of Example 2, fix the 1-norm on $\mathbb{C}^3$ and let $\delta_0 < r_{\mathbf{\Delta}_P}(A) = 0.396$. By Corollary 8, the perturbed linear systems $\dot{x}(t) = (A+\Delta(t))x(t)$ are asymptotically stable if the time-varying perturbations $\Delta(\cdot) : \mathbb{R}_+ \to \mathbf{\Delta}_P$ satisfy $\|\Delta(\cdot)\|_{\mathbf{\Delta}_{tv}} \leqslant \delta_0$. For any such system, let $\Phi_\Delta(t, s)$ denote the associated evolution operator. To find an upper bound for the transient gains $\sup_{t\geq 0} \|\Phi_\Delta(t, 0)\|_1$ of all these perturbed systems, we apply Algorithm 1 to the matrix $A + \delta_0 P$ and obtain an optimal left Lyapunov vector $\hat{w} = (1, 2.52, 1.64)^{\mathbb{T}}$ of $A + \delta_0 P$ with minimal condition number $\kappa(\hat{w}) = 2.52$. By Corollary 8, the condition number $\kappa(\hat{w})$ is an upper bound for the transient gains of all the systems $\dot{x}(t) = (A + \Delta(t))x(t)$ where $\Delta(\cdot) : \mathbb{R}_+ \to \mathbf{\Delta}_P$ satisfies $\|\Delta(\cdot)\|_{\mathbf{\Delta}_{tv}} \leqslant \delta_0$. This is illustrated in Figure 2 for two perturbations of norm $\delta_0 = 0.395$, one time-invariant and one time-varying,

$$\Delta_1(t) = \delta_0 P = 0.395 \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad \Delta_2(t) = 0.395 \begin{bmatrix} 0 & \cos(3t) & \sin(t)^2 \\ 1 & 0 & 0 \\ 0 & \cos(2t)^2 & 0 \end{bmatrix}$$

with associated evolution operators $\Phi_1(t,s) = e^{(A+\delta_0 P)(t-s)}$ and $\Phi_2(t,s) = \Phi_{\Delta_2}(t,s)$. Since $|\Delta_2(t)| \leqslant \delta_0 P$ (componentwise) for $t \geq 0$ it follows from Corollary 8 that $|\Phi_2(t,0)| \leqslant e^{(A+\delta_0 P)t}$ and hence $\|\Phi_2(t,0)\|_1 \leqslant \|e^{(A+\delta_0 P)t}\|_1$ for all $t \geq 0$, see Figure 2.

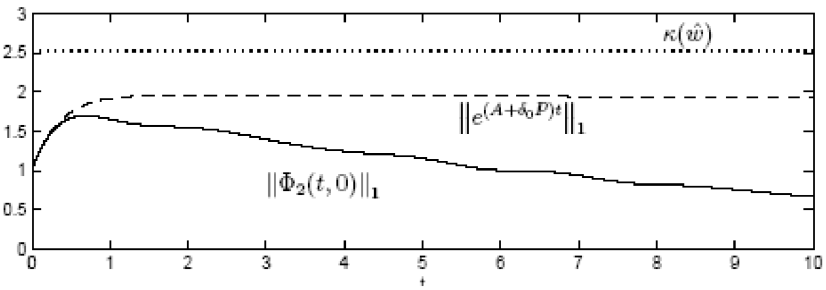

Fig. 2. Norms of the evolution operator/matrix exponential. and upper bound $\kappa(\hat{w})$

References

1. D. Bainov and P. Simeonov, *Integral Inequalities and Applications*, Kluwer Academic Publishers, Dordrecht, 1992.
2. A. Berman and R. J. Plemmons, *Nonnegative Matrices in Mathematical Sciences*, Academic Press, New York, 1979.
3. S. P. Bhattacharya and L. Keel, editors, *Robust Control*, CRC Press, Boca Raton, 1991.
4. C. A. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties*, Academic Press, New York, 1975.
5. N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, SIAM Publications, Philadelphia, PA, 2nd edition, 2002.
6. D. Hinrichsen and A. J. Pritchard, Stability radius for structured perturbations and the algebraic Riccati equation, *Syst. Control Lett.* **8** (1986)105–113.
7. D. Hinrichsen and A. J. Pritchard, *Mathematical Systems Theory I. Modelling, State Space Analysis, Stability and Robustness*, Springer-Verlag, Berlin, 2005.
8. D. Hinrichsen and A. J. Pritchard, Gershgorin–Brualdi perturbations and Riccati equations, *Lin. Alg. Appl.* **425** (2007) 374–403.
9. D. Hinrichsen and N.K. Son, μ-analysis and robust stability of positive linear systems, *Appl. Math. and Comp. Sci.* **8** (1998) 253–268.
10. R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1990.
11. R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
12. M. Karow, D.Hinrichsen, and A. J. Pritchard, Interconnected systems with uncertain couplings: explicit formulae for μ-values, spectral value sets and stability radii, *SIAM J. Control Optim.* **45** (2006) 856–884.

13. V. L. Kharitonov, Asymptotic stability of an equilibrium position of a family of systems of linear differential equation, *J. Diff. Eqns.* **14** (1979) 1483–1485.
14. D. G. Luenberger, *Introduction to Dynamic Systems: Theory, Models and Applications*, John Wiley & Sons, New York, 1979.
15. A. P. Molchanov and E. S. Pyatnitskij, Criteria of asymptotic stability of differential and difference inclusions encountered in control theory, *Syst. Control Lett.* **13** (1989) 59–64.
16. E. Plischke, *Transient Effects in Linear Dynamical Systems*, PhD thesis, Universität Bremen, Bremen, Germany, 2005.
17. E. Plischke and F. Wirth, Duality results for the joint spectral radius and transient behavior, *Lin. Alg. Appl.* (Accepted for publication).
18. R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
19. G. V. Smirnov, *Introduction to the Theory of Differential Inclusions*, American Mathematical Society, Providence, RI, 2002.
20. N. K. Son and D. Hinrichsen, Robust stability of positive continuous time systems, *Numer. Functional Anal. Optim.* **17** (1996) 649–659.
21. L. N. Trefethen, A. E. Trefethen, S. C. Reddy, and T. A. Driscol, Hydrodynamic stability without eigenvalues, *Science* **261** (1993) 578–584.
22. K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*, Prentice Hall, Upper Saddle River, NJ, 1996.