Proceedings of the 35th
Conference on Decision and Control
Kobe, Japan • December 1996

A dual of mixed μ and on the losslessness of (D, G)-scaling

Gjerrit Meinsma, Yash Shrivastava and Minyue Fu[1]

Abstract

This paper studies the mixed structured singular value, μ, and the well-known (D, G)-scaling upper bound, ν. A complete characterization of the losslessness of ν (i.e., ν being equal to μ) is derived in terms of the numbers of different perturbation blocks. Specifically, it is shown that ν is guaranteed to be lossless if and only if $2(m_r + m_c) + m_C \leq 3$, where m_r, m_c, and m_C are the numbers of repeated real scalar blocks, repeated complex scalar blocks and full complex blocks, respectively. The results hinge on a dual characterization of μ and ν, which intimately links μ with ν. Further, a special case of the aforementioned losslessness result leads to a variation of the well-known Kalman-Yakubovich-Popov lemma and Lyapunov inequalities.

1 Introduction

The paper [4] that coined the structured singular value is also the paper that introduced the D-scaling upper bound, which, to this date is still the most widely used upper bound of the structured singular value. As claimed in [9], D-scaling for complex structures is in practice close to the actual structured singular value (μ for short), which is somewhat surprising considering that computation of μ for complex structures is NP-hard [11]. Even more surprising is that for several nontrivial complex structures the D-scaling upper bound is lossless (i.e. equal to μ) (see [9]).

Similar claims and results are not known for the *mixed real/complex* μ. Mixed real/complex μ is an extension of μ that allows the structure to consist of complex *and real* parts. Such mixed structures arise, for example, if robust stability is to be tested with respect to parametric uncertainties. In Fan, Tits and Doyle [5] the (D, G)-scaling upper bound for mixed μ is introduced, but, unlike its pure complex counterpart, this upper bound—which we call ν—can be far from the actual mixed μ [13]. About losslessness of ν little is known. Fan, Tits and Doyle [5] have shown that ν is lossless if there is one *non-repeated* real scalar and one full complex block. Young [12] showed that ν is lossless for rank-one matrices.

In this paper we show that the upper bound ν of mixed μ is lossless if-and-only-if

$$2(m_r + m_c) + m_C \leq 3, \tag{1}$$

where m_r, m_c, and m_C are the numbers of repeated real scalar blocks, repeated complex scalar blocks and full complex blocks, respectively. It is an if-and-only-if condition in the sense that if the *number* of blocks violate (1) then—irrespective of the *size* of the blocks—always matrices M exist for which $\mu(M) < \nu(M)$. In this paper we also derive a transparent dual formulation of μ and ν which is a result of independent interest. It is partly based on Rantzer's proof of the Kalman-Yakubovich-Popov lemma [10].

Section 2 introduces notation and a few well established results. In Section 3 the dual characterizations of μ and ν are derived. As an example of the use of these dual results we show that $\mu(M) = \nu(M)$ if M has rank one (the proof is a substantial simplification compared to that of Young [12]). The dual characterizations are used in Section 4 to prove the losslessness of ν for the mentioned structures. In Section 5 we give examples that show that for all other structures ν is not guaranteed to be lossless. Section 6 is about a variation of the Kalman-Yakubovich-Popov lemma. Section 7 contains a more direct proof of losslessness of ν for the simpler case that the structure consists of one repeated real scalar block only. The proof relies on a variation of Lyapunov inequalities and on the notion of (antistable) square roots (ASRs) of non-Hermitian matrix. Interestingly, the definition of ASRs supersedes that of the ubiquitous square roots of Hermitian matrices. Basic properties of ASRs are reviewed in the appendix.

Sections 2 through 6 cover the same material as [7] save some details and proofs that have been omitted due to lack of space.

2 Notation and (D, G)-scaling

The norm $\|T\|$ of a matrix $T \in \mathbb{C}^{m\times n}$ is in this note the spectral norm. The Euclidean norm of T is denoted as $\|T\|_2$. T^H is the complex conjugate transpose of T, and He T is the Hermitian part T defined as

$$\operatorname{He} T = \frac{1}{2}(T + T^H).$$

For scalar T the Hermitian part is the real part. Given a subset $\mathbb{X}$ of $\mathbb{C}^{n\times n}$ the *(mixed) structured singular value* of $M \in \mathbb{C}^{n\times n}$ is denoted by $\mu_{\mathbb{X}}(M)$ and defined as

$$\mu_{\mathbb{X}}(M) = \frac{1}{\inf\{\|\Delta\| : I - \Delta M \text{ is singular and } \Delta \in \mathbb{X}\}}.$$

$\mu_{\mathbb{X}}(M)$ is set to zero if $I - \Delta M$ is nonsingular for every $\Delta \in \mathbb{X}$. Obviously $\mu_{\mathbb{X}}(M)$ depends on the "structure"

[1]Department of Electrical and Computer Engineering, The University of Newcastle, N.S.W. 2308, Australia. Ph: +61-49-217023, Fax: +61-49-216993, Email: meinsma@ee.newcastle.edu.au

$\mathbb{X}$. Whenever $\mu_{\mathbb{X}}(M)$ is used it is implicitly assumed that some structure $\mathbb{X}$ is given. Invariably $\mathbb{X}$ is assumed block-diagonal of the form

$$\begin{aligned}\mathbb{X} \;=\; \operatorname{diag}(&\mathbb{R}I_{k_1},\dots,\mathbb{R}I_{k_{m_r}},\\ &\mathbb{C}I_{l_1},\dots,\mathbb{C}I_{l_{m_c}},\\ &\mathbb{C}^{f_1\times f_1},\dots,\mathbb{C}^{f_{m_C}\times f_{m_C}}),\end{aligned} \tag{2}$$

where m_r, m_c and m_C are the number of repeated real scalar blocks, repeated complex scalar blocks and full complex blocks, respectively.

2.1 (D,G)-scaling

Let $\mathbb{H}^q$ denote the set of $q\times q$ Hermitian matrices and denote its subset of positive definite elements by $\mathbb{P}^q$. Given the structure $\mathbb{X}$ of (2), the sets $\mathcal{D}_{\mathbb{X}}$ and $\mathcal{G}_{\mathbb{X}}$ are defined as

$$\begin{aligned}\mathcal{D}_{\mathbb{X}} \;=\; \operatorname{diag}(&\mathbb{P}^{k_1},\dots,\mathbb{P}^{k_{m_r}},\\ &\mathbb{P}^{l_1},\dots,\mathbb{P}^{l_{m_c}},\\ &\mathbb{P}I_{f_1},\dots,\mathbb{P}I_{f_{m_C}}),\\ \mathcal{G}_{\mathbb{X}} \;=\; \operatorname{diag}(&\mathbb{H}^{k_1},\dots,\mathbb{H}^{k_{m_r}},\\ &0_{l_1\times l_1},\dots,0_{l_{m_c}\times l_{m_c}},\\ &0_{f_1\times f_1},\dots,0_{f_{m_C}\times f_{m_C}}).\end{aligned}$$

Given $M\in\mathbb{C}^{n\times n}$ and $\alpha\in\mathbb{R}$ the matrix function $\Phi_\alpha(D,G)$ is defined as

$$\Phi_\alpha(D,G) = M^H D M + j(GM - M^H G) - \alpha^2 D.$$

This notation is a bit different from that of [5]. Fan, Tits and Doyle [5] showed that $\mu_{\mathbb{X}}(M) < \alpha$ if $\alpha > 0$ and $\Phi_\alpha(D,G) < 0$ for some $D\in\mathcal{D}_{\mathbb{X}}$ and $G\in\mathcal{G}_{\mathbb{X}}$. The infimal α for which such D and G can be found is thus an upper bound of $\mu_{\mathbb{X}}(M)$ and this upper bound is denoted throughout as $\nu_{\mathbb{X}}(M)$, i.e.,

$$\nu_{\mathbb{X}}(M) = \inf_{D\in\mathcal{D}_{\mathbb{X}},\,G\in\mathcal{G}_{\mathbb{X}}}\{\alpha>0 : \Phi_\alpha(D,G)<0\}.$$

The importance of the upper bound $\nu_{\mathbb{X}}$ lies in the fact that it can be computed "efficiently" (in polynomial time) whereas computation of $\mu_{\mathbb{X}}$ is NP-hard. It may be verified that

$$\Phi_\alpha(D,G) = \operatorname{He}\,(M^H+\alpha I)(D+\frac{j}{\alpha}G)(M-\alpha I).$$

This allows to characterize $\nu_{\mathbb{X}}$ somewhat more compactly. Given α, any element E of $\mathcal{D}_{\mathbb{X}}+j\mathcal{G}_{\mathbb{X}}$ can be uniquely decomposed as $E = D + (j/\alpha)G$ with $D\in\mathcal{D}_{\mathbb{X}}$ and $G\in\mathcal{G}_{\mathbb{X}}$ (namely, take $D = \operatorname{He} E$ and let $(j/\alpha)G$ be the skew-Hermitian matrix $E - \operatorname{He} E$). Therefore

$$\begin{aligned}\nu_{\mathbb{X}}(M)<\alpha \quad\Leftrightarrow\quad &\exists E\in\mathcal{D}_{\mathbb{X}}+j\mathcal{G}_{\mathbb{X}} \text{ such that}\\ &\operatorname{He}\,(M^H+\alpha I)E(M-\alpha I)<0.\end{aligned}$$

This we use frequently.

3 Dual characterization of μ and ν

In this section we give a dual characterization of $\mu_{\mathbb{X}}$ and $\nu_{\mathbb{X}}$. In the next section we use these results to prove that $\mu_{\mathbb{X}} = \nu_{\mathbb{X}}$ for structures of the form $\mathbb{X} = \operatorname{diag}(\mathbb{R}I_m, \mathbb{C}^{p\times p})$. The dual characterizations of $\mu_{\mathbb{X}}$ and $\nu_{\mathbb{X}}$ that we present are easy and they are remarkably similar. It is tempting to think they have wider use than just the next sections, and this is exemplified by the fact that with these dual formulations it is easy to prove that $\mu_{\mathbb{X}}(M) = \nu_{\mathbb{X}}(M)$ for rank-one matrices M (irrespective of the structure). This is done at the end of this section.

The characterizations presented are dual in that they are an application of a duality argument for convex sets. The following preparatory result is in essence standard (see Boyd et.al. [2, page 29]).

Lemma 3.1 *Suppose $F(E)\in\mathbb{C}^{m\times m}$ depends affinely on $E\in\mathbb{C}^{n\times n}$. Let $\mathcal{E}$ be some convex subset of $\mathbb{C}^{n\times n}$. Then no $E\in\mathcal{E}$ exists for which*

$$\operatorname{He} F(E) < 0$$

iff there is a nonzero $W = W^H \geq 0$ such that

$$\operatorname{Re}\operatorname{tr} WF(E) \geq 0 \quad \forall E\in\mathcal{E}. \tag{3}$$

$\diamond$

Lemma 3.2 (Dual characterization of $\nu_{\mathbb{X}}$) *Let $\mathbb{X}$ be any structure (2). Then $\nu_{\mathbb{X}}(M)\geq\alpha$ iff there is a nonzero $W = W^H\geq 0$ such that*

$$\operatorname{Re}\operatorname{tr}\,(M-\alpha I)W(M^H+\alpha I)E \geq 0 \quad \forall E\in\mathcal{D}_{\mathbb{X}}+j\mathcal{G}_{\mathbb{X}}. \tag{4}$$

Proof: $\nu_{\mathbb{X}}(M)\geq\alpha$ iff no $E\in\mathcal{D}_{\mathbb{X}}+j\mathcal{G}_{\mathbb{X}}$ exists for which

$$\operatorname{He}\,(M^H+\alpha I)E(M-\alpha I)<0.$$

By Lemma 3.1 that is the case iff there is a $W = W^H\geq 0$ such that $\operatorname{Re}\operatorname{tr} W(M^H+\alpha I)E(M-\alpha I)\geq 0$ for all such E. The traces of $W(M^H+\alpha I)E(M-\alpha I)$ and $(M-\alpha I)W(M^H+\alpha I)E$ are the same. ■

We next reformulate this characterization of $\nu_{\mathbb{X}}$ without using E. To this end we partition E and $(M-\alpha I)W(M^H+\alpha I)$ compatible with structure $\mathbb{X}$ as

$$\begin{aligned}E &= \begin{bmatrix} E_1 & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & E_{m_r+m_c+m_C}\end{bmatrix},\\ (M-\alpha I)W(M^H+\alpha I) &= \begin{bmatrix} Z_1 & ? & ?\\ ? & \ddots & ?\\ ? & ? & Z_{m_r+m_c+m_C}\end{bmatrix}\end{aligned} \tag{5}$$

A "?" denotes an irrelevant entry. Varying E over all elements of $\mathcal{D}_{\mathbb{X}}+j\mathcal{G}_{\mathbb{X}}$ can be done by varying each block E_i independently of the other blocks, and as each block may be arbitrarily close to zero, we have that (4) holds iff for every $i\in\{1,\dots,m_r+m_c+m_C\}$ we have that

$$\operatorname{Re}\operatorname{tr}\, Z_i E_i \geq 0 \tag{6}$$

holds for all E_i in the appropriate sets.

Lemma 3.3 (Dual characterization of $\nu_{\mathbb{X}}$) *Let $\mathbb{X}$ be any structure (2). Then $\nu_{\mathbb{X}}(M) \geq \alpha$ iff there is a nonzero $W = W^H \geq 0$ such that*

$$\begin{cases} Z_i \text{ is Hermitian and} \geq 0 \;\; \forall i \in \{1, \ldots, m_r\}, \\ \mathrm{He}\, Z_i \geq 0 \;\; \forall i \in \{m_r+1, \ldots, m_r+m_c\}, \\ \mathrm{Re\,tr}\, Z_i \geq 0 \;\; \forall i \in \{m_r+m_c+1, \ldots, m_r+m_c+m_C\}. \end{cases} \tag{7}$$

Here Z_i is the i-th block on the diagonal of $(M - \alpha I)W(M^H + \alpha I)$ as shown in (5).

Proof: $\nu_{\mathbb{X}}(M) \geq \alpha$ iff (6) holds for all i and E_i in the appropriate sets. We distinguish three cases.

(Case 1:) If $i \in \{1, \ldots, m_r\}$ then E_i is any matrix whose Hermitian part He E_i is positive definite. For all such E_i we have that $\mathrm{Re\,tr}\, Z_iE_i \geq 0$ iff Z_i is Hermitian and ≥ 0. (See [7] for more details.)

(Case 2:) For $i \in \{m_r+1, \ldots, m_r+m_c\}$ the E_i is any Hermitian positive definite matrix. $\mathrm{Re\,tr}\, Z_iE_i \geq 0$ for all such E_i iff He $Z_i \geq 0$.

(Case 3:) For $i \in \{m_r+m_c+1, \ldots, m_r+m_c+m_C\}$ the E_i is any matrix of the form $E_i = d_iI$ with $0 < d_i \in \mathbb{R}$. $\mathrm{Re\,tr}\, Z_iE_i \geq 0$ for all such E_i iff $\mathrm{Re\,tr}\, Z_i \geq 0$. ∎

Next we derive a characterization of $\mu_{\mathbb{X}}$ in similar terms. For that we need the following lemma. It is readily proved.

Lemma 3.4 (Three little lemmas) *Let f, g be two column vectors of the same dimension.*

1. *$(f-g)(f+g)^H$ is Hermitian and ≥ 0 iff $g = \delta f$ for some $\delta \in [-1, 1]$.*
2. *The Hermitian part of $(f-g)(f+g)^H$ is ≥ 0 iff $g = \delta f$ for some $\delta \in \mathbb{C}$ with $|\delta| \leq 1$.*
3. *$\mathrm{Re\,tr}\,(f-g)(f+g)^H = \|f\|_2^2 - \|g\|_2^2$. Hence $\mathrm{Re\,tr}\,(f-g)(f+g)^H \geq 0$ iff $g = \Delta f$ for some matrix Δ with $\|\Delta\| \leq 1$.*

◇

Lemma 3.5 (Dual characterization of $\mu_{\mathbb{X}}$) *Let $\mathbb{X}$ be any structure (2). Then $\mu_{\mathbb{X}}(M) \geq \alpha$ iff there is a nonzero vector $t \in \mathbb{C}^n$ such that*

$$\mathrm{Re\,tr}\,(M - \alpha I)tt^H(M^H + \alpha I)E \geq 0 \quad \forall E \in \mathcal{D}_{\mathbb{X}} + j\mathcal{G}_{\mathbb{X}}. \tag{8}$$

Equivalently, $\mu_{\mathbb{X}}(M) \geq \alpha$ iff there is a nonzero vector $t \in \mathbb{C}^n$ such that

$$\begin{cases} Z_i \text{ is Hermitian and} \geq 0 \;\; \forall i \in \{1, \ldots, m_r\}, \\ \mathrm{He}\, Z_i \geq 0 \;\; \forall i \in \{m_r+1, \ldots, m_r+m_c\}, \\ \mathrm{Re\,tr}\, Z_i \geq 0 \;\; \forall i \in \{m_r+m_c+1, \ldots, m_r+m_c+m_C\}. \end{cases} \tag{9}$$

Here Z_i is the i-th block on the diagonal of $(M - \alpha I)tt^H(M^H + \alpha I)$ as partitioned compatibly with $\mathbb{X}$. ◇

Proof: The equivalence of (8) and (9) was shown earlier. We prove that $\mu_{\mathbb{X}}(M) \geq \alpha$ iff (9) holds. Note that Z_i can be written as the product of a column vector and a row vector as

$$Z_i = \begin{bmatrix} M_{i1} \cdots M_{i(i-1)} & M_{ii} - \alpha I & M_{i(i+1)} \cdots \end{bmatrix} t \tag{10}$$
$$\left(\begin{bmatrix} M_{i1} \cdots M_{i(i-1)} & M_{ii} + \alpha I & M_{i(i+1)} \cdots \end{bmatrix} t \right)^H.$$

We will formulate necessary and sufficient conditions for (9) to hold. We distinguish the three cases.

(Case 1.) Let $i \in \{1, \ldots, m_r\}$. By Lemma 3.4, Item 1 we have that (10) is Hermitian and ≥ 0 iff $\begin{bmatrix} 0 \cdots 0 & \alpha I & 0 \cdots 0 \end{bmatrix} t = \delta_i \begin{bmatrix} M_{i1} & \cdots & M_{i(m_r+m_c+m_C)} \end{bmatrix} t$ for some $\delta_i \in [-1, 1]$.

(Case 2.) Let $i \in \{m_r+1, \ldots, m_r+m_c\}$. By Lemma 3.4, Item 2 we have that the Hermitian part of (10) is ≥ 0 iff $\begin{bmatrix} 0 \cdots 0 & \alpha I & 0 \cdots 0 \end{bmatrix} t = \delta_i \begin{bmatrix} M_{i1} & \cdots & M_{i(m_r+m_c+m_C)} \end{bmatrix} t$ for some $\delta_i \in \mathbb{C}$ with $|\delta_i| \leq 1$.

(Case 3.) Let $i \in \{m_r+m_c+1, \ldots, m_r+m_c+m_C\}$. By Lemma 3.4, Item 3 the real part of the trace of (10) is ≥ 0 iff $\begin{bmatrix} 0 \cdots 0 & \alpha I & 0 \cdots 0 \end{bmatrix} t = \Delta_i \begin{bmatrix} M_{i1} & \cdots & M_{i(m_r+m_c+m_C)} \end{bmatrix} t$ for some Δ_i with $\|\Delta_i\| \leq 1$.

The three cases combined show that there is a nonzero t that satisfies (9) if and only if $(\alpha I - \Delta M)t = 0$ for some t and some $\Delta = \mathrm{diag}(\delta_1, \ldots, \delta_{m_r}, \delta_{m_r+1}, \ldots, \delta_{m_r+m_c}, \widehat{\Delta}_1, \ldots, \widehat{\Delta}_{m_C}) \in \mathbb{X}$ with $\|\Delta\| \leq 1$, i.e., iff $\mu_{\mathbb{X}}(M) \geq \alpha$. ∎

In summary, the results in this section say that $\nu_{\mathbb{X}}(M) \geq \alpha$ iff a nonzero $W = W^H \geq 0$ exists with certain properties (7), and that $\mu_{\mathbb{X}}(M) \geq \alpha$ iff that W can be chosen to have rank 1. Another interpretation, and which is more in line with that of Packard and Doyle [9] and Rantzer [10], is as follows. Fix α. The set

$$\{(M - \alpha I)W(M^H + \alpha I) : W = W^H \geq 0\}$$

is the convex hull of the set

$$\Theta := \{(M - \alpha I)tt^H(M^H + \alpha I) : t \in \mathbb{C}^n\}.$$

Therefore $\nu_{\mathbb{X}}(M) \geq \alpha$ iff the convex hull of Θ has certain properties, whereas $\mu_{\mathbb{X}}(M) \geq \alpha$ iff Θ itself has those properties.

We end this section with an application which shows the potential of the dual characterizations. Young [12] was the first to prove the following lemma, but whereas his proof is rather cumbersome, the proof based on dual characterizations is a few lines only.

Lemma 3.6 $\mu_{\mathbb{X}}(M) = \nu_{\mathbb{X}}(M)$ *if M has rank one.*

Proof: It suffices to show that $\nu_{\mathbb{X}}(M) \geq \alpha$ implies $\mu_{\mathbb{X}}(M) \geq \alpha$. Suppose $\nu_{\mathbb{X}}(M) \geq \alpha$. Therefore there is a nonzero nonnegative definite $W = W^H$ for which $(M - \alpha I)W(M^H + \alpha I)$ satisfies the positivity conditions

(7). Let $x, y \in \mathbb{C}^n$ be such that $M = xy^H$, and decompose W compatibly with that as

$$W = tt^H + \widehat{W}$$

in which $\widehat{W}y = 0$, $\widehat{W} = \widehat{W}^H \geq 0$ and $t \in \mathbb{C}^n$. Such a decomposition always exists. Then we have

$$\begin{aligned}(M - \alpha I)tt^H(M^H + \alpha I) \\ &= (xy^H - \alpha I)(W - \widehat{W})(yx^H + \alpha I) \\ &= (M - \alpha I)W(M^H + \alpha I) + \alpha^2 \widehat{W}.\end{aligned}$$

By assumption, $(M - \alpha I)W(M^H + \alpha I)$ satisfies the positivity conditions (7), but then so does $(M - \alpha I)tt^H(M^H + \alpha I)$ because $\alpha^2 \widehat{W}$ is Hermitian and ≥ 0. Hence $\mu_{\mathbb{X}}(M) \geq \alpha$, which is what we needed to prove. (Note that the vector t is nonzero, because, otherwise, $(M - \alpha I)W(M^H + \alpha I) = -\alpha^2 \widehat{W} \leq 0$ which would have contradicted (7).) ∎

4 The case $\mathbb{X} = \operatorname{diag}(\mathbb{R}I_m, \mathbb{C}^{p \times p})$

In this section we prove that $\mu_{\mathbb{X}} = \nu_{\mathbb{X}}$ if the structure has the form $\mathbb{X} = \operatorname{diag}(\mathbb{R}I_m, \mathbb{C}^{p \times p})$. A straightforward application of Lemma 3.3 and Lemma 3.5 is as follows.

Corollary 4.1 *Let $\mathbb{X} = \operatorname{diag}(\mathbb{R}I_m, \mathbb{C}^{p \times p})$. Then $\nu_{\mathbb{X}}(M) \geq \alpha$ iff a nonzero Hermitian $W \geq 0$ exists for which*

$$\begin{cases} \begin{bmatrix} M_{11} - \alpha I_m & M_{12} \end{bmatrix} W \begin{bmatrix} M_{11}^H + \alpha I_m \\ M_{12}^H \end{bmatrix} \text{ is Hermitian} \geq 0, \\ \operatorname{Re} \operatorname{tr} \begin{bmatrix} M_{21} & M_{22} - \alpha I_p \end{bmatrix} W \begin{bmatrix} M_{21}^H \\ M_{22}^H + \alpha I_p \end{bmatrix} \geq 0. \end{cases} \tag{11}$$

Moreover $\mu_{\mathbb{X}}(M) \geq \alpha$ iff a $W = W^H \geq 0$ of rank one exists with these properties (11). ◇

The following is a technical result that we need later. It can be proved with induction (see [7] for proof).

Lemma 4.2 (Technical result) *Let F and G be complex matrices of the same dimensions. If a $W = W^H \geq 0$ of rank n is such that*

$$FWG^H \text{ is Hermitian and} \geq 0,$$

then there exist n column vectors t_k such that $W = \sum_{k=1}^{k=n} t_k t_k^H$ and $F t_k t_k^H G^H$ is Hermitian and ≥ 0 $\forall k$. ◇

Theorem 4.3 $\mu_{\mathbb{X}} = \nu_{\mathbb{X}}$ *if* $\mathbb{X} = \operatorname{diag}(\mathbb{R}I_m, \mathbb{C}^{p \times p})$.

Proof: Since $\mu_{\mathbb{X}}(M) \leq \nu_{\mathbb{X}}(M)$ it suffices to prove that $\nu_{\mathbb{X}}(M) \geq \alpha$ implies $\mu_{\mathbb{X}}(M) \geq \alpha$.

Suppose $\nu_{\mathbb{X}}(M) \geq \alpha$. Then by Corollary 4.1 there is a nonzero $W = W^H \geq 0$ that satisfies (11). By Lemma 4.2 we can write this W as $W = \sum_k t_k t_k^H$ such that for all k

$$\begin{bmatrix} M_{11} - \alpha I & M_{12} \end{bmatrix} t_k t_k^H \begin{bmatrix} M_{11}^H + \alpha I \\ M_{12}^H \end{bmatrix} \text{ is Hermitian} \geq 0.$$

Since $W = \sum_k t_k t_k^H$ satisfies (11), there is at least one index k, say $k = l$, for which $t_l \neq 0$ and

$$\operatorname{Re} \operatorname{tr} \begin{bmatrix} M_{21} & M_{22}^H - \alpha I \end{bmatrix} t_l t_l^H \begin{bmatrix} M_{21}^H \\ M_{22}^H + \alpha I \end{bmatrix} \geq 0.$$

Hence $W := t_l t_l^H$ is a rank one matrix that satisfies (11) so that $\mu_{\mathbb{X}}(M) \geq \alpha$. ∎

5 Losslessness of (D, G)-scaling

Table 1: When $\mu_{\mathbb{X}} = \nu_{\mathbb{X}}$ is guaranteed.

	$m_C = 0$	$m_C = 1$	$m_C = 2$	$m_C = 3$
$m_r = 0$, $m_c = 0$		Yes. See [9]	Yes. See [9]	Yes. See [9]
$m_r = 0$, $m_c = 1$	Yes. See [9]	Yes. See [9]	No. See [9]	
$m_r = 1$, $m_c = 0$	Yes. Thm. 4.3	Yes. Thm. 4.3	No. Ex. 5.1	
$m_r = 1$, $m_c = 1$	No. Ex. 5.2			
$m_r = 0$, $m_c = 2$	No. See [9]			
$m_r = 2$, $m_c = 0$	No. See [8]			

Building on work by [4], Packard and Doyle [9] showed that $\mu_{\mathbb{X}} = \nu_{\mathbb{X}}$ whenever

$$m_r = 0, \quad 2m_c + m_C \leq 3. \tag{12}$$

Together with the results of Section 4 we thus have that $\mu_{\mathbb{X}} = \nu_{\mathbb{X}}$ for any of the structures $\mathbb{X}$ for which

$$2(m_r + m_c) + m_C \leq 3. \tag{13}$$

Packard and Doyle [9] further show by examples that $\mu_{\mathbb{X}}(M) < \nu_{\mathbb{X}}(M)$ can occur within any complex structure with $2m_c + m_C > 3$ (and $m_r = 0$). For $m_r = 2$ it is possible to construct 2-by-2 matrices M for which $\mu_{\mathbb{X}}(M) < \nu_{\mathbb{X}}(M)$ (see [8]). Table 1 details (13) and gives references for the various cases. In this section we give two examples that complete the picture in that they—combined with the other examples—show that for any structure $\mathbb{X}$ that violates (13) there exist matrices M such that $\mu_{\mathbb{X}}(M) < \nu_{\mathbb{X}}(M)$.

Example 5.1 Let $\mathbb{X} = \operatorname{diag}(\mathbb{R}, \mathbb{C}, \mathbb{C})$ and take

$$M = \begin{bmatrix} 0 & 1 & j \\ j & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

We show that $\mu_{\mathbb{X}}(M) = 1$ and that $\nu_{\mathbb{X}}(M) = \sqrt{3}$. The spectral norm of M is $\|M\| = \sqrt{3}$, so we have $\nu_{\mathbb{X}}(M) \leq \sqrt{3}$. Furthermore, for the Hermitian nonnegative definite W defined as

$$W = \begin{bmatrix} 2 & 1 & -j \\ 1 & 2 & j \\ j & -j & 2 \end{bmatrix}$$

we have that

$$(M - \sqrt{3}I)W(M^H + \sqrt{3}I) = \begin{bmatrix} 0 & ? & ? \\ ? & 0 & 0 \\ ? & 0 & 0 \end{bmatrix}. \qquad (14)$$

(The "?" denotes an irrelevant entry.) Since all diagonal entries of (14) are zero it follows from Lemma 3.3 that $\nu_{\mathbb{X}}(M) \geq \sqrt{3}$. Hence $\nu_{\mathbb{X}}(M) = \sqrt{3}$.

Calculation shows that $I_3 - \operatorname{diag}(\delta_1, \delta_2, \delta_3)M$ is singular iff

$$\delta_2\delta_3 + j\delta_1(\delta_2 + \delta_3) - 1 = 0. \qquad (15)$$

Suppose $\delta_1 \in [-1, 1]$ and that $|\delta_3| < 1$. Then the δ_2 for which (15) holds, equals $\delta_2 = (1 - j\delta_1\delta_3)/(j\delta_1 + \delta_3)$ and satisfies

$$|\delta_2|^2 = \left|\frac{1 - j\delta_1\delta_3}{j\delta_1 + \delta_3}\right|^2 = \frac{1 + \delta_1^2|\delta_3|^2 - 2\delta_1 \operatorname{Im}(\delta_3)}{\delta_1^2 + |\delta_3|^2 - 2\delta_1 \operatorname{Im}(\delta_3)} \geq 1.$$

Therefore $\max_i |\delta_i| \geq 1$ for every solution of (15). Since $(\delta_1, \delta_2, \delta_3) = (1, j, -j)$ is a solution of (15) for which $\max_i |\delta_i| = 1$ we have that $\mu_{\mathbb{X}}(M) = 1$. ◇

Example 5.2 Let $\mathbb{X} = \operatorname{diag}(\mathbb{R}, \mathbb{C}I_2)$ and take the same M as in the previous example. From (14) we infer that also for this structure $\nu_{\mathbb{X}}(M) \geq \sqrt{3}$. Since $\nu_{\mathbb{X}}(M) \leq \|M\| = \sqrt{3}$ we have, again, that $\nu_{\mathbb{X}}(M) = \sqrt{3}$. It further follows from the previous example that $I_3 - \operatorname{diag}(\delta_1, \delta_2 I_2)M$ is singular iff

$$\delta_2^2 + j2\delta_1\delta_2 - 1 = 0.$$

The solutions $\delta_2 = -j\delta_1 \pm \sqrt{1 - \delta_1^2}$ have absolute value 1 for every $\delta_1 \in [-1, 1]$. Hence $\mu_{\mathbb{X}}(M) = 1 < \nu_{\mathbb{X}}(M)$. ◇

6 A variation of the KYP lemma

The conditions of the Kalman-Yakubovich-Popov lemma (KYP lemma) and the bounded real lemma are known to be equivalent to the fact that $\mu_{\mathbb{X}} = \nu_{\mathbb{X}}$ for the complex structures $\mathbb{X} = \operatorname{diag}(\mathbb{C}I_m, \mathbb{C}^{p\times p})$ (see [9]). In this section we rephrase Theorem 4.3 as a KYP type result. More precisely, Item 2 of the lemma below is this reformulation (for a proof see [7]). Items 1 and 3 have been included for comparison only: Item 1 is an application of the famous KYP lemma in a strict inequality version (see e.g. [1]) and Item 3 is an immediate consequence of Lemma 3.1 of [6].

Lemma 6.1 *Let G is a square rational matrix with realization $G \stackrel{s}{=} [A, B, C, D]$, and consider the following LMI in P:*

$$\begin{bmatrix} PA + A^H P^H & PB - C^H \\ -C + B^H P^H & -D - D^H \end{bmatrix} < 0. \qquad (16)$$

Then

1. *There is a Hermitian $P = P^H > 0$ that satisfies (16) iff A is stable and $G(s) + [G(s)]^H > 0$ for all s in the closed RHP including ∞.*
2. *There is a P (Hermitian or not) with $\operatorname{He} P > 0$ that satisfies (16) iff A has no eigenvalues on the positive real line $[0, \infty)$ and $G(s) + [G(s)]^H > 0$ for all $s \in [0, \infty)$ including ∞.*
3. *There is a P (Hermitian or not) that satisfies (16) iff A is nonsingular and $G(s) + [G(s)]^H > 0$ at $s = 0$ and $s = \infty$.*

7 A special case: $\mathbb{X} = \mathbb{R}I_m$

As a special case of Theorem 4.3 we have that $\mu_{\mathbb{X}} = \nu_{\mathbb{X}}$ if $\mathbb{X} = \mathbb{R}I_n$. In this section we indicate a more direct and insightful proof of this result. The proof relies on the notion of (antistable) square roots of non-Hermitian matrices (cf. Dasgupta and Anderson [3]) and a variation of a Lyapunov inequalities.

Lemma 7.1 (ASR) *If $T \in \mathbb{C}^{n\times n}$ has no eigenvalues on the negative real axis $(-\infty, 0]$, then there is a unique antistable Z such that $T = Z^2$.* ◇

Such a Z will be called the *antistable square root* (ASR) of T. Antistable means that all its eigenvalues lie in the open RHP. The definition of ASR supersedes that of the usual square root of positive definite matrices. (See Appendix for review of ASR results.) The notion of antistable square roots is used to prove a variation of a Lyapunov stability condition:

Lemma 7.2 *$A \in \mathbb{C}^{n\times n}$ has no eigenvalues on the positive real line $[0, \infty)$ iff $\exists P \in \mathbb{C}^{n\times n}$ such that*

$$PA + A^H P^H < 0, \quad \operatorname{He} P > 0. \qquad (17)$$

Proof: If $v \in \mathbb{C}^n$ satisfies $Av = sv$ with $s \in [0, \infty)$ then $v^H(PA + A^H P^H)v = sv^H(P + P^H)v$ which contradicts (17). Conversely, if A has no eigenvalues on $[0, \infty)$, then by Lemma 7.1 we have $-A = Z^2$ for some antistable Z. Let Q be any of the many Hermitian matrices for which $QZ + Z^H Q > 0$. (It follows from standard Lyapunov theory that such $Q = Q^H$ exist because Z is antistable.) Then $P := QZ^{-1}$ satisfies (17) because $PA + A^H P^H = -QZ - Z^H Q < 0$ and $P + P^H = Z^{-H}(Z^H Q + QZ)Z^{-1} > 0$. ■

Theorem 7.3 *Let $\mathbb{X} = \mathbb{R}I_m$. The following are equivalent.*

1. *$\mu_{\mathbb{X}}(M) < 1$, that is, $I - \delta M$ is nonsingular for every $\delta \in [-1, 1]$.*
2. *$(M - I)$ is nonsingular and $(M - I)^{-1}(I + M)$ has no eigenvalues on the positive real axis $[0, \infty)$.*
3. *There is a $C \in \mathbb{C}^{n\times n}$ such that $\operatorname{He} C(I - \delta M) < 0$ for all $\delta \in [-1, 1]$.*
4. *$\nu_{\mathbb{X}}(M) < 1$, i.e., $\exists E \in \mathcal{D}_{\mathbb{X}} + j\mathcal{G}_{\mathbb{X}} = \{E : \operatorname{He} E > 0\}$ such that $\operatorname{He}\,(I + M^H)E(M - I) < 0$.*

Proof: Define $A = (M - I)^{-1}(I + M)$. We prove $1 \implies 2 \implies 3 \implies 4 \implies 1$. (In what follows conv(U, V) denotes the set of convex combinations of U and V.)

(1 $\Longrightarrow$ 2.) $\{I-\delta M : \delta \in [-1,1]\} = \mathrm{conv}(I-M, I+M) = (I-M)\mathrm{conv}(I,-A)$. The set $\mathrm{conv}(I,-A)$ has a singular element iff some eigenvalue of A lies on $[0,\infty)$.

(2 $\Longrightarrow$ 3.) By Lemma 7.2 there is a P such that $\mathrm{He}\, PA < 0$ and $\mathrm{He}\, P > 0$. Let $C = P(M-I)^{-1}$. Then

$$\begin{aligned}\{C(I-\delta M) : \delta \in [-1,1]\} &= C\,\mathrm{conv}(I-M, I+M) \\ &= P\,\mathrm{conv}(-I, A) = \mathrm{conv}(-P, PA).\end{aligned}$$

$\mathrm{He}\,\mathrm{conv}(-P, PA) < 0$ because both $-P$ and PA have negative definite Hermitian part.

(3 $\Longrightarrow$ 4.) E defined as $E = -(I+M^H)^{-1}C$ works.

(4 $\Longrightarrow$ 1.) Direct, since $\mu_{\mathbf{X}}(M) \leq \nu_{\mathbf{X}}(M)$.

■

Acknowledgement: We thank Peter Young for his comments on the losslessness of (D,G)-scaling and Soura Dasgupta for some discussions on the single real block case.

8 Appendix: Square roots

This appendix contains a more or less self-contained exposition of square roots of non-Hermitian matrices. Recall that a $Z \in \mathbb{C}^{n\times n}$ is called an *antistable square root* (ASR) of $T \in \mathbb{C}^{n\times n}$ if Z is antistable and $T = Z^2$.

Lemma 8.1 *Let $T \in \mathbb{C}^{n\times n}$. The following holds.*

1. *The antistable square root of T is unique.*
2. *If Z is an ASR of T, then λ is an eigenvalue of T if and only if $\sqrt{\lambda}$ is an eigenvalue of Z. (Here we assume that the square root $\sqrt{\lambda}$ is taken to be in the closed RHP.)*
3. *T has an antistable square root iff T has no eigenvalues on the negative real axis $(-\infty, 0]$.*
4. *If $T > 0$ then the well known positive definite square root $T^{1/2}$ is the same as the ASR of T.*

Proof:

1. Suppose Z_1 and Z_2 are two ASRs of T. Define $E = Z_1 - Z_2$. Then
$$EZ_1 + Z_2E = (Z_1 - Z_2)Z_1 + Z_2(Z_1 - Z_2) = 0.$$
That is, E satisfies the Lyapunov equation $EZ_1 + Z_2E = 0$. Since both Z_1 and Z_2 are antistable it follows from standard Lyapunov theory that E is unique: $E = 0$, i.e., $Z_1 = Z_2$.
2. Let $\sqrt{\lambda}$ denote the square root *in the closed RHP*. We have that
$$(T - \lambda I_n) = (Z^2 - \lambda I_n) = (Z + \sqrt{\lambda} I_n)(Z - \sqrt{\lambda} I_n).$$
Since $Z + \sqrt{\lambda} I_n$ is nonsingular—its eigenvalues lie in the open RHP—we have that $T - \lambda I_n$ is singular if and only if $Z - \sqrt{\lambda} I_n$ is singular.
3. Item 2 implies that T can have an ASR only if none of its eigenvalues lie on the negative real axis $(-\infty, 0]$. Let $J = Q^{-1}TQ$ be a Jordan normal form similar to T (that is Q is nonsingular and J is diagonal possibly with some entries 1 just above the diagonal). Note that none of the diagonal entries of J lie on the negative real axis $(-\infty, 0]$. It follows trivially by construction that then J has an antistable square root $\hat{Z}$. Then $Z := Q\hat{Z}Q^{-1}$ is an ASR of T because it is antistable and $Z^2 = Q\hat{Z}Q^{-1}Q\hat{Z}Q^{-1} = Q\hat{Z}^2Q^{-1} = QJQ^{-1} = T$.
4. The standard "square root" $T^{1/2}$ of a positive definite T is by definition > 0. Hence the eigenvalues of $T^{1/2}$ are real > 0, i.e., $T^{1/2}$ is antistable.

■

References

[1] B.D.O. Anderson and S. Vongpanitlerd. *Network Analysis and Synthesis*. Prentice Hall, 1973.

[2] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. SIAM, Philadelphia, 1994.

[3] S. Dasgupta and B.D.O. Anderson. Square root of linear time varying systems with applications. *IEEE Trans. on Circuits and Systems*. To appear.

[4] J. C. Doyle. Analysis of feedback systems with structured uncertainties. *IEE Proc., Part D*, 129:242–250, 1982.

[5] M. Fan, A. Tits, and J. Doyle. Robustness in the presence of joint parametric uncertainty and unmodeled dynamics. *IEEE Trans. on Aut. Control*, 36(1):25–38, 1991.

[6] P. Gahinet and P. Apkarian. A linear matrix inequality approach to $\mathcal{H}_\infty$ control. *Int. J. of Robust and Nonlinear Control*, 4:421–448, 1994.

[7] G. Meinsma, Y. Shrivastava, and M. Fu. A dual formulation of mixed μ and the losslessness of (D,G)-scaling. Technical Report EE9541, University of Newcastle, Australia, 1995. Accepted for publication in *IEEE TAC*.

[8] G. Meinsma, Y. Shrivastava, and M. Fu. Some properties of an upper bound of μ. 1995. Accepted for publication in *IEEE TAC*.

[9] A. Packard and J. Doyle. The complex structured singular value. *Automatica*, 29(1):71–109, 1993.

[10] A. Rantzer. A note on the Kalman-Yacubovich-Popov lemma. In *Proceedings of the 3rd European Control Conference*, pages 1792–1795, 1995.

[11] O. Toker and H. Ozbay. On the NP-hardness of the purely complex μ computation, analysis/synthesis, and some related problems in multidimensional systems. In *Proceedings of the American Control Conference*, pages 447–451, 1995.

[12] P. M. Young. The rank one mixed μ problem and 'Kharitonov-type' analysis. *Automatica*, 30(12):1899–1911, 1994.

[13] P.M. Young, M.P. Newlin, and J.C. Doyle. Let's get real. In *Robust Control Theory*, pages 143–174. Springer Verlag, 1995.