INTERNATIONAL JOURNAL OF ROBUST AND NONLINEAR CONTROL, VOL. 7, 187–209 (1997)

COMPLEX AND REAL PERFORMANCE RADII AND THEIR COMPUTATION

TING-SHU HU* AND LI QIU

Department of Electrical and Electronic Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

SUMMARY

This paper considers the problem of robust performance of a linear time-invariant system in the $\mathscr{H}_\infty$ norm. The concepts of complex and real performance radii are introduced to describe the smallest size of dynamic or parametric perturbations to a feedback system that either destabilize the system or destroy a performance bound in a certain closed-loop transfer matrix of the system. An algorithm to compute the complex performance radius is given. For the real performance radius, a lower bound, which often turns out to be exact, is obtained.

International Journal of Robust and Nonlinear Control vol. 7, 187–209 (1997)
(No. of Figures: 4 No. of Tables: 0 No. of Refs: 15)

Key words: uncertain systems, robust stability, robust performance, $\mathscr{H}_\infty$ norm

1. INTRODUCTION

This paper concerns the robust $\mathscr{H}_\infty$ performance of a linear time-invariant (LTI) system under dynamic or parametric perturbation. Consider the uncertain system shown in Figure 1. Let $\mathscr{RH}_\infty$ denote the ring of real rational functions in $\mathscr{H}_\infty$. Assume that $G = \left[\begin{smallmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{smallmatrix}\right] \in \mathscr{RH}_\infty^{(q+p)\times(l+m)}$ is a transfer matrix representing an LTI system and $\Delta \in \mathbb{F}^{m\times p}$, where $\mathbb{F}$ is either the complex field $\mathbb{C}$ or the real field $\mathbb{R}$. The transfer matrix from w to z is then given by the following linear fractional transformation:

$$\mathscr{F}(G,\Delta) = G_{11} + G_{12}(I - \Delta G_{22})^{-1}\Delta G_{21}$$

The system is said to be internally stable if $(I - \Delta G_{22})^{-1}$ exists and belongs to $\mathscr{H}_\infty^{m\times m}$.

Now assume that $\|G_{11}\|_\infty < 1$. Define the performance radius of G to be

$$\begin{aligned} pr_{\mathbb{F}}(G) := \sup\{r\colon & \mathscr{F}(G,\Delta) \text{ is internally stable and} \\ & \|\mathscr{F}(G,\Delta)\|_\infty < 1 \text{ for all } \Delta \in \mathbb{F}^{m\times p} \text{ with } \|\Delta\| < r\} \end{aligned} \tag{1}$$

Here the norm of Δ is the spectral norm, i.e., the largest singular value. $pr_{\mathbb{C}}(G)$ will be called the complex performance radius of G and $pr_{\mathbb{R}}(G)$ the real performance radius. The purpose of this paper is to study the computation of $pr_{\mathbb{C}}(G)$ and $pr_{\mathbb{R}}(G)$. A complete solution to the problem of computing $pr_{\mathbb{C}}(G)$ is given. Roughly speaking, $pr_{\mathbb{C}}(G)$ can be computed via the computation of an

* On leave from Shanghai Jiao Tong University.

Contract grant sponsor: Hong Kong Research Grants Council. Contract grant number: HKUST 552/94E.

CCC 1049-8923/97/020187–23

Received 2 November 1995
Revised 4 May 1996

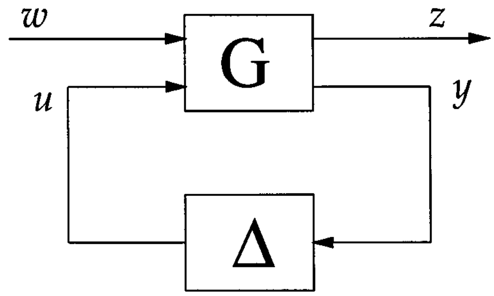

Figure 1. Uncertain control system

$\mathscr{H}_\infty$ norm and a frequency sweep of a function which can be evaluated by performing a one-dimensional convex optimization. The problem of computing $pr_{\mathbb{R}}$ is however not completely solved in this paper. Instead a lower bound is obtained. This lower bound can be obtained via the computation of a real stability radius and a frequency sweep of a function which can be evaluated by performing a two-dimensional nonlinear minimization. It is also shown that the function to be minimized has only one local minimum. Numerical experience shows that this lower bound is often tight.

It is clear that $pr_{\mathbb{R}}(G)$ gives a measure to the robust performance of the system shown in Figure 1 under parametric uncertainty. It can be shown using standard techniques[1] that

$$pr_{\mathbb{C}}(G) = \sup\{r\colon \mathscr{F}(G,\Delta) \text{ is internally stable and } \|\mathscr{F}(G,\Delta)\|_\infty < 1 \text{ for all } \Delta \in \mathscr{H}_\infty^{m\times p} \text{ with } \|\Delta\|_\infty < r\}$$

Hence $pr_{\mathbb{C}}(G)$ gives a measure to the robust performance of the feedback system under LTI dynamic perturbation. Furthermore, if we define the norm of a stable nonlinear system F, i.e., a bounded nonlinear causal operator from $\mathscr{L}_2[0,\infty)$ to $\mathscr{L}_2[0,\infty)$, by

$$\|F\| = \sup_{u\in\mathscr{L}_2[0,\infty), u\neq 0} \frac{\|Fu\|_2}{\|u\|_2}$$

Then we can also show by using the small gain theorem that

$$pr_{\mathbb{C}}(G) = \sup\{r\colon \mathscr{F}(G,\Delta) \text{ is internally stable and } \|\mathscr{F}(G,\Delta)\| < 1 \text{ for all nonlinear time-varying systems } \Delta \text{ with } \|\Delta\| < r\}$$

This says that $pr_{\mathbb{C}}(G)$ also gives a measure to the robust performance of the feedback system under nonlinear time-varying perturbations.

The robust performance measure $pr_{\mathbb{F}}(G)$ is also connected to the robust stability of LTI systems under structured perturbations, which has been studied extensively in the μ framework, see e.g., References 2–5. Consider the uncertain system shown in Figure 2. The small gain theorem implies that this uncertain system is internally stable for all $\Delta \in \mathscr{RH}_\infty^{m\times p}$ with $\|\Delta\|_\infty < r$ and $\tilde{\Delta} \in \mathscr{RH}_\infty^{l\times q}$ with $\|\tilde{\Delta}\|_\infty < 1$ if and only if $pr_{\mathbb{C}}(G) \geqslant r$. In the μ framework, algorithms are available to determine if the uncertain system in Figure 2 is internally stable for all $\Delta \in \mathscr{RH}_\infty^{m\times p}$ with $\|\Delta\|_\infty < r$ and $\tilde{\Delta} \in \mathscr{RH}_\infty^{l\times q}$ with $\|\tilde{\Delta}\|_\infty < r$. Notice the slight difference between the formulation of the complex performance radius and that of μ. Of course, a moment's thought reveals

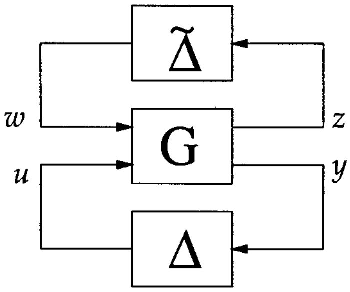

Figure 2. Augmented uncertain control system

that the μ algorithm can be iteratively used to find $pr_{\mathbb{C}}(G)$. In this paper, we propose an algorithm specifically tailored to $pr_{\mathbb{C}}(G)$ so that it is computed directly without this additional iteration.

The complex performance radius, though not called so, is also studied in Reference 6. An algorithm for its computation, free from the additional iteration, is given in Reference 6. The algorithm is based on the quasi-convex univariate minimization of the largest generalized eigenvalue of a semi-definite matrix pair. Our algorithm, however, is different. It is based on the convex univariate minimization of the largest eigenvalue of a Hermitian matrix. Owing to the better convexity property of our minimization problem, faster and more reliable line search methods can be applied. Also notice that generalized eigenvalue problems have higher computational complexity than Hermitian eigenvalue problems (with roughly the same size).

The emphasis of this paper is the real performance radius. Again the small gain theorem implies that the uncertain system in Figure 2 is internally stable for all $\Delta \in \mathbb{R}^{m\times p}$ with $\|\Delta\| < r$ and $\tilde{\Delta} \in \mathscr{H}_\infty^{l\times q}$ with $\|\tilde{\Delta}\|_\infty < 1$ if and only if $pr_{\mathbb{R}}(G) \geqslant r$. This type of robust stability problem with one complex full block and one real full block can also be studied in the μ framework. However, no algorithm is readily available to compute the corresponding μ value. The past literature in μ favours to model parametric perturbations in the form of scalar times identity,[4] which usually ends up with exponential time algorithms. Our conviction is that modelling parameter perturbations in the form of full matrices may in some cases yield easier solutions. Although we are not able to solve completely the computation problem of the real performance radius at this moment, we believe that it is potentially solvable.

The computation of $pr_{\mathbb{F}}$ can also be used to find the worst $\mathscr{H}_\infty$ performance when the perturbation bound is given. The connection again is made possible by the small gain theorem. In the dynamic perturbation case, it is actually the same problem as the computation of $pr_{\mathbb{C}}$:

$$\sup_{\Delta\in\mathscr{RH}_\infty^{m\times p},\ \|\Delta\|_\infty<1} \|\mathscr{F}(G,\Delta)\|_\infty = pr_{\mathbb{C}}\left(\begin{bmatrix} 0 & I \\ I & 0\end{bmatrix} G \begin{bmatrix} 0 & I \\ I & 0\end{bmatrix}\right)$$

In the parametric perturbation case, it can be done by iteratively computing $pr_{\mathbb{R}}$.

$$\sup_{\Delta\in\mathscr{R}^{m\times p},\ \|\Delta\|<1} \|\mathscr{F}(G,\Delta)\|_\infty = \sup\left\{\gamma : pr_{\mathbb{R}}\left(\begin{bmatrix} 1/\gamma I & 0 \\ 0 & I\end{bmatrix} G\right) > 1\right\}$$

The paper is organized as follows. Section 2 is for the preliminary development. We will convert the computation of the performance radius to well-defined linear algebra problems. Section 3 gives a formula for the complex performance radius. Section 4 gives a lower bound of the real performance radius and Section 5 gives some properties of the lower bound. In Section 6, two examples are presented. Section 7 is the conclusion.

In the following, we define some notation used in this paper. For $X \in \mathbb{C}^{m \times p}$, the real and imaginary parts of X are denoted by $\operatorname{Re} X$ and $\operatorname{Im} X$ respectively. The singular values of X are denoted by $\sigma_i(X)$, assuming non-increasing order. The largest singular value of X is also denoted by $\bar{\sigma}(X)$. We always set $\|X\| = \bar{\sigma}(X)$. If X is Hermitian, then the eigenvalues of X are denoted by $\lambda_i(X)$, also assuming non-increasing order.

2. DEVELOPMENT

Recall from References 1, 7 and 8, the definition of the stability radius of $F \in \mathscr{RH}_\infty^{p \times m}$:

$$r_{\mathbb{F}}(F) = \inf\{\|\Delta\| : \Delta \in \mathbb{F}^{m \times p} \text{ and } (I - \Delta F)^{-1} \notin \mathscr{H}_\infty\}$$

For $G \in \mathscr{RH}_\infty^{(q+p) \times (l+m)}$ with $\|G_{11}\|_\infty < 1$, define

$$p_{\mathbb{F}}(G) = \inf_{\omega \in [0, \infty]} \inf\left\{\|\Delta\| : \Delta \in \mathbb{F}^{m \times p}, \quad \det\left\{I - \begin{bmatrix} 0 & \Delta \\ \Delta^* & 0 \end{bmatrix} M(j\omega)\right\} = 0\right\}$$

where

$$M(j\omega) = \begin{bmatrix} S(j\omega) & N(j\omega) \\ N^*(j\omega) & R(j\omega) \end{bmatrix} \tag{2}$$

and

$$S(j\omega) = G_{12}^*(j\omega)[I - G_{11}(j\omega)G_{11}^*(j\omega)]^{-1} G_{12}(j\omega) \tag{3}$$

$$N(j\omega) = G_{22}^*(j\omega) + G_{12}^*(j\omega)G_{11}(j\omega)[I - G_{11}^*(j\omega)G_{11}(j\omega)]^{-1} G_{21}^*(j\omega) \tag{4}$$

$$R(j\omega) = G_{21}(j\omega)[I - G_{11}^*(j\omega)G_{11}(j\omega)]^{-1} G_{21}^*(j\omega) \tag{5}$$

Note that $S(j\omega) \geqslant 0$ and $R(j\omega) \geqslant 0$ for all $\omega \in [-\infty, \infty]$.

In fact, M equals the transfer matrix from $\left[\begin{smallmatrix} u_1 \\ u_2 \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix}\right]$ in Figure 3. By the definition of star product,[9] we also have $M = [(JG^\sim J) \star G]J$, where, $J = \left[\begin{smallmatrix} 0 & I \\ I & 0 \end{smallmatrix}\right]$.

Proposition 1

$pr_{\mathbb{F}}(G) = \min\{r_{\mathbb{F}}(G_{22}), p_{\mathbb{F}}(G)\}$.

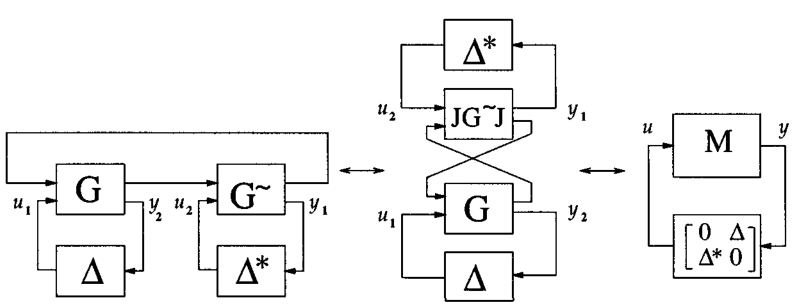

Figure 3. System theoretic interpretation of M

Proof. If $\|\Delta\| < r_{\mathbb{F}}(G_{22})$, then $I - \Delta G_{22}(j\omega)$ is non-singular for all $\omega \in [-\infty, \infty]$. In this case, $\|\mathscr{F}(G, \Delta)\|_\infty = 1$ if and only if for some $\omega \in [-\infty, \infty]$,

$$\det \begin{bmatrix} I & \mathscr{F}[G(j\omega), \Delta] \\ \mathscr{F}^*[G(j\omega), \Delta)] & I \end{bmatrix}$$

$$= \det \left\{ \begin{bmatrix} I & G_{11}(j\omega) \\ G_{11}^*(j\omega) & I \end{bmatrix} - \begin{bmatrix} G_{12}(j\omega) & 0 \\ 0 & -G_{21}^*(j\omega)\Delta^* \end{bmatrix} \begin{bmatrix} 0 & I - G_{22}^*(j\omega)\Delta^* \\ I - \Delta G_{22}(j\omega) & 0 \end{bmatrix}^{-1} \begin{bmatrix} G_{12}^*(j\omega) & 0 \\ 0 & -\Delta G_{21}(j\omega) \end{bmatrix} \right\}$$

$$= 0$$

By using the formula $\det \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] = \det(D)\det(A - BD^{-1}C)$, we see that the above equality is true if and only if

$$\det \begin{bmatrix} I & G_{11}(j\omega) & G_{12}(j\omega) & 0 \\ G_{11}^*(j\omega) & I & 0 & -G_{21}^*(j\omega)\Delta^* \\ G_{12}^*(j\omega) & 0 & 0 & I - G_{22}^*(j\omega)\Delta^* \\ 0 & -\Delta G_{21}(j\omega) & I - \Delta G_{22}(j\omega) & 0 \end{bmatrix}$$

$$= \det \left\{ \begin{bmatrix} I & G_{11}(j\omega) & G_{12}(j\omega) & 0 \\ G_{11}^*(j\omega) & I & 0 & 0 \\ G_{12}^*(j\omega) & 0 & 0 & I \\ 0 & 0 & I & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & G_{21}^*(j\omega) \\ 0 & G_{22}^*(j\omega) \\ I & 0 \end{bmatrix} \begin{bmatrix} 0 & \Delta \\ \Delta^* & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & I \\ 0 & G_{21}(j\omega) & G_{22}(j\omega) & 0 \end{bmatrix} \right\}$$

$$= 0$$

Direct computation shows that

$$M(j\omega) = \begin{bmatrix} 0 & 0 & 0 & I \\ 0 & G_{21}(j\omega) & G_{22}(j\omega) & 0 \end{bmatrix} \begin{bmatrix} I & G_{11}(j\omega) & G_{12}(j\omega) & 0 \\ G_{11}^*(j\omega) & I & 0 & 0 \\ G_{12}^*(j\omega) & 0 & 0 & I \\ 0 & 0 & I & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 0 \\ 0 & G_{21}^*(j\omega) \\ 0 & G_{22}^*(j\omega) \\ I & 0 \end{bmatrix}$$

Hence, if $\|\Delta\| < r_{\mathbb{F}}(G_{22})$, then $\|\mathscr{F}(G, \Delta)\|_\infty = 1$ is equivalent to

$$\det \left\{ I - \begin{bmatrix} 0 & \Delta \\ \Delta^* & 0 \end{bmatrix} M(j\omega) \right\} = 0$$

for some $\omega \in [-\infty, \infty]$.

Notice that in the definition of $p_{\mathbb{F}}(G)$, taking the infimum over $[0,\infty]$ or over $[-\infty,\infty]$ does not make any difference. Therefore, by the definition of $p_{\mathbb{F}}(G)$, if $\|\Delta\| < \min\{r_{\mathbb{F}}(G_{22}), p_{\mathbb{F}}(G)\}$, then $\|\mathscr{F}(G,\Delta)\|_\infty \neq 1$. Since $\|\mathscr{F}(G,\Delta)\|_\infty$ is continuous in Δ when $\|\Delta\| < r_{\mathbb{F}}(G_{22})$, we know that if $\|\Delta\| < \min\{r_{\mathbb{F}}(G_{22}), p_{\mathbb{F}}(G)\}$, then $\mathscr{F}(G,\Delta)$ is internally stable and $\|\mathscr{F}(G,\Delta)\|_\infty < 1$. This shows that $pr_{\mathbb{F}}(G) \geqslant \min\{r_{\mathbb{F}}(G_{22}), p_{\mathbb{F}}(G)\}$.

From the definitions and the above proof, it is easy to see that there exists a $\Delta \in \mathbb{F}^{m\times p}$, with $\|\Delta\| = \min\{r_{\mathbb{F}}(G_{22}), p_{\mathbb{F}}(G)\}$, such that either internal stability is destroyed or $\|\mathscr{F}(G,\Delta)\|_\infty = 1$. This shows $pr_{\mathbb{F}}(G) \leqslant \min\{r_{\mathbb{F}}(G_{22}), p_{\mathbb{F}}(G)\}$. $\square$

The computation of $r_{\mathbb{F}}(G_{22})$ has been studied for a long time and the following formulas are now well known:[1,7,8,10]

$$r_{\mathbb{C}}(G_{22}) = \|G_{22}\|_\infty^{-1} \tag{6}$$

$$r_{\mathbb{R}}(G_{22}) = \left\{\sup_{\omega\in[0,\infty]} \inf_{\gamma\in(0,1]} \sigma_2\left(\begin{bmatrix} \operatorname{Re} G_{22}(j\omega) & -\gamma \operatorname{Im} G_{22}(j\omega) \\ \gamma^{-1}\operatorname{Im} G_{22}(j\omega) & \operatorname{Re} G_{22}(j\omega)\end{bmatrix}\right)\right\}^{-1} \tag{7}$$

Hence, we only need to focus on the computation of $p_{\mathbb{F}}(G)$ in this paper.

For Hermitian matrix $M = \left[\begin{smallmatrix} S & N \\ N^* & R\end{smallmatrix}\right] \in \mathbb{C}^{(m+p)\times(m+p)}$ with $S \geqslant 0$ and $R \geqslant 0$, define

$$\psi_{\mathbb{F}}(M) = \inf\left\{\|\Delta\| : \Delta \in \mathbb{F}^{m\times p}, \quad \det\left\{I - \begin{bmatrix} 0 & \Delta \\ \Delta^* & 0\end{bmatrix} M\right\} = 0\right\} \tag{8}$$

then it follows that $p_{\mathbb{F}}(G) = \inf_{\omega\in[0,\infty]} \psi_{\mathbb{F}}[M(j\omega)]$. Therefore, if $\psi_{\mathbb{F}}(M)$ can be obtained for each $M = \left[\begin{smallmatrix} S & N \\ N^* & R\end{smallmatrix}\right]$ with $S \geqslant 0$ and $R \geqslant 0$, then $pr_{\mathbb{F}}(G)$ can be computed by a frequency sweep.

The next two sections are dedicated to the computation of $\psi_{\mathbb{C}}$ and $\psi_{\mathbb{R}}$ respectively. Before going into the comutation of $\psi_{\mathbb{F}}$, we observe that if G is given in terms of a state space realization:

$$G(s) \simeq \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right] \tag{9}$$

then a formula for $M(j\omega)$ is given by

$$M(j\omega) = \begin{bmatrix} 0 & 0 & B_2^{\mathrm{T}} & D_{12}^{\mathrm{T}} \\ C_2 & D_{21} & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & -j\omega I - A^{\mathrm{T}} & -C_1^{\mathrm{T}} \\ 0 & I & -B_1^{\mathrm{T}} & -D_{11}^{\mathrm{T}} \\ j\omega I - A & -B_1 & 0 & 0 \\ -C_1 & -D_{11} & 0 & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & C_2^{\mathrm{T}} \\ 0 & D_{21}^{\mathrm{T}} \\ B_2 & 0 \\ D_{12} & 0 \end{bmatrix}$$

3. COMPLEX PERFORMANCE RADIUS

It follows from the development in Section 2 that $pr_{\mathbb{C}}(G)$ can be obtained via the computation of $\psi_{\mathbb{C}}(M)$ for each $M = \left[\begin{smallmatrix} S & N \\ N^* & R\end{smallmatrix}\right]$ with $S \geqslant 0$ and $R \geqslant 0$. Hence, this section is dedicated to the computation of $\psi_{\mathbb{C}}(M)$. In the following, we show that $\psi_{\mathbb{C}}(M)$ is the infimum of a convex univariate function.

Note that

$$\det\left\{I-\begin{bmatrix}0 & \Delta\\ \Delta^* & 0\end{bmatrix}M\right\}=\det\left\{I-\begin{bmatrix}\sqrt{\gamma}I & 0\\ 0 & I/\sqrt{\gamma}\end{bmatrix}\begin{bmatrix}0 & \Delta\\ \Delta^* & 0\end{bmatrix}M\begin{bmatrix}\sqrt{\gamma}I & 0\\ 0 & I/\sqrt{\gamma}\end{bmatrix}^{-1}\right\}$$

$$=\det\left\{I-\begin{bmatrix}0 & \Delta\\ \Delta^* & 0\end{bmatrix}\begin{bmatrix}S/\gamma & N\\ N^* & \gamma R\end{bmatrix}\right\} \tag{10}$$

Theorem 1

Let M $= [\begin{smallmatrix}S & N\\ N^* & R\end{smallmatrix}]$ satisfy $S \geqslant 0$ and $R \geqslant 0$. Then

$$\psi_{\mathbb{C}}(M)=\left\{\inf_{\gamma>0}\lambda_1\begin{bmatrix}S/\gamma & N\\ N^* & \gamma R\end{bmatrix}\right\}^{-1}$$

Furthermore, $\lambda_1[\begin{smallmatrix}S/\gamma & N\\ N^* & \gamma R\end{smallmatrix}]$ is a convex function of γ on $(0, \infty)$.

Several lemmas are needed for the proof of Theorem 1.

Lemma 1 (See Reference 11, p. 149)

Let $F(\gamma)\in\mathbb{C}^{n\times n}$ be a Hermitian matrix function analytic on an open set $\Gamma\subset\mathbb{R}$. Then there exist a unitary matrix function $\tilde{V}(\gamma)=[\tilde{v}_1(\gamma),\ldots,\tilde{v}_n(\gamma)]\in\mathbb{C}^{n\times n}$ and a diagonal matrix function $\tilde{\Lambda}(\gamma)=\operatorname{diag}[\tilde{\lambda}_1(\gamma),\ldots,\tilde{\lambda}_n(\gamma)]\in\mathbb{C}^{n\times n}$, both analytic on Γ, such that

$$F(\gamma)=\tilde{V}(\gamma)\tilde{\Lambda}(\gamma)\tilde{V}^*(\gamma)$$

Furthermore,

$$\frac{\mathrm{d}\tilde{\lambda}_i(\gamma)}{\mathrm{d}\gamma}=\tilde{v}_i^*(\gamma)\frac{\mathrm{d}F(\gamma)}{\mathrm{d}\gamma}\tilde{v}_i(\gamma) \tag{11}$$

Lemma 2

Let $F(\gamma)\in\mathbb{C}^{n\times n}$ be a Hermitian matrix function analytic on an open set $\Gamma\subset\mathbb{R}$. Let $\lambda_1(\gamma)\geqslant\lambda_2(\gamma)\geqslant\cdots\geqslant\lambda_n(\gamma)$ be its ordered eigenvalues. If $\lambda_i(\gamma)$ has a local extremum at $\gamma_0\in\Gamma$, then $F(\gamma_0)$ has an eigenvector $v\in\mathbb{C}^n$ corresponding to $\lambda_i(\gamma_0)$ such that $v^*(\mathrm{d}F(\gamma_0)/\mathrm{d}\gamma)v=0$.

Proof. If the multiplicity of $\lambda_i(\gamma_0)$ is one, then $\lambda_i(\gamma)$ is equal to $\tilde{\lambda}_j(\gamma)$ given in Lemma 1 in an open neighbourhood of γ_0. Thus γ_0 is also a stationary point of $\tilde{\lambda}_j(\gamma)$. Let $\tilde{v}_j(\gamma)$ be an analytic eigenvector corresponding to $\tilde{\lambda}_j(\gamma)$. Then (11) gives

$$\tilde{v}_j^*(\gamma_0)\frac{\mathrm{d}F(\gamma_0)}{\mathrm{d}\gamma}\tilde{v}_j(\gamma_0)=0$$

If instead the multiplicity of $\lambda_i(\gamma_0)$ is greater than one, then we can assume, without loss of generality, that in an open neighbourhood of γ_0, $\lambda_i(\gamma)=\tilde{\lambda}_{j1}(\gamma)$ for $\gamma\leqslant\gamma_0$ and $\lambda_i(\gamma)=\tilde{\lambda}_{j2}(\gamma)$ for $\gamma\geqslant\gamma_0$. If $j1=j2$, then $\lambda_i(\gamma_0)$ must be a local extremum of $\tilde{\lambda}_{j1}(\gamma)$, so we get the result by applying

(11). Otherwise let $\tilde{v}_k(\gamma)$, $k = j1, j2$, be the analytic eigenvectors of $F(\gamma)$ corresponding to $\tilde{\lambda}_k(\gamma)$. Then (11) gives

$$\frac{\mathrm{d}\tilde{\lambda}_{j1}(\gamma_0)}{\mathrm{d}\gamma} = \tilde{v}^*_{j1}(\gamma_0)\frac{\mathrm{d}F(\gamma_0)}{\mathrm{d}\gamma}\tilde{v}_{j1}(\gamma_0)$$

$$\frac{\mathrm{d}\tilde{\lambda}_{j2}(\gamma_0)}{\mathrm{d}\gamma} = \tilde{v}^*_{j2}(\gamma_0)\frac{\mathrm{d}F(\gamma_0)}{\mathrm{d}\gamma}\tilde{v}_{j2}(\gamma_0)$$

Put $v_\alpha = \alpha\tilde{v}_{j1} + (1-\alpha^2)^{1/2}\tilde{v}_{j2}$ for $\alpha \in [0, 1]$. Then $v_\alpha(\gamma_0)$ is also a unit length eigenvector of $F(\gamma_0)$ corresponding to $\lambda_i(\gamma_0)$. Define

$$f(\alpha) = v^*_\alpha(\gamma_0)\frac{\mathrm{d}F(\gamma_0)}{\mathrm{d}\gamma}v_\alpha(\gamma_0)$$

Since γ_0 is a local extremum of $\lambda_i(\gamma)$, we must have $f(0)f(1) = (\mathrm{d}\tilde{\lambda}_{j1}(\gamma_0)/\mathrm{d}\gamma)(\mathrm{d}\tilde{\lambda}_{j2}(\gamma_0)/\mathrm{d}\gamma) \leqslant 0$. By continuity, $f(\alpha) = 0$ has a solution in $[0, 1]$. This proves the lemma. □

Lemma 3

Let $F = \left[\begin{smallmatrix} X & Y \\ Y^* & Z \end{smallmatrix}\right] \in \mathbb{C}^{(m+p)\times(m+p)}$ be a Hermitian matrix. Suppose $X \geqslant 0$ and $Z \geqslant 0$, then for i $= 1, 2, \ldots, \min\{m, p\}$,

$$\lambda_i(F) \geqslant -\lambda_{m+p-i+1}(F)$$

Proof. Let $F_1 = \left[\begin{smallmatrix} 0 & Y \\ Y^* & 0 \end{smallmatrix}\right]$ and $F_2 = \left[\begin{smallmatrix} X & 0 \\ 0 & Z \end{smallmatrix}\right]$. Then $F = F_1 + F_2$ and for $i = 1, 2, \ldots, \min\{m, p\}$,

$$\lambda_i(F_1) = -\lambda_{m+p-i+1}(F_1) = \sigma_i(Y)$$

Since $F_2 \geqslant 0$, it follows that $\lambda_i(F) \geqslant \lambda_i(F_1)$ for $i = 1, 2, \ldots, m+p$. Hence, for $i = 1, 2, \ldots, \min\{m, p\}$,

$$\lambda_i(F) \geqslant \lambda_i(F_1) = -\lambda_{m+p-i+1}(F_1) \geqslant -\lambda_{m+p-i+1}(F)$$ □

Proof of Theorem 1. Denote $F(\gamma) = \left[\begin{smallmatrix} S/\gamma & N \\ N^* & \gamma R \end{smallmatrix}\right]$. Then from Lemma 3, $\lambda_1[F(\gamma)] \geqslant -\lambda_{m+p}[F(\gamma)]$ for all $\gamma > 0$, hence $\lambda_1[F(\gamma)] = \bar{\sigma}[F(\gamma)]$. We see from (10) that

$$\psi_{\mathbb{C}}(M) \geqslant \left\{\inf_{\gamma>0} \lambda_1[F(\lambda)]\right\}^{-1}$$

The fact that $\bar{\sigma}[F(\gamma)]$ is convex follows from Reference 12.

Let $\lambda_0 = \inf_{\gamma>0}\lambda_1[F(\gamma)]$. The rest of the proof is to show that $\psi_{\mathbb{C}}(M) \leqslant \lambda_0^{-1}$.

If $\lambda_0 = 0$, we certainly have $\psi_{\mathbb{C}}(M) \leqslant \lambda_0^{-1}$. So we assume that $\lambda_0 > 0$. There are two cases.

Case 1: $\inf_{\gamma>0}\lambda_1[F(\gamma)]$ is attained at $\gamma_0 \in (0, \infty)$.

By Lemma 2, there exist $v = \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$, $x \in \mathbb{C}^m$ and $y \in \mathbb{C}^p$ such that $F(\gamma_0)v = \lambda_0 v$, i.e.,

$$\frac{1}{\gamma_0}Sx + Ny = \lambda_0 x \tag{12}$$

$$N^*x + \gamma_0 Ry = \lambda_0 y \tag{13}$$

and

$$v^* \frac{\mathrm{d}F(\gamma_0)}{\mathrm{d}\gamma} v = -\frac{1}{\gamma_0^2} x^*Sx + y^*Ry = 0$$

Multiplying (12) and (13) from the left by x^* and y^* respectively, subtracting the resulting equations, and noting that x^*Ny must be real, we get

$$\lambda_0(x^*x - y^*y) = \frac{1}{\gamma_0} x^*Sx - \gamma_0 y^*Ry = 0$$

Hence, $x^*x = y^*y$. Now construct $\Delta = (\lambda_0)^{-1}xy^*/x^*x$, then it is easy to verify that $\|\Delta\| = (\lambda_0)^{-1}$, and

$$\left\{ I - \begin{bmatrix} 0 & \Delta \\ \Delta^* & 0 \end{bmatrix} \begin{bmatrix} S/\gamma_0 & N \\ N^* & \gamma_0 R \end{bmatrix} \right\} \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

which means that $\det\{I - [\begin{smallmatrix} 0 & \Delta \\ \Delta^* & 0 \end{smallmatrix}]M\} = 0$. From the definition of $\psi_{\mathbb{C}}(M)$, we have $\psi_{\mathbb{C}}(M) \leqslant (\lambda_0)^{-1}$.

Case 2: $\lambda_0 = \lim_{\gamma \to 0} \lambda_1[F(\gamma)]$ or $\lambda_0 = \lim_{\gamma \to \infty} \lambda_1[F(\gamma)]$.
This case occurs only when $S = 0$ or $R = 0$. Assume $S = 0$, then

$$\det \left\{ I - \begin{bmatrix} 0 & \Delta \\ \Delta^* & 0 \end{bmatrix} M \right\} = \det(I - \Delta N^*) \det(I - \Delta^* N)$$

which shows $\psi_{\mathbb{C}}(M) = [\bar{\sigma}(N)]^{-1}$.

On the other hand, since $R \geqslant 0$, we have

$$\lambda_0 = \inf_{\gamma > 0} \lambda_1 \begin{bmatrix} 0 & N \\ N^* & \gamma R \end{bmatrix} = \bar{\sigma}(N)$$

Therefore, $\psi_{\mathbb{C}}(M) = \lambda_0^{-1}$ in this case. □

4. A LOWER BOUND OF THE REAL PERFORMANCE RADIUS

Recall from Section 2 that $pr_{\mathbb{R}}(G) = \min\{r_{\mathbb{R}}(G_{22}), p_{\mathbb{R}}(G)\}$. The computation of $r_{\mathbb{R}}(G_{22})$ can be done using formula (7). The computation of $p_{\mathbb{R}}(G)$ depends on the computation of $\psi_{\mathbb{R}}(M)$ for a given Hermitian matrix $M = [\begin{smallmatrix} S & N \\ N^* & R \end{smallmatrix}]$ with $S \geqslant 0$ and $R \geqslant 0$. Unfortunately, we are not able to compute $\psi_{\mathbb{R}}(M)$ at this moment. In this section, a lower bound of $\psi_{\mathbb{R}}(M)$ will be given.

Let

$$S_r = \operatorname{Re} S, \quad S_i = \operatorname{Im} S, \quad R_r = \operatorname{Re} R, \quad R_i = \operatorname{Im} R, \quad N_r = \operatorname{Re} N, \quad N_i = \operatorname{Im} N$$

Then

$$S_r = S_r^{\mathrm{T}}, \quad R_r = R_r^{\mathrm{T}}, \quad S_i = -S_i^{\mathrm{T}}, \quad R_i = -R_i^{\mathrm{T}} \tag{14}$$

and $S_r \geqslant 0$, $R_r \geqslant 0$.

We follow the idea in Reference 10 to convert the problem of computing $\psi_{\mathbb{R}}(M)$, one with complex data and realness constraint, into a pure real problem. Let

$$P_0 = \begin{bmatrix} \operatorname{Re} M & -\operatorname{Im} M \\ \operatorname{Im} M & \operatorname{Re} M \end{bmatrix} = \begin{bmatrix} S_r & N_r & -S_i & -N_i \\ N_r^{\mathrm{T}} & R_r & N_i^{\mathrm{T}} & -R_i \\ S_i & N_i & S_r & N_r \\ -N_i^{\mathrm{T}} & R_i & N_r^{\mathrm{T}} & R_r \end{bmatrix}, \quad \Delta_a = \begin{bmatrix} 0 & \Delta & 0 & 0 \\ \Delta^{\mathrm{T}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \Delta \\ 0 & 0 & \Delta^{\mathrm{T}} & 0 \end{bmatrix}$$

The following lemma will be frequently used.

Lemma 4

Given Hermitian matrix $Z = X + jY$ with $X, Y \in \mathbb{R}^{n \times n}$, let $Q = [\begin{smallmatrix} X & -Y \\ Y & X \end{smallmatrix}]$. Then $\lambda_{2i-1}(Q)$, $\lambda_{2i}(Q) = \lambda_i(Z)$.

Proof. This follows since Q is similar to $[\begin{smallmatrix} X+jY & 0 \\ 0 & X-jY \end{smallmatrix}]$, and $X + jY$, $X - jY$ share the same eigenvalues. □

For $\alpha, \beta > 0$, define scaling matrix

$$D(\alpha,\beta) = \operatorname{diag}\left[\sqrt{\alpha\beta} I,\ \frac{1}{\sqrt{\alpha\beta}} I,\ \sqrt{\frac{\beta}{\alpha}} I,\ \sqrt{\frac{\alpha}{\beta}} I\right] \tag{15}$$

then $D^{-1}(\alpha,\beta)\Delta_a D^{-1}(\alpha,\beta) = \Delta_a$ and

$$P(\alpha,\beta) := D(\alpha,\beta) P_0 D(\alpha,\beta) = \begin{bmatrix} \alpha\beta S_r & N_r & -\beta S_i & -\alpha N_i \\ N_r^{\mathrm{T}} & \frac{1}{\alpha\beta} R_r & \frac{1}{\alpha} N_i^{\mathrm{T}} & -\frac{1}{\beta} R_i \\ \beta S_i & \frac{1}{\alpha} N_i & \frac{\beta}{\alpha} S_r & N_r \\ -\alpha N_i^{\mathrm{T}} & \frac{1}{\beta} R_i & N_r^{\mathrm{T}} & \frac{\alpha}{\beta} R_r \end{bmatrix} \tag{16}$$

Assume $P_0 = P(1,1)$ has π positive and v negative eigenvalues. Lemma 4 says that π and v are even. Also observe that $\pi \geqslant v$ which follows from applying Lemma 3 and Lemma 4 to a similarly permuted version of P_0:

$$\begin{bmatrix} S_r & -S_i & N_r & -N_i \\ S_i & S_r & N_i & N_r \\ N_r^{\mathrm{T}} & N_i^{\mathrm{T}} & R_r & -R_i \\ -N_i^{\mathrm{T}} & N_r^{\mathrm{T}} & R_i & R_r \end{bmatrix}$$

By the law of inertia, $P(\alpha,\beta)$ also has π positive and v negative eigenvalues for all $\alpha, \beta > 0$ and its second eigenvalue is always non-negative.

Another interesting property of $P(\alpha,\beta)$ is that

$$\lambda_i[P(\alpha,\beta)] = \lambda_i\left[P\left(\frac{1}{\alpha}, \beta\right)\right] \tag{17}$$

for $i = 1, 2, \ldots, 2(m + p)$, $\alpha, \beta \neq 0$. This is because $P(\alpha, \beta)$ is similar to

$$\begin{bmatrix} \alpha\beta S_r & \alpha N_i & -\beta S_i & N_r \\ \alpha N_i^{\mathrm{T}} & \frac{\alpha}{\beta} R_r & -N_r^{\mathrm{T}} & -\frac{1}{\beta} R_i \\ \beta S_i & -N_r & \frac{\beta}{\alpha} S_r & \frac{1}{\alpha} N_i \\ N_r^{\mathrm{T}} & \frac{1}{\beta} R_i & \frac{1}{\alpha} N_i^{\mathrm{T}} & \frac{1}{\alpha\beta} R_r \end{bmatrix} \tag{18}$$

and

$$\begin{bmatrix} \beta S_i & -N_r \\ N_r^{\mathrm{T}} & \frac{1}{\beta} R_i \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} -\beta S_i & N_r \\ -N_r^{\mathrm{T}} & -\frac{1}{\beta} R_i \end{bmatrix}.$$

Theorem 2

$\psi_{\mathbb{R}}(M) \geqslant \{\inf_{\alpha \in (0,1], \beta > 0} \lambda_2[P(\alpha, \beta)]\}^{-1}$.

To prove Theorem 2, we need two lemmas.

Lemma 5 (See Reference 13, p. 203)

Let $M, \Delta \in \mathbb{C}^{n \times n}$ be Hermitian matrices. Denote the eigenvalues of M as $\eta_1 \geqslant \eta_2 \geqslant \cdots \geqslant \eta_n$ and the eignevalues of $M + \Delta$ as $\xi_1 \geqslant \xi_2 \geqslant \cdots \geqslant \xi_n$. Then

$$|\xi_i - \eta_i| \leqslant \bar{\sigma}(\Delta)$$

Lemma 6

For $A \in \mathbb{C}^{m \times p}$ and $B \in \mathbb{C}^{p \times m}$,

$$m - \operatorname{rank}(I + AB) = p - \operatorname{rank}(I + BA)$$

Proof. It can be verified that

$$\begin{bmatrix} I_m & -A \\ 0 & I_p \end{bmatrix} \begin{bmatrix} I_m + AB & 0 \\ B & I_p \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ B & I_p + BA \end{bmatrix} \begin{bmatrix} I_m & -A \\ 0 & I_p \end{bmatrix}$$

Hence,

$$\operatorname{rank} \begin{bmatrix} I_m + AB & 0 \\ B & I_p \end{bmatrix} = \operatorname{rank} \begin{bmatrix} I_m & 0 \\ B & I_p + BA \end{bmatrix}$$

which implies that rank $(I_m + AB) + p = \operatorname{rank}(I_p + BA) + m$. The lemma then follows. □

Proof of Theorem 2. It follows from (17) that

$$\inf_{\alpha \in (0,1], \beta > 0} \lambda_2[P(\alpha, \beta)] = \inf_{\alpha, \beta > 0} \lambda_2[P(\alpha, \beta)]$$

Since rank $(P_0) = \pi + \nu$, it follows from Lemma 4 that $\operatorname{rank}(M) = (\pi + \nu)/2$. Suppose that M is decomposed as $M = U_M \Lambda_M U_M^*$, where $U_M^* U_M = I$, $U_M \in \mathbb{C}^{(m+p) \times ((\pi+\nu)/2)}$, and $\Lambda_M \in \mathbb{R}^{((\pi+\nu)/2) \times ((\pi+\nu)/2)}$ is diagonal and non-singular. Then

$$P_0 = \begin{bmatrix} \operatorname{Re} M & -\operatorname{Im} M \\ \operatorname{Im} M & \operatorname{Re} M \end{bmatrix} = U \Lambda U^{\mathrm{T}}$$

where

$$U = \begin{bmatrix} \mathrm{Re}\, U_M & -\mathrm{Im}\, U_M \\ \mathrm{Im}\, U_M & \mathrm{Re}\, U_M \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \Lambda_M & 0 \\ 0 & \Lambda_M \end{bmatrix}$$

Let $X(\alpha,\beta) = D(\alpha,\beta)U$, then $X(\alpha,\beta)$ has full column rank. Carry out the Gram–Schmidt orthonormalization to the columns of $X(\alpha,\beta)$, we get $X(\alpha,\beta) = V(\alpha,\beta)R(\alpha,\beta)$, where $V^{\mathrm{T}}(\alpha,\beta)V(\alpha,\beta) = I$ and $R(\alpha,\beta)$ is non-singular. It is easy to see from the orthonormalization process that the maps from X to V and R are analytic when X has full column rank. Hence, $V(\alpha,\beta)$ and $R(\alpha,\beta)$ are analytic in (α,β).

Let $E(\alpha,\beta) = R(\alpha,\beta)\Lambda R^{\mathrm{T}}(\alpha,\beta)$, then $E(\alpha,\beta)$ is analytic. From $P_0 = U\Lambda U^{\mathrm{T}}$, we get

$$P(\alpha,\beta) = D(\alpha,\beta)P_0 D^{\mathrm{T}}(\alpha,\beta) = V(\alpha,\beta)E(\alpha,\beta)V^{\mathrm{T}}(\alpha,\beta) \tag{19}$$

From $D(1,1) = I$, we get $V(1,1) = U$ and $E(1,1) = \Lambda$.

Since $V(\alpha,\beta)$ is orthonormal, the eigenvalues of $E(\alpha,\beta)$ are equal to the non-zero eigenvalues of $P(\alpha,\beta)$.

Note that

$$\begin{aligned} \mathrm{rank}\,(I - \Delta_a P_0) &= \mathrm{rank}\,\{I - \Delta_a P(\alpha,\beta)\} \\ &= \mathrm{rank}\,\{I - \Delta_a V(\alpha,\beta)E(\alpha,\beta)V^{\mathrm{T}}(\alpha,\beta)\} \end{aligned}$$

By Lemma 6, we obtain

$$\begin{aligned} 2(m+p) - \mathrm{rank}(I - \Delta_a P_0) &= \pi + v - \mathrm{rank}\,\{I - V^{\mathrm{T}}(\alpha,\beta)\Delta_a V(\alpha,\beta)E(\alpha,\beta)\} \\ &= \pi + v - \mathrm{rank}\,\{E^{-1}(\alpha,\beta) - V^{\mathrm{T}}(\alpha,\beta)\Delta_a V(\alpha,\beta)\} \end{aligned} \tag{20}$$

Denote

$$H(\alpha,\beta,\Delta) = E^{-1}(\alpha,\beta) - V^{\mathrm{T}}(\alpha,\beta)\Delta_a V(\alpha,\beta) \tag{21}$$

then by (20), $\mathrm{rank}[H(\alpha,\beta,\Delta)]$ is independent of (α,β), and

$$\mathrm{rank}[H(\alpha,\beta,\Delta)] = \pi + v - \{2(m+p) - \mathrm{rank}(I - \Delta_a P_0)\} \tag{22}$$

In the following, we will show that if $\bar{\sigma}(\Delta) < \{\inf_{\alpha,\beta>0} \lambda_2[P(\alpha,\beta)]\}^{-1}$, then $\mathrm{rank}[H(\alpha,\beta,\Delta)] = \pi + v$, which leads to $\psi_{\mathbb{R}}(M) \geqslant \{\inf_{\alpha,\beta>0} \lambda_2[P(\alpha,\beta)]\}^{-1}$ by (22).

Since $E(\alpha,\beta)$ is analytic and non-singular and $V(\alpha,\beta)$ is analytic, it follows that for a fixed Δ, the eigenvalues of $H(\alpha,\beta,\Delta)$ are continuous in (α,β). Since the rank of $H(\alpha,\beta,\Delta)$ is independent of (α,β), we conclude that for a fixed Δ, the inertia of $H(\alpha,\beta,\Delta)$ are independent of (α,β). Consequently, it can be denoted by $\{\pi_\Delta, v_\Delta, \zeta_\Delta\}$. Furthermore, π_Δ, v_Δ and ζ_Δ are even numbers, since by Lemma 4 the eigenvalues of

$$H(1,1,\Delta) = \begin{bmatrix} \Lambda_M^{-1} & 0 \\ 0 & \Lambda_M^{-1} \end{bmatrix} - \begin{bmatrix} \mathrm{Re}\, U_M & -\mathrm{Im}\, U_M \\ \mathrm{Im}\, U_M & \mathrm{Re}\, U_M \end{bmatrix}^{\mathrm{T}} \Delta_a \begin{bmatrix} \mathrm{Re}\, U_M & -\mathrm{Im}\, U_M \\ \mathrm{Im}\, U_M & \mathrm{Re}\, U_M \end{bmatrix}$$

have even multiplicity.

The eigenvalues of $H(\alpha,\beta,0) = E^{-1}(\alpha,\beta)$, which are $\lambda_i^{-1}[P(\alpha,\beta)]$, $i = 1,\ldots,\pi$ and $i = 2(m+p) - v + 1, \ldots, 2(m+p)$, satisfy

$$\cdots \geqslant \lambda_2^{-1}[P(\alpha,\beta)] \geqslant \lambda_1^{-1}[P(\alpha,\beta)] > 0 > \lambda_{2(m+p)}^{-1}[P(\alpha,\beta)] \geqslant \lambda_{2(m+p)-1}^{-1}[P(\alpha,\beta)] \geqslant \cdots$$

Now consider the inertia of $H(\alpha,\beta,\Delta)$ under perturbation. If $\|\bar{\sigma}(\Delta)\| < \{\inf_{\alpha,\beta>0}\lambda_2[P(\alpha,\beta)]\}^{-1}$, there exists α_0,β_0 such that $\bar{\sigma}(\Delta) < \lambda_2^{-1}[P(\alpha_0,\beta_0)]$. Note that the eigenvalues of $P(\alpha,\beta)$ are the same as those of

$$\begin{bmatrix} \alpha\beta S_r & -\beta S_i & N_r & -\alpha N_i \\ \beta S_i & \frac{\beta}{\alpha}S_r & \frac{1}{\alpha}N_i & N_r \\ N_r^{\mathrm{T}} & \frac{1}{\alpha}N_i^{\mathrm{T}} & \frac{1}{\alpha\beta}R_r & -\frac{1}{\beta}R_i \\ -\alpha N_i^{\mathrm{T}} & N_r^{\mathrm{T}} & \frac{1}{\beta}R_i & \frac{\alpha}{\beta}R_r \end{bmatrix} \tag{23}$$

Hence, if $v \geqslant 2$, then Lemma 3 implies that $\bar{\sigma}(\Delta) < -\lambda_{2(m+p)-1}^{-1}[P(\alpha_0,\beta_0)]$. Since $\bar{\sigma}[V^{\mathrm{T}}(\alpha_0,\beta_0)\Delta_a V(\alpha_0,\beta_0)] \leqslant \bar{\sigma}(\Delta)$, it follows from Lemma 5 that $\pi_\Delta > \pi - 2$ and $v_\Delta > v - 2$. Since π_Δ and v_Δ are even numbers, we must have $\pi_\Delta = \pi$, $v_\Delta = v$. This is also true if $v = 0$. Therefore if $\|\bar{\sigma}(\Delta)\| < \{\inf_{\alpha,\beta>0}\lambda_2[P(\alpha,\beta)]\}^{-1}$, then $\operatorname{rank}[H(\alpha,\beta,\Delta)] = \pi + v$, which shows $\operatorname{rank}(I - \Delta_a P_0) = 2(m+p)$. Then it follows from Lemma 4 and the definition of $\psi_{\mathbb{R}}(M)$ that $\psi_{\mathbb{R}}(M) \geqslant \{\inf_{\alpha,\beta>0}\lambda_2[P(\alpha,\beta)]\}^{-1}$. □

Remark 1

More general scaling matrices D than the one used in (15) can be used so that $D^{-1}\Delta_a D^{-1} = \Delta_a$ and DP_0D is symmetric. For example,

$$D_1(\alpha,\beta) = \operatorname{diag}\left[\alpha I,\ \frac{1}{\alpha}I,\ \beta I,\ \frac{1}{\beta}I\right]$$

and

$$D_2(\alpha,\beta,\gamma) = \begin{bmatrix} \alpha I & 0 & \gamma I & 0 \\ 0 & \alpha_1 I & 0 & \gamma_1 I \\ \gamma I & 0 & \beta I & 0 \\ 0 & \gamma_1 I & 0 & \beta_1 I \end{bmatrix}, \quad \text{where } \begin{bmatrix} \alpha_1 & \gamma_1 \\ \gamma_1 & \beta_1 \end{bmatrix} = \begin{bmatrix} \alpha & \gamma \\ \gamma & \beta \end{bmatrix}^{-1}$$

It is easy to see that $D(\alpha,\beta) = D_1(\sqrt{\alpha\beta}, \sqrt{\beta/\alpha})$ and one can verify that $D_2(\alpha,\beta,\gamma)P_0D_2(\alpha,\beta,\gamma)$ is similar to $D_1(\lambda_1,\lambda_2)P_0D_1(\lambda_1,\lambda_2)$, where λ_1 and λ_2 are the eigenvalues of $[\begin{smallmatrix}\alpha & \gamma \\ \gamma & \beta\end{smallmatrix}]$. So these scaling matrices are equivalent to (15) as conservatism reduction is concerned.

Remark 2

Let $\sigma_0 = \inf_{\alpha,\beta>0}\sigma_2[P(\alpha,\beta)]$, then it is easy to see from the above proof that another lower bound for $\psi_{\mathbb{R}}(M)$ is σ_0^{-1}. It is obvious that $\sigma_0 \geqslant \inf_{\alpha,\beta>0}\lambda_2[P(\alpha,\beta)]$. If $\sigma_0 = \inf_{\alpha,\beta>0}\sigma_2[P(\alpha,\beta)]$ is obtained at (α_0,β_0), and σ_0 is a singular value of $P(\alpha_0,\beta_0)$ with multiplicity greater than one, then σ_0^{-1} may be strictly smaller than $\{\inf_{\alpha,\beta>0}\lambda_2[P(\alpha,\beta)]\}^{-1}$, if this multiplicity is caused by the intersection of $\lambda_2[P(\alpha,\beta)]$ and $-\lambda_{2(m+p)}[P(\alpha,\beta)]$, as is illustrated in Section 6.

5. PROPERTIES OF THE LOWER BOUND

A lower bound of $\psi_{\mathbb{R}}(M)$, given in terms of a two-dimensional nonlinear minimization problem, is derived in the last section. Two questions remain to be answered:

1. How tight is the lower bound?
2. Is the minimization problem numerically tractable?

The first question is partially answered in this section; we give sufficient conditions under which the lower bound is tight (Theorem 3). The second issue is completely answered; we show that the function $\lambda_2[P(\alpha,\beta)]$, which is to be minimized in the lower bound, is unimodal in the area $(0,1]\times(0,\infty)$. The definition of a multivariate function being unimodal is non-standard. We adopt the following definition: a real valued function is said to be unimodal if the inverse image of $(-\infty, y)$ is connected for all $y\in\mathbb{R}$.

Theorem 3

If $\lambda_2[P(\alpha,\beta)]$ has a local minimum λ_0 at (α_0,β_0) with $(\alpha_0,\beta_0)\in(0,1)\times(0,\infty)$ and if λ_0 is a simple eigenvalue of $P(\alpha_0,\beta_0)$, then $\psi_{\mathbb{R}}(M)=\lambda_0^{-1}$.

Theorem 4

$\lambda_2[P(\alpha,\beta)]$ is unimodal in the area $\alpha\in(0,1]$, $\beta>0$.

Numerical experience shows that the condition in Theorem 3 is often satisfied. Numerical examples also show that $\lambda_2[P(\alpha,\beta)]$ has a rather weak convexity property: it is not quasi-convex. For fixed α_0, the univariate function $\lambda_2[P(\alpha_0,\beta)]$ is not unimodal in general. Nevertheless, it is enough that $\lambda_2[P(\alpha,\beta)]$ is unimodal. Many non-smooth local optimization methods can be applied to find the global infimum of $\lambda_2[P(\alpha,\beta)]$. Numerical experience shows that the simplex method works quite well.

To prove Theorem 3, we need the following lemma.

Lemma 7

Let $U\in\mathbb{R}^{p\times k}$ and $V\in\mathbb{R}^{m\times k}$. If $U^{\mathrm{T}}U=V^{\mathrm{T}}V\neq 0$, then

(1) $\bar{\sigma}[VU^{\dagger}]=1$,
(2) $VU^{\dagger}U=V$,
(3) $(U^{\dagger})^{\mathrm{T}}V^{\mathrm{T}}V=U$.

Proof. (1) and (2) are directly from Lemma 2 of Reference 10. Now, we prove (3). Since $V^{\mathrm{T}}V=U^{\mathrm{T}}U$, so,

$$(U^{\dagger})^{\mathrm{T}}V^{\mathrm{T}}V=(U^{\dagger})^{\mathrm{T}}U^{\mathrm{T}}U=U$$

Here, the last equality is obtained by applying the definition of Moore–Penrose inverse directly. □

Proof of Theorem 3. Since λ_0 is a simple eigenvalue of $P(\alpha_0,\beta_0)$, it follows from Reference 13, p. 185, Corollary 2.4 that $\lambda_2[P(\alpha,\beta)]$ is analytic in a neighbourhood of (α_0,β_0). Denote the eigenvector of $P(\alpha_0,\beta_0)$ corresponding to λ_0 as v, then $P(\alpha_0,\beta_0)v=\lambda_0 v$.

Since $\lambda_0=\lambda_2[P(\alpha_0,\beta_0)]$ is a local minimum, we must have

$$\frac{\partial\lambda_2[P(\alpha_0,\beta_0)]}{\partial\alpha}=v^{\mathrm{T}}\frac{\partial P(\alpha_0,\beta_0)}{\partial\alpha}v=0$$

$$\frac{\partial\lambda_2[P(\alpha_0,\beta_0)]}{\partial\beta}=v^{\mathrm{T}}\frac{\partial P(\alpha_0,\beta_0)}{\partial\beta}v=0$$

Now partition v in accordance with $P(\alpha,\beta)$ as $v=[v_1^{\mathrm{T}}\ v_2^{\mathrm{T}}\ v_3^{\mathrm{T}}\ v_4^{\mathrm{T}}]^{\mathrm{T}}$, where $v_1, v_3 \in \mathbb{R}^m$ and $v_2, v_4 \in \mathbb{R}^p$. Then we will first show that

$$[v_1\ \ v_3]^{\mathrm{T}}[v_1\ \ v_3] = [v_2\ \ v_4]^{\mathrm{T}}[v_2\ \ v_4] \tag{24}$$

When this is satisfied, a special Δ can be constructed as $\Delta = \lambda_0^{-1}[v_1\ v_3][v_2\ v_4]^{\dagger}$, with $\bar{\sigma}(\Delta) = \lambda_0^{-1}$ and $\det\{I - [\begin{smallmatrix} 0 & \Delta \\ \Delta^{\mathrm{T}} & 0\end{smallmatrix}][\begin{smallmatrix} S & N \\ N^* & R\end{smallmatrix}]\} = 0$, which leads to $\psi_{\mathbb{R}}(M) \leqslant \lambda_0^{-1}$.

Note that

$$\begin{aligned}
\frac{\partial\lambda_2[P(\alpha_0,\beta_0)]}{\partial\alpha} &= v^{\mathrm{T}}\frac{\partial P(\alpha_0,\beta_0)}{\partial\alpha}v = 2v^{\mathrm{T}}\frac{\partial D(\alpha_0,\beta_0)}{\partial\alpha}P_0 D(\alpha_0,\beta_0)v \\
&= \frac{1}{\alpha_0}v^{\mathrm{T}}\begin{bmatrix} I & 0 & 0 & 0 \\ 0 & -I & 0 & 0 \\ 0 & 0 & -I & 0 \\ 0 & 0 & 0 & I \end{bmatrix} D(\alpha_0,\beta_0)P_0D(\alpha_0,\beta_0)v \\
&= \frac{\lambda_0}{\alpha_0}(v_1^{\mathrm{T}}v_1 - v_2^{\mathrm{T}}v_2 - v_3^{\mathrm{T}}v_3 + v_4^{\mathrm{T}}v_4)
\end{aligned}$$

Similarly,

$$\frac{\partial\lambda_2[P(\alpha_0,\beta_0)]}{\partial\beta} = \frac{\lambda_0}{\beta_0}(v_1^{\mathrm{T}}v_1 - v_2^{\mathrm{T}}v_2 + v_3^{\mathrm{T}}v_3 - v_4^{\mathrm{T}}v_4)$$

Since $\partial\lambda_2[P(\alpha_0,\beta_0)]/\partial\alpha = \partial\lambda_2[P(\alpha_0,\beta_0)]/\partial\beta = 0$, we have $v_1^{\mathrm{T}}v_1 - v_2^{\mathrm{T}}v_2 = v_3^{\mathrm{T}}v_3 - v_4^{\mathrm{T}}v_4$ and $v_1^{\mathrm{T}}v_1 - v_2^{\mathrm{T}}v_2 = -(v_3^{\mathrm{T}}v_3 - v_4^{\mathrm{T}}v_4)$, which result in

$$v_1^{\mathrm{T}}v_1 = v_2^{\mathrm{T}}v_2 \quad \text{and} \quad v_3^{\mathrm{T}}v_3 = v_4^{\mathrm{T}}v_4 \tag{25}$$

From $P(\alpha_0,\beta_0)v = \lambda_0 v$, i.e.,

$$\begin{bmatrix} \alpha_0\beta_0 S_r & N_r & -\beta_0 S_i & -\alpha_0 N_i \\ N_r^{\mathrm{T}} & \frac{1}{\alpha_0\beta_0}R_r & \frac{1}{\alpha_0}N_i^{\mathrm{T}} & -\frac{1}{\beta_0}R_i \\ \beta_0 S_i & \frac{1}{\alpha_0}N_i & \frac{\beta_0}{\alpha_0}S_r & N_r \\ -\alpha_0 N_i^{\mathrm{T}} & \frac{1}{\beta_0}R_i & N_r^{\mathrm{T}} & \frac{\alpha_0}{\beta_0}R_r \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \lambda_0\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}$$

we obtain

$$\begin{aligned}
2\lambda_0(v_3^{\mathrm{T}}v_1 - v_4^{\mathrm{T}}v_2) &= [v_3^{\mathrm{T}}\ \ -v_4^{\mathrm{T}}\ \ v_1^{\mathrm{T}}\ \ -v_2^{\mathrm{T}}]^{\mathrm{T}}P(\alpha_0,\beta_0)v \\
&= \left(\alpha_0 + \frac{1}{\alpha_0}\right)\left(\beta_0 v_3^{\mathrm{T}}S_r v_1 - \frac{1}{\beta_0}v_2^{\mathrm{T}}R_r v_4 - v_4^{\mathrm{T}}N_i^{\mathrm{T}}v_3 + v_1^{\mathrm{T}}N_i v_2\right)
\end{aligned} \tag{26}$$

and

$$\begin{aligned}
0 &= [v_3^{\mathrm{T}}\ \ v_4^{\mathrm{T}}\ \ -v_1^{\mathrm{T}}\ \ -v_2^{\mathrm{T}}]^{\mathrm{T}}P(\alpha_0,\beta_0)v \\
&= \left(\alpha_0 - \frac{1}{\alpha_0}\right)\left(\beta_0 v_3^{\mathrm{T}}S_r v_1 - \frac{1}{\beta_0}v_2^{\mathrm{T}}R_r v_4 - v_4^{\mathrm{T}}N_i^{\mathrm{T}}v_3 + v_1^{\mathrm{T}}N_i v_2\right)
\end{aligned} \tag{27}$$

Here we use the fact that $v_4^{\mathrm{T}}R_i v_4 = v_2^{\mathrm{T}}R_i v_2 = v_1^{\mathrm{T}}S_i v_1 = v_3^{\mathrm{T}}S_i v_3 = 0$.

Since $\lambda_0 \neq 0$ and $\alpha_0 \neq 1$, from (26) and (27), we get $v_3^{\mathrm{T}} v_1 = v_4^{\mathrm{T}} v_2$. Combining with (25), we get (24).

Now construct Δ as $\Delta = \lambda_0^{-1}[v_1 \;\; v_3][v_2 \;\; v_4]^{\dagger}$, then from Lemma 7, $\bar{\sigma}(\Delta) = \lambda_0^{-1}$ and

$$\lambda_0 \Delta [v_2 \;\; v_4] = [v_1 \;\; v_3][v_2 \;\; v_4]^{\dagger}[v_2 \;\; v_4] = [v_1 \;\; v_3]$$

$$\lambda_0 \Delta^{\mathrm{T}} [v_1 \;\; v_3] = ([v_2 \;\; v_4]^{\dagger})^{\mathrm{T}}[v_1 \;\; v_3]^{\mathrm{T}}[v_1 \;\; v_3] = [v_2 \;\; v_4]$$

Therefore,

$$[I - \Delta_a P(\alpha_0, \beta_0)]v = v - \lambda_0 \Delta_a v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} - \lambda_0 \begin{bmatrix} 0 & \Delta & 0 & 0 \\ \Delta^{\mathrm{T}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \Delta \\ 0 & 0 & \Delta^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = 0$$

and

$$D^{-1}(\alpha_0, \beta_0)(I - \Delta_a P_0) D(\alpha_0, \beta_0) v = 0 \tag{28}$$

Partition $D(\alpha_0, \beta_0)v$ as $[x_1^{\mathrm{T}} \;\; x_2^{\mathrm{T}} \;\; x_3^{\mathrm{T}} \;\; x_4^{\mathrm{T}}]^{\mathrm{T}}$, with appropriate dimensions, then (28) is equivalent to

$$\left(I - \begin{bmatrix} 0 & \Delta \\ \Delta^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} S & N \\ N^* & R \end{bmatrix} \right) \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + j \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} \right) = 0$$

which means that $\det\{I - [\begin{smallmatrix} 0 & \Delta \\ \Delta^{\mathrm{T}} & 0 \end{smallmatrix}][\begin{smallmatrix} S & N \\ N^* & R \end{smallmatrix}]\} = 0$. From the definition of $\psi_{\mathbb{R}}(M)$, we have $\psi_{\mathbb{R}}(M) \leqslant \bar{\sigma}(\Delta) = \lambda_0^{-1} \leqslant \{\inf_{\alpha \in (0,1], \beta > 0} \lambda_2 [P(\alpha, \beta)]\}^{-1}$, since $\lambda_0 \geqslant \inf_{\alpha \in (0,1], \beta > 0} \lambda_2 [P(\alpha, \beta)]$. Combining with Theorem 2, we get $\psi_{\mathbb{R}}(M) = \lambda_0^{-1}$. □

The seemingly innocent Theorem 4 has a long and dry proof, to which the rest of this section is devoted. We break the proof into several lemmas. In the sequal, we denote the inertia of $[\begin{smallmatrix} S_r & N_i \\ N_i^{\mathrm{T}} & R_r \end{smallmatrix}]$ as (π, v, ζ).

Lemma 8

For each fixed $\beta_0 > 0$,

1. $\lambda_1[P(\alpha, \beta_0)]$ is convex on $(0, \infty)$, its minimum is attained at $\alpha = 1$;
2. $\lambda_2[P(\alpha, \beta_0)]$ is unimodal on $(0, 1]$.

Proof. The convexity of $\lambda_1[P(\alpha, \beta_0)]$ is obvious by noting that $\lambda_1[P(\alpha, \beta_0)] = \bar{\sigma}[P(\alpha, \beta_0)]$ from Lemma 3. The minimum is attained at $\alpha = 1$ since $\lambda_1[P(\alpha, \beta_0)] = \lambda_1[P(\frac{1}{\alpha}, \beta_0)]$.

The rest is for the proof of 2. We can write $\det[\lambda I - P(\alpha, \beta_0)] = \alpha^{-(\pi + v)} f(\lambda, \alpha)$, where $f(\lambda, \alpha)$ is a polynomial in α with degree $2(\pi + v)$ and in λ with degree $2(m + p)$, note that $P(\alpha, \beta)$ is similar to the matrix in (18). So for almost all λ, there exist at most $2(\pi + v)$ non-zero α's (counting multiplicity) such that $\det[\lambda I - P(\alpha, \beta_0)] = 0$. In other words, for almost all $c > 0$, the intersection of the straight line $\lambda = c$ and the curves $\lambda = \lambda_i[P(-\alpha, \beta_0)]$ and $\lambda = \lambda_i[P(\alpha, \beta_0)]$, $0 < \alpha < \infty$, $i = 1, 2, \ldots, 2(m + p)$, consists at most $2(\pi + v)$ points.

Denote $c_1 = \inf_{\alpha \in (0,1]} \lambda_2[P(\alpha, \beta_0)]$ and $c_2 = \inf_{\alpha \in (0,1]} \lambda_2[P(-\alpha, \beta_0)]$. It is easy to see that $\lambda_i[P(-\alpha, \beta_0)] = -\lambda_{2(m+p)-i+1}[P(\alpha, \beta_0)]$, so by permutting $P(\alpha, \beta)$ to (23) and applying Lemma 3, we have $c_1 \geqslant c_2$.

As $\alpha \to 0$ (or ∞), $\lambda_i[P(\alpha, \beta_0)]$, $i \leqslant \pi$, and $\lambda_i[P(-\alpha, \beta_0)]$, $i \leqslant v$, go to infinity. Hence for any $c \geqslant \inf_{\alpha \in (0,1]} \lambda_i[P(\alpha, \beta_0)]$, $i \leqslant \pi$, the intersection of $\lambda = c$ and $\lambda = \lambda_i[P(\alpha, \beta_0)]$ consists at least two points on $(0, \infty)$. The same applies to $\lambda_i[P(-\alpha, \beta_0)]$, $i \leqslant v$.

From Lemma 4, $\lambda_1[P(-1, \beta_0)] = \lambda_2[P(-1, \beta_0)]$. Since $\lambda_i[P(-\alpha, \beta_0)]$ is continuous in α, it follows that the intersection of $\lambda = c$, $c \geqslant c_1 \geqslant c_2$, and $\lambda = \lambda_i[P(-\alpha, \beta_0)]$, $i \leqslant v$ has at least $2v$ points.

Now, suppose on the contrary that $\lambda_2[P(\alpha, \beta_0)]$ is not unimodal in $(0, 1]$, i.e., it has another local minimum c_1' in additional to c_1, on the interval $(0, 1]$. Then there must exist a local maximum $c_1'' > c_1'$. For all $c \in (c_1', c_2'')$, the straight line $\lambda = c$ must have at least six (if $\pi > 1$), or four (if $\pi = 1$) crossings with $\lambda = \lambda_2[P(\alpha, \beta_0)]$ for $\alpha \in (0, \infty)$, noting the property (17). Adding to the number of crossings with other curves, $\lambda = \lambda_i[P(\alpha, \beta_0)]$, $3 \leqslant i \leqslant \pi$ and $\lambda = \lambda_i[P(-\alpha, \beta_0)]$, $i \leqslant v$, we get more than $2(\pi + v)$ total crossings. Therefore, we conclude that $\lambda_2[P(\alpha, \beta_0)]$ is unimodal in $(0, 1]$. □

For $0 < \beta < \infty$, define

$$\underline{\lambda}(\beta) := \inf_{\alpha \in (0,1]} \lambda_2[P(\alpha, \beta)]$$

Lemma 9

$\underline{\lambda}(\beta) = \lim_{\alpha \to 0} \lambda_2[P(\alpha, \beta)]$ if and only if $\pi \leqslant 1$. In this case, $\underline{\lambda}(\beta)$ is a constant over $(0, \infty)$.

Proof. If $\pi > 1$, then $\lim_{\alpha \to 0} \lambda_2[P(\alpha, \beta_0)] = \infty$, so it cannot be equal to $\underline{\lambda}(\beta)$. The case when $\pi = 0$ is trivial; $P(\alpha, \beta)$ is actually a constant matrix in this case. For the case when $\pi = 1$, denote $c_0 = \lim_{\alpha \to 0} \lambda_2[P(\alpha, \beta_0)]$ (which is finite) for some fixed $\beta_0 > 0$. If $c_0 > \underline{\lambda}(\beta_0)$, then for any $c \in (\underline{\lambda}(\beta_0), c_0)$, $\lambda = \lambda_i[P(\alpha, \beta_0)]$, $i = 1, 2$ have at least four crossings with $\lambda = c$, which is impossible (cf. the proof of Lemma 8). Therefore, $\underline{\lambda}(\beta_0) = c_0$. In this case, owing to $S, R \geqslant 0$, we must have $\operatorname{rank} S_r$, $\operatorname{rank} R_r$, $\operatorname{rank} N_i \leqslant 1$ and $S_i = 0$, $R_i = 0$. Hence, by noting $v \leqslant \pi$, we see that $\left[\begin{smallmatrix} \beta S_r & N_i \\ N_i^T & R_r/\beta \end{smallmatrix}\right]$ is similar to

$$\begin{bmatrix} \beta a & c & 0 \\ c & b/\beta & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

for some real a, b, c with $a, b \geqslant 0$, $ab \leqslant c^2$. If ab < c^2, then as $\alpha \to 0$, the $2(m + p) - 2$ finite eigenvalues of $P(\alpha, \beta)$ approach some constants independent of β. If $ab = c^2$, then it can be verified that $\lim_{\alpha \to 0} \lambda_2[P(\alpha, \beta)] = \bar{\sigma}(N_r)$. □

In the following, we will assume that $\pi \geqslant 2$. Define

$$\underline{\alpha}(\beta) := \{\alpha \in (0, 1] : \lambda_2[P(\alpha, \beta)] = \underline{\lambda}(\beta)\}$$

and

$$\alpha_M(\beta) := \max\{\underline{\alpha}(\beta)\} \quad \alpha_m(\beta) := \min\{\underline{\alpha}(\beta)\}$$

With the assumption that $\pi \geqslant 2$, the set $\underline{\alpha}(\beta)$ is the non-empty and is closed interval for each $\beta > 0$. Hence $\alpha_m(\beta)$ and $\alpha_M(\beta)$ are well defined.

Lemma 10

When $\pi \geqslant 2$, $\underline{\lambda}(\beta)$ is continuous and piecewise analytic, and $\alpha_M(\beta)$, $\alpha_m(\beta)$ are piecewise analytic.

Proof. The continuity of $\underline{\lambda}(\beta)$ can be proved similarly to the proof of the Proposition in Reference 10. By applying Lemma 8 and the continuity of $\underline{\lambda}(\beta)$ and $\lambda_2[P(\alpha,\beta)]$, it can be shown that $\alpha_M(\beta)$ and $\alpha_m(\beta)$ are piecewise continuous. Let

$$\det[\lambda I - P(\alpha,\beta)] = \alpha^{-k_1}\beta^{-k_2}\prod_{i=1}^{N_1} e_i^{n_{ei}}(\lambda)\prod_{i=1}^{N_2} f_i^{n_{fi}}(\lambda,\beta)\prod_{i=1}^{N_3} g_i^{n_{gi}}(\lambda,\alpha)\prod_{i=1}^{N_4} h_i^{n_{hi}}(\lambda,\alpha,\beta)$$

where $e_i(\lambda)$, $i=1,\ldots,N_1$, are different prime polynomials in λ with positive degrees, $f_i(\lambda,\beta)$, $i=1,\ldots,N_2$, are different prime polynomials in λ and β, with positive degrees, etc. It should be noted that $\det[\lambda I - P(\alpha,\beta)]$ cannot have a divisor in the form of $f(\alpha,\beta)$, $f(\alpha)$ or $f(\beta)$.

Denote

$$x(\lambda,\alpha,\beta) = \prod_{i=1}^{N_1} e_i(\lambda)\prod_{i=1}^{N_2} f_i(\lambda,\beta)\prod_{i=1}^{N_3} g_i(\lambda,\alpha)\prod_{i=1}^{N_4} h_i(\lambda,\alpha,\beta)$$

$$y(\lambda,\alpha,\beta) = \prod_{i=1}^{N_3} g_i(\lambda,\alpha)\prod_{i=1}^{N_4} h_i(\lambda,\alpha,\beta)$$

Then for each $\alpha,\beta\neq 0$, the roots of $x(\lambda,\alpha,\beta)$ are the same as $\lambda_i[P(\alpha,\beta)]$, but with decreased multiplicities. And the roots of $y(\lambda,\alpha,\beta)$ are the eigenvalues of $P(\alpha,\beta)$ after excluding those that are independent of α.

It is easy to see that $x(\lambda,\alpha,\beta)$ and $\partial x(\lambda,\alpha,\beta)/\partial\lambda$ have no common divisor with positive degree, so from Theorem 5.5 and Theorem 5.6 of Reference 14, there exist polynomials $u_1(\lambda,\alpha,\beta)$, $v_1(\lambda,\alpha,\beta)$ and $t_1(\alpha,\beta)\neq 0$ such that

$$x(\lambda,\alpha,\beta)u_1(\lambda,\alpha,\beta) + \frac{\partial x(\lambda,\alpha,\beta)}{\partial\lambda}v_1(\lambda,\alpha,\beta) = t_1(\alpha,\beta) \tag{29}$$

Similarly, there exist $u_2(\lambda,\alpha,\beta)$, $v_2(\lambda,\alpha,\beta)$ and $t_2(\alpha,\beta)\neq 0$ such that

$$y(\lambda,\alpha,\beta)u_2(\lambda,\alpha,\beta) + \frac{\partial y(\lambda,\alpha,\beta)}{\partial\alpha}v_2(\lambda,\alpha,\beta) = t_2(\alpha,\beta) \tag{30}$$

Case 1: $\underline{\alpha}(\beta_0)$ is a single point. Then $\lambda_2[P(\alpha,\beta_0)]$ must be a root of $y(\lambda,\alpha,\beta_0)$ in a neighbourhood of $\underline{\alpha}(\beta_0)$. If $\underline{\lambda}(\beta_0)$ is a simple root of $y(\lambda,\underline{\alpha}(\beta_0),\beta_0)$, we have

$$\frac{\partial y(\underline{\lambda}(\beta_0),\underline{\alpha}(\beta_0),\beta_0)}{\partial\alpha} = -\frac{\partial y(\underline{\lambda}(\beta_0),\underline{\alpha}(\beta_0),\beta_0)}{\partial\lambda}\frac{\mathrm{d}\lambda_2[P(\underline{\alpha}(\beta_0),\beta_0)]}{\mathrm{d}\alpha} = 0$$

If $\underline{\lambda}(\beta_0)$ is not a simple root, then $\partial y(\underline{\lambda}(\beta_0),\underline{\alpha}(\beta_0),\beta_0)/\partial\lambda = 0$, which results in $\partial x(\underline{\lambda}(\beta_0),\underline{\alpha}(\beta_0),\beta_0)/\partial\lambda = 0$, since y divides x.

Case 2: $\underline{\alpha}(\beta_0)$ is an interval. It is easy to see that the multiplicity of $\underline{\lambda}(\beta_0)$ at $\alpha_M(\beta_0)$ is greater than one, among the roots of $x(\lambda,\alpha_M(\beta_0),\beta_0)$. So we also get $\partial x(\underline{\lambda}(\beta_0),\underline{\alpha}(\beta_0),\beta_0)/\partial\lambda = 0$. The same applies to $\alpha_m(\beta_0)$.

Combining the above two cases, we have

$$\begin{gathered}\{[(\underline{\lambda}(\beta),\alpha_M(\beta),\beta]:\beta>0\}\\ \subset\left\{(\lambda,\alpha,\beta):\frac{\partial x(\lambda,\alpha,\beta)}{\partial\lambda} = x(\lambda,\alpha,\beta) = 0 \text{ or } \frac{\partial y(\lambda,\alpha,\beta)}{\partial\alpha} = y(\lambda,\alpha,\beta) = 0\right\}\end{gathered}$$

Noticing (29) and (30), we obtain

$$\{[\alpha_M(\beta), \beta] : \beta > 0\} \subset \{(\alpha, \beta) : t_1(\alpha, \beta) = 0 \text{ or } t_2(\alpha, \beta) = 0\}$$

Denote the roots of $t_1(\alpha, \beta)$ and $t_2(\alpha, \beta)$ for fixed β as $\alpha_i(\beta)$. Then $\alpha_i(\beta)$ is piecewise analytic on the intervals where it takes values in the real field. Since $\alpha_M(\beta)$ is piecewise continuous when $\pi \geqslant 2$, we conclude that it is piecewise analytic. The same applies to $\alpha_m(\beta)$. Hence $P(\alpha_M(\beta), \beta)$ is piecewise analytic. It follows that $\underline{\lambda}(\beta)$ is also piecewise analytic. $\square$

Lemma 11

Let $F(\gamma) \in \mathbb{C}^{n \times n}$ be a Hermitian matrix function analytic on an open set $\Gamma \subset \mathbb{R}$. Suppose $\lambda_i(\gamma_0)$ is an eigenvalue of $F(\gamma_0)$ with multiplicity k, with corresponding eigenvector matrix $W \in \mathbb{C}^{n \times k}$, then the inertia of $W^*(\mathrm{d}F(\gamma_0)/\mathrm{d}\gamma)W$ represents the numbers of eigenvalues (among the k ones clustered at γ_0) that increase, decrease and stationary, respectively.

Proof. By Lemma 1, there exists $U(\gamma) \in \mathbb{C}^{n \times k}$, $U^*(\gamma)U(\gamma) = I_k$, and diagonal $\Lambda_k(\gamma) \in \mathbb{C}^{k \times k}$, both analytic on $\Gamma \subset \mathbb{R}$, such that

$$F(\gamma)U(\gamma) = U(\gamma)\Lambda_k(\gamma)$$

where $\Lambda_k(\gamma_0) = \lambda_i(\gamma_0)I_k$. Differentiating both sides and multiplying from the left with $U^*(\gamma)$, we get

$$U^*(\gamma_0)\frac{\mathrm{d}F(\gamma_0)}{\mathrm{d}\gamma}U(\gamma_0) = \frac{\mathrm{d}\Lambda_k(\gamma_0)}{\mathrm{d}\gamma}$$

This is for the special analytic eigenvector matrix. For arbitrarily selected eigenvector matrix W, $W = U(\lambda_0)Q$ for some non-singular Q. So,

$$W^*\frac{\mathrm{d}F(\gamma_0)}{\mathrm{d}\gamma}W = Q^*\frac{\mathrm{d}\Lambda_k(\gamma_0)}{\mathrm{d}\gamma}Q \qquad \square$$

and the inertia is invariant.

Lemma 12

When $\pi \geqslant 2$, $\underline{\lambda}(\beta)$ is unimodal.

Proof. We prove this by contradiction. Suppose that $\underline{\lambda}(\beta)$ has two local minima at β_1 and β_2. Then by Theorem 2, $\psi_{\mathbb{R}}(M) \geqslant \max\{\underline{\lambda}^{-1}(\beta_1), \underline{\lambda}^{-1}(\beta_2)\}$. Between β_1 and β_2, $\underline{\lambda}(\beta)$ must have a local maximum, say at β_0, with $\underline{\lambda}(\beta_0) > \max\{\underline{\lambda}(\beta_1), \underline{\lambda}(\beta_2)\}$. If this is true, then we can show that $\psi_{\mathbb{R}}(M) \leqslant \underline{\lambda}^{-1}(\beta_0)$, which is a contradiction.

By Lemma 10, $\alpha_M(\beta)$ is piecewise analytic, we can denote

$$\lim_{\beta \to \beta_0^-}\frac{\mathrm{d}\alpha_M(\beta)}{\mathrm{d}\beta} = t^- \quad \text{and} \quad \lim_{\beta \to \beta_0^+}\frac{\mathrm{d}\alpha_M(\beta)}{\mathrm{d}\beta} = t^+$$

If β_0 is a strict local maximizer of $\underline{\lambda}(\beta)$, then neither t^+ nor t^- can be infinity, since $\underline{\lambda}(\beta_0)$ is a local minimum of $\lambda_2[P(\alpha, \beta_0)]$, it cannot be a strict local maximum of $\lambda_2[P(\alpha, \beta)]$ along $\alpha_M(\beta)$. If β_0 is not a strict local maximizer of $\underline{\lambda}(\beta)$, then β_0 can always be chosen such that neither t^+ nor t^- is infinity.

Case 1: $\lim_{\beta\to\beta_0^-}\alpha_M(\beta)=\lim_{\beta\to\beta_0^+}\alpha_M(\beta)=\alpha_0$.

By applying Lemma 11 and observing that $\lambda_1[P(\alpha,\beta_0)]$ is convex and takes its minimum at $\alpha=1$, we can show that $\underline{\lambda}(\beta)$ cannot take its local maximum at β_0 with $\alpha_M(\beta_0)=1$, or with $\lambda_1[P(\alpha_0,\beta_0)]=\lambda_2[P(\alpha_0,\beta_0)]$.

Let the multiplicity of $\lambda_2[P(\alpha_0,\beta_0)]$ be k, then λ_2 and all of the k clustered eigenvalues increase along $\alpha_M(\beta)$ for $\beta<\beta_0$ and decrease for $\beta>\beta_0$. Let the corresponding eigenvector matrix be W, then by Lemma 11,

$$W^{\mathrm{T}}\left[\frac{\partial P(\alpha_0,\beta_0)}{\partial\alpha}t^-+\frac{\partial P(\alpha_0,\beta_0)}{\partial\beta}\right]W\geqslant 0,\quad W^{\mathrm{T}}\left[\frac{\partial P(\alpha_0,\beta_0)}{\partial\alpha}t^++\frac{\partial P(\alpha_0,\beta_0)}{\partial\beta}\right]W\leqslant 0 \tag{31}$$

Since α_0 is a local minimizer of $\lambda_2[P(\alpha,\beta_0)]$, so from Lemma 2, there exists a unit vector $z\in\mathbb{R}^k$ such that $z^{\mathrm{T}}W^{\mathrm{T}}(\partial P(\alpha_0,\beta_0)/\partial\alpha)Wz=0$. From (31), we must have $z^{\mathrm{T}}W^{\mathrm{T}}(\partial P(\alpha_0,\beta_0)/\partial\beta)Wz=0$.

Denote

$$Wz=\begin{bmatrix}v_1\\v_2\\v_3\\v_4\end{bmatrix},$$

then like the proof of Theorem 3, we have

$$[v_1\ \ v_3]^{\mathrm{T}}[v_1\ \ v_3]=[v_2\ \ v_4]^{\mathrm{T}}[v_2\ \ v_4]$$

So there exists a $\Delta\in\mathbb{R}^{m\times p}$ with $\bar{\sigma}(\Delta)=\underline{\lambda}^{-1}(\beta_0)$ and $\det\{I-[\begin{smallmatrix}0&\Delta\\\Delta^{\mathrm{T}}&0\end{smallmatrix}][\begin{smallmatrix}S&N\\N^*&R\end{smallmatrix}]\}=0$. Thus

$$\psi_{\mathbb{R}}(M)\leqslant\underline{\lambda}^{-1}(\beta_0)<\max\{\underline{\lambda}^{-1}(\beta_1),\underline{\lambda}^{-1}(\beta_2)\}\leqslant\psi_{\mathbb{R}}(M)$$

which is a contradiction.

Case 2: $\lim_{\beta\to\beta_0^-}\alpha_M(\beta)\neq\lim_{\beta\to\beta_0^+}\alpha_M(\beta)$.

Let $\alpha_1=\lim_{\beta\to\beta_0^-}\alpha_M(\beta)$ and $\alpha_2=\lim_{\beta\to\beta_0^+}\alpha_M(\beta)$. Then for $\alpha\in[\alpha_1,\alpha_2]$, $\lambda_2[P(\alpha,\beta_0)]$ is a constant, since $\lambda_2[P(\alpha,\beta_0)]$ is unimodal and $\underline{\lambda}(\beta)$ is continuous. As in case 1, we can show that $\alpha_1,\alpha_2<1$ and $\lambda_1[P(\alpha_i,\beta_0)]\neq\lambda_2[P(\alpha_i,\beta_0)]$, $i=1,2$. Let the multiplicity of the constant eigenvalues be m_1 and the multiplicity of the eigenvalues clustered at α_1 be m_1+m_2. Denote the analytic eigenvectors corresponding to these m_1+m_2 eigenvalues as $[W_1(\alpha)W_2(\alpha)]\in\mathbb{R}^{2(m+p)\times(m_1+m_2)}$, then $W_1^{\mathrm{T}}(\alpha)(\partial P(\alpha,\beta_0)/\partial\alpha)W_1(\alpha)=0$. Since β_0 is a local maximizer of $\underline{\lambda}(\beta)$, by Lemma 11, we have

$$\begin{bmatrix}W_1^{\mathrm{T}}(\alpha_1)\\W_2^{\mathrm{T}}(\alpha_1)\end{bmatrix}\left[\frac{\partial P(\alpha_1,\beta_0)}{\partial\alpha}t^-+\frac{\partial P(\alpha_1,\beta_0)}{\partial\beta}\right][W_1(\alpha_1)\ \ W_2(\alpha_1)]\geqslant 0$$

Hence $W_1^{\mathrm{T}}(\alpha_1)(\partial P(\alpha_1,\beta_0)/\partial\beta)W_1(\alpha_1)\geqslant 0$ and similarly $W_1^{\mathrm{T}}(\alpha_2)(\partial P(\alpha_2,\beta_0)/\partial\beta)W_1(\alpha_2)\leqslant 0$.

By continuity, there exists $\alpha_0\in[\alpha_1,\alpha_2]$ such that $W_1^{\mathrm{T}}(\alpha_0)(\partial P(\alpha_0,\beta_0)/\partial\beta)W_1(\alpha_0)$ has a zero eigenvalue. Therefore, there exists a unit vector $z\in\mathbb{R}^{m_1}$ such that

$$z^{\mathrm{T}}W_1^{\mathrm{T}}(\alpha_0)\frac{\partial P(\alpha_0,\beta_0)}{\partial\beta}W_1(\alpha_0)z=0$$

What is left is similar to case 1. $\square$

Proof of Theorem 4. Combining Lemma 8, Lemma 9, Lemma 10 and Lemma 12, we get the result. □

6. EXAMPLES

Example 1

Assume that a state space description of G of the form (9) is given by

$$\left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right]$$

$$= \left[\begin{array}{cccc|cccccc} 79 & 20 & -30 & -20 & -0{\cdot}5 & -0{\cdot}35 & 0 & 0 & 0{\cdot}3 & 0 \\ -41 & -12 & 17 & 13 & 0 & 0{\cdot}15 & 0{\cdot}2 & 0{\cdot}4 & 0 & 0{\cdot}2 \\ 167 & 40 & -60 & -38 & 0{\cdot}3 & 0 & 0 & 0 & 0 & 0 \\ 33{\cdot}5 & 9 & -14{\cdot}5 & -11 & 0 & 0{\cdot}3 & 0 & 0 & 0 & 0{\cdot}2 \\ \hline 0{\cdot}25 & 0 & 0 & 0 & 0{\cdot}1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0{\cdot}1 & 0 & 0{\cdot}1 & 0 & 0{\cdot}1 & 0 & 0{\cdot}2 & 0 & 0 \\ 0{\cdot}4 & 0 & 0{\cdot}5 & 0 & 0{\cdot}2 & 0 & 0{\cdot}1 & 0 & 0 & 0 \\ 0 & -0{\cdot}5 & 0 & 0 & -0{\cdot}2 & 0{\cdot}1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0{\cdot}2 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

The perturbation matrix Δ is 3×3. The computation result is:

$$r_{\mathbb{C}}(G_{22}) = 0{\cdot}5006; \quad p_{\mathbb{C}}(G) = 0{\cdot}1700$$

$$r_{\mathbb{R}}(G_{22}) = 1{\cdot}0432; \quad p_{\mathbb{R}}(G) \geqslant 0{\cdot}3998$$

Hence $pr_{\mathbb{C}}(G) = 0{\cdot}1700$; $\quad pr_{\mathbb{R}}(G) \geqslant 0{\cdot}3998$.

In Figure 4, the dashed line is $\psi_{\mathbb{C}}[M(j\omega)]$ and the solid line is the lower bound for $\psi_{\mathbb{R}}[M(j\omega)]$. For this example, the condition in Theorem 3 is satisfied at all ω. Therefore, the solid line in Figure 4 is actually the plot of $\psi_{\mathbb{R}}[M(j\omega)]$ and the real performance radius is exactly computed as $pr_{\mathbb{R}}(G) = 0{\cdot}3998$.

It is remarked in Section 4 that another lower bound of $\psi_{\mathbb{R}}(M)$ is $\{\inf_{\alpha,\beta>0} \sigma_2[P(\alpha,\beta)]\}^{-1}$. As a comparison, $\{\inf_{\alpha,\beta>0} \sigma_2[P(\alpha,\beta)]\}^{-1}$, where $P(\alpha,\beta)$ is formed from $M(j10{\cdot}11)$, is computed and the value is 0·2491, whereas the lower bound given by Theorem 2 is 0·3998. This shows that the lower bound of $\psi_{\mathbb{R}}(M)$ given by $\{\inf_{\alpha,\beta>0} \sigma_2[P(\alpha,\beta)]\}^{-1}$ is more conservative.

Example 2

It is of interest to know how tight the lower bound in Theorem 2 is. A sufficient condition is given in Theorem 3 for this lower bound to be the exact value of $\psi_{\mathbb{R}}(M)$. Hence the probability that the condition in Theorem 3 is satisfied gives an indication of how often the lower bound is tight. 4000 complex matrices $G = \left[\begin{smallmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{smallmatrix}\right]$ are randomly generated with $\bar{\sigma}(G_{11}) < 1$ and

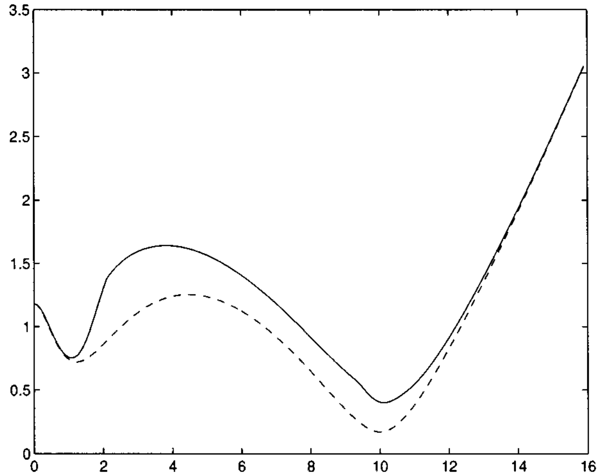

Figure 4. The solid line is a lower bound for $\psi_{\mathbb{R}}[M(j\omega)]$ and the dashed line is $\psi_{\mathbb{C}}[M(j\omega)]$

M is computed from G by using (2)–(5). The number of G matrices for which the condition of Theorem 3 is satisfied is 3 661. This shows that the probability of the lower bound being tight is over 90%.

7. CONCLUSION

In this paper, the concepts of complex and real performance radii are introduced to measure the performance robustness of an LTI system with respect to dynamic and parametric perturbations respectively. An algorithm for the computation of the complex performance radius is given, which complements the algorithms in the μ framework (which treats more general set-ups) for robust performance analysis. A computationally tractable lower bound for the real stability radius is given, which often turns out to be tight. Further study is needed for the exact computation of the real performance radius.

The analysis in this paper is for continuous time systems. The results can be easily adapted for discrete time systems.

ACKNOWLEDGEMENTS

This research was supported by the Hong Kong Research Grants Council grant HKUST 552/94E.

REFERENCES

1. Chen, M.J. and C.A. Desoer, 'Necessary and sufficient condition for robust stability of linear distributed systems', *Int. J. Control*, **35**, 255–267 (1982).
2. Doyle, J.C., 'Analysis of feedback systems with structured uncertainty', *IEE Proc. Part D*, **129**, 242–250 (1982).
3. Safonov, M.G., 'Stability margin of diagonally perturbed multivariable feedback systems', *Proc. IEE*, *Part D*, **129**, 251–256 (1982).
4. Fan, M.K., A.L. Tits and J.C. Doyle, 'Robustness in the presence of mixed parametric uncertainty and unmodeled dynamics', *IEEE Trans. Automat. Control*, **36**, 25–38 (1991).
5. Packard, A. and J.C. Doyle, 'The complex structured singular value', *Automatica*, **29**, 71–110 (1993).

6. Fan, M.K. and A.L. Tits, 'A measure of worst-case $\mathscr{H}_\infty$ performance and of largest acceptable uncertainty', *Systems & Control Letters*, **18**, 409–421 (1992).
7. Doyle, J.C. and G. Stein, 'Multivariable feedback design: concepts for a classical/modern synthesis', *IEEE Trans. Automat. Control*. **AC-26**, 4–16 (1981).
8. Hinrichsen, D. and A.J. Pritchard, 'Stability radius for structured perturbations and the algebraic Riccati equation', *Systems & Control Letters*, **8**, 105–113 (1986).
9. Zhou, K., J.C. Doyle and K. Glover, *Robust and Optimal Control*, Prentice Hall, Upper Saddle river, N.J. 1996.
10. Qiu, L., B. Bernhardsson, A. Rantzer, E.J. Davison, P.M. Young and J.C. Doyle, 'A formula for computation of the real stability radius', *Automatica*, **31**, 879–890 (1995).
11. Baumgärtel, H., *Analytic Perturbation Theory for Matrix Operators*, Birkhäuser, Basel, 1985.
12. Sezginer, R.S. and M.L. Overton, 'The largest singular value of $e^X A_0 e^{-X}$ is convex on convex sets of commuting matrices', *IEEE Trans. Automat. Control*, **35**, 229–230 (1990).
13. Stewart, G.W. and J.G. Sun, *Matrix Perturbation Theory*, Academic Press, San Diego, 1990.
14. Marcus, M., *Introduction to Modern Algebra*, Marcel, Inc., New York, 1978.