Proceedings of the 28th Conference
on Decision and Control
Tampa, Florida • December 1989

A Unified Approach for the Stability Robustness of Polynomials in a Convex Set*

L. Qiu and E.J. Davison
Department of Electrical Engineering
University of Toronto
Toronto, Ontario, Canada M5S 1A4

Abstract

Consider a polynomial $p(s,k)$ which is affine in the parameter perturbation k; assume that the vector k is uncertain but belongs to a convex set which contains the origin, and call a polynomial stable if all of its roots are contained in a pre-specified stability region in the complex plane. Then the stability robustness of $p(s,k)$ can be measured by the maximal nonnegative number ρ with the property that if the gauge (or the Minkowski functional) of k with respect to the convex set is less than ρ, the polynomial $p(s,k)$ is always stable. This paper develops a unified approach to compute the robustness measure ρ. The approach imbeds the problem considered into the framework of convex analysis so that some powerful tools in convex analysis can be used. The approach is very general because of two reasons: i) the stability region in the complex plane can be an arbitrary open set, which includes the open left half plane and the open unit disc as special cases; ii) the convex set in which k is contained can be assumed to have an arbitrary shape, which includes polytopes and ellipsoids as special cases. The computational procedure to compute ρ which results from this approach is easy to implement. Various examples are included to illustrate the type of results which may be obtained.

1 Introduction

Consider an n-th degree real polynomial in a complex variable s. Assume that its coefficients are affine functions of a vector $k \in \mathbf{R}^m$, whose entries represent independent physical parameters. This polynomial can be written as

$$p(s,k) = a_0(k)s^n + a_1(k)s^{n-1} + \cdots + a_{n-1}(k)s + a_n(k), \qquad (1)$$

where $a_i(k)$, $i = 0, 1, \ldots, n$, are affine functionals of k, i.e. there exist a matrix $F \in \mathbf{R}^{(n+1)\times m}$ and a vector $g \in \mathbf{R}^{n+1}$ such that

$$a(k) := [a_0(k)\ a_1(k) \cdots a_n(k)]' = Fk + g. \qquad (2)$$

In many applications, the parameter k is somewhat uncertain but is known to be contained in a given set $S \subset \mathbf{R}^m$. In this paper, the set S is assumed to be a closed convex set. With no loss of generality, we also assume that S contains the origin.

In many control problems, a desired property of a polynomial is that all of its roots are located in a pre-specified area in the complex plane. This pre-specified area will be called the stability region and a polynomial is said to be stable if all of its roots are located there. A stability robustness problem which naturally arises for the polynomial of the form (1)-(2) is to determine whether or not the polynomial $p(s,k)$ is stable for all k contained in S; this problem will be called the robustness checking problem. However, an answer to the robustness checking problem is often not enough. It is desired to have a quantitative robustness measure of the polynomial model (1)-(2). If we define the gauge (or Minkowski functional) of the set S to be a function μ_S: $\mathbf{R}^m \to \mathbf{R} \cup \{\infty\}$ such that $\mu_S(k) = \inf\{\alpha > 0 : \alpha^{-1}k \in S\}$, then a relative size of k is provided by $\mu_S(k)$ and a robustness measure is provided by the maximal nonnegative number ρ with the property that the polynomial $p(s,k)$ is stable for all k with $\mu_S(k) < \rho$. The problem to find such robustness measure ρ is called the maximal robustness problem. It is clear that an answer to the maximal robustness problem automatically provides an answer to the robustness checking problem, but to do the converse requires an extra one-dimensional search.

The purpose of this paper is to find a procedure to compute the robustness measure ρ when F, g, S and the stability region are given. In this paper, a simple and numerically feasible procedure is obtained to compute ρ. The procedure can be applied to very general cases: i) the stability region in the complex plane can be an arbitrary open set; ii) the convex set S can be assumed to have an arbitrary shape.

A special case of the above problems occurs when each column and row of F have at most one nonzero element, the stability region is the open left half part of the complex plane, and the convex set S is a hyperrectangle. In this case, the polynomial $p(s,k)$ is called an interval polynomial. The remarkable theorem of Kharitonov [10] gives an elegant solution to the robustness checking problem of an interval matrix. It says that $p(s,k)$ is stable for all k in S if and only if four specially constructed polynomials are stable. The maximal robustness problem for an interval polynomial can be easily solved in terms of Kharitonov's four polynomials using the Hurwitz stability criteria [9].

The stability of interval polynomials with respect to the open left half part of the complex plane is a very restricted special case of the general problem, in which the matrix F, the stability region and the convex set S are arbitrary. Although some attempts have been made to generalize the result of Kharitonov to the general problem (see [1], [2], [6] and the references in [3]), no results obtained have the same level of simplicity as Kharitonov's theorem. Recently, a considerable amount of research has lead to the development of feasible numerical methods to solve the maximal robustness problem and the robustness checking problem for polynomials whose coefficients are general affine functions of uncertain parameters. References [1] and [2] consider the robustness checking problem for the case when the set S is a polytope. The methods used in [1] and [2] are based on the concept of the value set, which is the set of complex numbers $p(s,k)$ with s fixed and k varying in a known set. The computation required in the methods of [1] and [2] are combinatorially explosive with respect to the number of uncertain parameters, so that they are only realistic to apply when the number of uncertain parameters is small. The maximal robustness problem is solved in [5] and [8] for the case when the set S is a hyperellipsoid by using Euclidean space projection theory. In this case the gauge of S is actually the weighted Hölder 2-norm. The case when S is a polytope is solved in [7] by using the geometry of the value sets and in [17] by using a linear programming method. The method in [7] completely eliminates the combinatorical explosion of the computational complexity with respect to the number of uncertain parameters. A new method based on the Hahn-Banach theorem is developed in [11] to solve the maximal robustness problem for the case when the stability robustness is measured by norms, which are just special kinds of gauges. In this paper, the method used in [11] is extended to the case when S is any closed convex set containing the origin.

Due to space limitations, all of the proofs are excluded from this paper; a complete treatment of the material covered in this paper is given in [12].

The following notation is used throughout this paper:

$\mathbf{R}$(or $\mathbf{C}$)	the field of real (or complex) numbers
$\mathbf{R}^+$	the set of nonnegative real numbers
$\overline{\mathbf{R}}$	$\mathbf{R} \cup \{-\infty, \infty\}$

*This work has been supported by the National Sciences and Engineering Research Council of Canada under grant no. A4396.

CH2642-7/89/0000-0030$1.00 © 1989 IEEE

$\overline{R}^+$	$R^+ \cup \{\infty\}$
$\Re(s)$(or $\Im(s)$)	the real (or complex) part of $s \in C$
cl(S)	closure of the set S
co(S)	convex hull of the set S

In this paper, we have to consider algebraic operations in $\overline{R}^+$. If these operations involve ∞, the conventional rules are used. In addition to the obvious rules, these rules include: $0\infty = \infty 0 = 0$ and $\inf \varnothing = \infty$. However, the following operations are avoided: $\frac{\infty}{\infty}$, $\frac{0}{0}$, and $\infty - \infty$.

2 Preliminaries

Assume in the following development that the vector space $\mathbf{R}^m$ is equipped with the usual inner product $\langle x, y\rangle = x'y$, $\forall x, y \in \mathbf{R}^m$. This inner product induces a norm (the Hölder 2-norm) $\|x\|_2 = \langle x, x\rangle^{\frac{1}{2}}$, $\forall x \in \mathbf{R}^m$. Let S, T be subsets of $\mathbf{R}^m$ and α, β be scalers in R. We use αS to denote the set $\{\alpha x : x \in S\}$ and $\alpha S + \beta T$ to denote the set $\{\alpha x + \beta y : x \in S,\ y \in T\}$. Then S is *convex* if and only if $\alpha S + (1-\alpha)S \subset S$ for all $\alpha \in [0, 1]$. This section introduces some preliminary results required in our development. The proofs of the results presented can be found in many standard references, e.g. [4], [13], [14]. Reference [12] also provides a complete and concise source of these proofs.

Definition 1 *A set $S \subset \mathbf{R}^m$ is said to be*

(a) *bounded if there exists $\alpha > 0$ such that $\|x\|_2 \leq \alpha$ for every $x \in S$,*

(b) *absorbing if for each $x \in \mathbf{R}^m$, there exists $\alpha > 0$ such that $x \in \beta S$ for every $\beta \geq \alpha$,*

(c) *balanced if $\alpha S \subset S$ for every $\alpha \in [-1, 1]$,*

(d) *ellipsoidal if $S = \{x \in \mathbf{R}^m : \langle x, Mx\rangle \leq 1\}$ for some positive definite matrix $M \in \mathbf{R}^{m\times m}$,*

(e) *polyhedral if S is the intersection of a finite number of sets of the form $\{x \in \mathbf{R}^m : \langle x, y\rangle \leq \alpha\}$ for some $y \in \mathbf{R}^m$ and some $\alpha \in \mathbf{R}$.*

If $y \neq 0$, a set of the form $\{x \in \mathbf{R}^m : \langle x, y\rangle \leq \alpha\}$ is usually called a *closed halfspace*; its boundary $\{x \in \mathbf{R}^m : \langle x, y\rangle = \alpha\}$ is called a *hyperplane*. In the case when $y = 0$, the set $\{x \in \mathbf{R}^m : \langle x, y\rangle \leq \alpha\}$ is either $\varnothing$ or $\mathbf{R}^m$. A bounded nonempty polyhedral set is called a *polytope*; equivalently, we can define a polytope to be the convex hull of a finite set of points.

Let $S \subset \mathbf{R}^m$ be a closed convex set containing the origin.

Definition 2 *The gauge (or Minkowski functional) of S is the function $\mu_S : \mathbf{R}^m \to \overline{\mathbf{R}}^+$, defined by*

$$\mu_S(x) = \inf\{\alpha > 0 : \alpha^{-1}x \in S\}. \tag{3}$$

Some properties of the gauge are given by the following proposition.

Proposition 1 *Suppose that $S \subset \mathbf{R}^m$ is a closed convex set containing the origin. Then*

(a) $\mu_S(x + y) \leq \mu_S(x) + \mu_S(y)$,

(b) $\mu_S(\alpha x) = \alpha\mu_S(x)$ *if* $\alpha \geq 0$,

(c) $\mu_S(x) = 0$ *is equivalent to* $x = 0$ *if S is bounded,*

(d) $\mu_S(x) < \infty$ *for every* $x \in \mathbf{R}^m$ *if S is absorbing,*

(e) μ_S *is a norm if S is bounded, absorbing and balanced,*

(f) $S = \{x \in \mathbf{R}^m : \mu_S(x) \leq 1\}$,

(g) μ_S *is lower semi-continuous; it is continuous if S is absorbing.*

It is seen from Proposition 1 that the gauge is a generalization of the norm. The set S is the generalized unit ball. For example, if $S = \{x : \sum_{i=1}^m |x_i|^p = 1\}$, then μ_S is the Hölder p-norm. When $p = 2$, the set S is ellipsoidal. When $p = 1$ or ∞, the set S is polyhedral (actually a polytope).

Let $S \subset \mathbf{R}^m$ be an arbitrary set.

Definition 3 *The polar of S, denoted by S°, is defined by*

$$S^\circ = \{y \in \mathbf{R}^m : \langle x, y\rangle \leq 1,\ \forall x \in S\}. \tag{4}$$

Proposition 2 *The polar S° of any set $S \subset \mathbf{R}^m$ is a closed convex set and contains the origin. If S is absorbing, then S° is bounded. If S is bounded, then S° is absorbing. If S is balanced, ellipsoidal or polyhedral, so is S°.*

Since S° is also a set in $\mathbf{R}^m$, it has a polar which is written as $S^{\circ\circ}$ instead of $(S^\circ)^\circ$. The set $S^{\circ\circ}$ is sometimes called the *bipolar* of S.

Proposition 3 *For any set $S \subset \mathbf{R}^m$, $S^{\circ\circ} = $ cl[co$(\{0\} \cup S)$].*

Corollary 1 *If $S \subset \mathbf{R}^m$ is a closed convex set containing the origin, then $S^{\circ\circ} = S$.*

Since S° is a closed convex set containing the origin, its gauge is defined.

Definition 4 *Let μ_S be the gauge of a closed convex set $S \subset \mathbf{R}^m$ containing the origin. Then the gauge μ_{S° of S° is called the dual gauge of μ_S.*

The dual gauge of μ_S can also be related to μ_S through the following proposition.

Proposition 4 *Let μ_S be the gauge of a closed convex set $S \subset \mathbf{R}^m$ containing the origin. Then its dual gauge μ_{S° satisfies*

$$\mu_{S^\circ}(y) = \inf\{\alpha \geq 0 : \langle x, y\rangle \leq \alpha\mu_S(x),\ \forall x \in \mathbf{R}^m\}. \tag{5}$$

If μ_S is a norm, the right hand side of (5) is the definition of the dual norm. Hence, the concept of the dual gauge is a generalization of that of the dual norm. For example, if μ_S is the Holder p-norm, then μ_{S° is the Holder q-norm, where $\frac{1}{p} + \frac{1}{q} = 1$.

A key result in our development is a direct application of the well-known Hahn-Banach theorem [15].

Lemma 1 (Hahn-Banach Theorem) *Suppose*

(a) $\phi : \mathbf{R}^m \to \mathbf{R}^+$ *is a function satisfying* $\phi(x + y) \leq \phi(x) + \phi(y)$ *and* $\phi(\alpha x) = \alpha\phi(x)$ *for all* $x, y \in \mathbf{R}^m$, $\alpha \geq 0$,

(b) f *is a linear functional on $\mathcal{V}$ and* $f(x) \leq \phi(x)$ *for all* $x \in \mathcal{V}$.

Then there exists a linear functional $\hat{f}$ on $\mathbf{R}^m$ such that $\hat{f}(x) = f(x)$ for all $x \in \mathcal{V}$ and $\hat{f}(x) \leq \phi(x)$ for all $x \in \mathbf{R}^m$.

Note that the condition (a) of Lemma 1 is always satisfied by the gauge of any absorbing convex set S.

3 Problem Formulation

Let $\overline{\mathbf{C}}$ be the one point compactification of $\mathbf{C}$. It is known that $\overline{\mathbf{C}}$ is homeomorphic to the Riemann sphere. Partition $\overline{\mathbf{C}}$ into two disjoint subsets $\mathbf{C}_g$ and $\mathbf{C}_b$, i.e. $\overline{\mathbf{C}} = \mathbf{C}_g \dot{\cup} \mathbf{C}_b$, such that $\mathbf{C}_g$ is open. The reason why we consider $\overline{\mathbf{C}}$ instead of $\mathbf{C}$ is to avoid ambiguity at ∞; we want ∞ to belong to either $\mathbf{C}_g$ or $\mathbf{C}_b$ but not both. Let $\mathcal{P}$ be the space of all real polynomials with degree not larger than n. The number of roots of each polynomial in $\mathcal{P}$ is made to be n by supplementing an appropriate number of ∞'s as its roots. We use $p(s, k)$ to denote the image of $k \in \mathbf{R}^m$ under a fixed affine map from $\mathbf{R}^m$ to $\mathcal{P}$ which is completely characterized by a matrix $F \in \mathbf{R}^{(n+1)\times m}$ and a vector $g \in \mathbf{R}^{n+1}$ as

$$p(s, k) = [s^n\ s^{n-1} \cdots 1](Fk + g). \tag{6}$$

A polynomial in $\mathcal{P}$ is said to be *stable* if all its roots are in $\mathbf{C}_g$. For any given closed convex set $S \subset \mathbf{R}^m$ containing the origin, define

$$\rho := \inf\{\mu_S(k) : k \in \mathbf{R}^m \text{ and } p(s, k) \text{ is unstable}\}. \tag{7}$$

The purpose of this paper is to find a procedure to compute ρ when F, g, $\mathbf{C}_g$ and S are given. If $p(s, 0)$ is unstable, we must have $\rho = 0$, so it is always assumed in the following that $p(s, 0)$ is stable. Alternatively we can write ρ as

$$\rho = \inf\{\mu_S(k) : k \in \mathbf{R}^m \text{ and } \exists s \in \mathbf{C}_b \text{ such that } p(s, k) = 0\}. \tag{8}$$

Denote the boundary of $\mathbf{C}_g$ by $\partial\mathbf{C}_g$, i.e. $\partial\mathbf{C}_g = \mathbf{C}_b \cap \text{cl}(\mathbf{C}_g)$. Then simple continuity arguments show that

$$\rho = \inf\{\mu_S(k) : k \in \mathbf{R}^m \text{ and } \exists s \in \partial\mathbf{C}_g \text{ such that } p(s,k) = 0\}, \quad (9)$$

which can be rewritten as

$$\rho = \inf_{s\in\partial\mathbf{C}_g}\{\inf\{\mu_S(k) : k \in \mathbf{R}^m \text{ and } p(s,k) = 0\}\}. \quad (10)$$

Define a function $\tau(s) : \partial\mathbf{C}_g \to \overline{R}^+$ by

$$\tau(s) = \inf\{\mu_S(k) : k \in \mathbf{R}^m \text{ and } p(s,k) = 0\}. \quad (11)$$

It is seen from (10) that the computation of ρ can be accomplished in two phases. The first phase is to find $\tau(s)$ for any fixed $s \in \partial\mathbf{C}_g$. The second phase is to carry out a search over all points in $\partial\mathbf{C}_g$, which is usually a one-dimensional curve in $\overline{\mathbf{C}}$, to find $\inf_{s\in\partial\mathbf{C}_g}\tau(s)$.

Now let $s \in \partial\mathbf{C}_g$ be fixed. The equation $p(s,k) = 0$ becomes

$$[s^n\ s^{n-1} \cdots 1]Fk = -[s^n\ s^{n-1} \cdots 1]g. \quad (12)$$

The assumption on the stability of $p(s,0)$ implies that the right hand side of (12) is nonzero. Let

$$w(s) := -\frac{F'[s^n\ s^{n-1} \cdots 1]'}{[s^n\ s^{n-1} \cdots 1]g}$$

and let $u(s) := \Re[w(s)]$, $v(s) := \Im[w(s)]$, the real and the imaginary part of w respectively. Then $u(s), v(s) \in \mathbf{R}^m$ and (12) is equivalent to

$$u(s)'k = 1 \quad\text{and}\quad v(s)'k = 0. \quad (13)$$

Consequently, we have

$$\tau(s) = \inf\{\mu_S(k) : k \in \mathbf{R}^m \text{ and } u(s)'k = 1, v(s)'k = 0\}. \quad (14)$$

Therefore, the first phase of the problem to compute ρ becomes a specialization of the following problem: find

$$\inf\{\mu_S(k) : k \in \mathbf{R}^m \text{ and } u'k = 1, v'k = 0\} \quad (15)$$

for any given $u, v \in \mathbf{R}^m$. A straightforward method to solve this problem can be directly obtained since this is just a nonlinear programming problem with a convex cost function $\mu_S(k)$ and linear constraints. It can be shown that this nonlinear programming problem is reduced to a linear programming problem if S is a polytope. This is basically the method used in [17]. The purpose of this paper is not to pursue this direction; instead, we will simplify this problem to a form which is much more tractable numerically.

The second phase of the problem to compute ρ is usually carried out by a "brute force search" over $\partial\mathbf{C}_g$. The details and the possible numerical difficulties of this search are discussed in Section 7.

Finally, we note that in most applications $\partial\mathbf{C}_g$ is symmetric to the real axis. Since $p(s,k) = 0$ if and only if $p(\bar{s},k) = 0$, where $\bar{s}$ means the conjugate of s, it is sufficient in this case to carry out the search over the intersection of $\partial\mathbf{C}_g$ with the closed upper half of the complex plane.

4 The Main Result

The purpose of this section is to analyze the quantity $\inf\{\mu_S(k) : k \in \mathbf{R}^m \text{ and } u'k = 1, v'k = 0\}$, which is essential in the computation of $\tau(s)$ defined in the previous section. We only consider in the following sections the case when S is absorbing and bounded; in other words, the set S is a bounded convex set containing the origin as an interior point. In this case, both $\mu_S(x)$ and $\mu_{S^\circ}(x)$ are less than infinity for all x and strictly positive for nonzero x. The main theorem given in this section simplifies the quantity $\inf\{\mu_S(k) : k \in \mathbf{R}^m \text{ and } u'k = 1, v'k = 0\}$ to a form more readily computable.

From the theory of linear equations it is known that if $\text{rank}[u\ v] \neq \text{rank}\begin{bmatrix} u & v \\ 1 & 0 \end{bmatrix}$, then there exist no $k \in \mathbf{R}^m$ such that the equations $u'k = 1$ and $v'k = 0$ are satisfied. Thus in this case

$$\inf\{\mu_S(k) : k \in \mathbf{R}^m \text{ and } u'k = 1, v'k = 0\} = \infty.$$

Now assume that $\text{rank}[u\ v] = \text{rank}\begin{bmatrix} u & v \\ 1 & 0 \end{bmatrix} = 1$ or 2. Any k satisfying $\langle k,u\rangle = 1$ and $\langle k,v\rangle = 0$ defines a linear functional f on the subspace $\mathcal{V} := \text{span}\{u,v\} \subset \mathbf{R}^m$ by $f(\alpha u + \beta v) = \alpha$. Define a function $\phi : \mathbf{R}^m \to \mathbf{R}^+$ by

$$\phi(x) = \mu_{S^\circ}(x) \sup_{\substack{\alpha,\beta\in\mathbf{R} \\ \alpha u+\beta v\neq 0}} \frac{\alpha}{\mu_{S^\circ}(\alpha u + \beta v)}$$

Then $\phi(x)$ satisfies the conditions in Lemma 1. Therefore there exists a linear functional $\overline{f}$ on $\mathbf{R}^m$ such that $\overline{f}(x) = f(x)$ for all $x \in \mathcal{V}$, and $\overline{f}(x) \leq \phi(x)$ for all $x \in \mathbf{R}^m$. This means, by the Riesz representation theorem ([15]), that there exists $k \in \mathbf{R}^m$ such that $\langle k,u\rangle = 1$, $\langle k,v\rangle = 0$, and $\langle k,x\rangle \leq \phi(x)$ for all $x \in \mathbf{R}^m$. Furthermore, the inequality becomes an equality for some $x \in \mathcal{V}$. By Proposition 4, we obtain

$$\begin{aligned}
\mu_S(k) &= \mu_{S^{\circ\circ}}(k) \\
&= \sup_{\substack{\alpha,\beta\in\mathbf{R} \\ \alpha u+\beta v\neq 0}} \frac{\alpha}{\mu_{S^\circ}(\alpha u + \beta v)} \\
&= \max\{\sup_{\substack{\alpha\in\mathbf{R} \\ u+\alpha v=0}} \frac{1}{\mu_{S^\circ}(u + \alpha v)}, \sup_{\substack{\alpha\in\mathbf{R} \\ -u+\alpha v=0}} \frac{1}{\mu_{S^\circ}(-u + \alpha v)}\}.
\end{aligned}$$

A special case happens when u, v are linearly dependent. Since we have already assumed that $\text{rank}[u\ v] = \text{rank}\begin{bmatrix} u & v \\ 1 & 0 \end{bmatrix}$, then u, v are linearly dependent if and only if $u \neq 0$ and $v = 0$. In this case, we have

$$\sup_{\substack{\alpha\in\mathbf{R} \\ u+\alpha v\neq 0}} \frac{1}{\mu_{S^\circ}(u + \alpha v)} = \frac{1}{\mu_{S^\circ}(u)}$$

and

$$\sup_{\substack{\alpha\in\mathbf{R} \\ -u+\alpha v\neq 0}} \frac{1}{\mu_{S^\circ}(-u + \alpha v)} = \frac{1}{\mu_{S^\circ}(-u)}.$$

The following theorem summarizes the above development.

Theorem 1 *For any given $u, v \in \mathbf{R}^m$,*

$$\inf\{\mu_S(k) : k \in \mathbf{R}^m \text{ and } u'k = 1, v'k = 0\}$$
$$= \begin{cases} \infty & if\ \text{rank}[u\,v] \neq \text{rank}\begin{bmatrix} u & v \\ 1 & 0 \end{bmatrix} \\ \max\{\frac{1}{\mu_{S^\circ}(u)}, \frac{1}{\mu_{S^\circ}(-u)}\} & if\ u \neq 0\ and\ v = 0 \\ \max\{\sup_{\alpha\in\mathbf{R}} \frac{1}{\mu_{S^\circ}(u + \alpha v)}, \\ \quad \sup_{\alpha\in\mathbf{R}} \frac{1}{\mu_{S^\circ}(-u + \alpha v)}\} & if\ \text{rank}[u\,v] = 2. \end{cases} \quad (16)$$

In the special case when S is balanced, the gauges μ_S and μ_{S° are actually norms. Theorem 1 can be simplified as follows in this case:

Corollary 2 *If S is a balanced closed convex set containing the origin, then for any given $u, v \in \mathbf{R}^m$,*

$$\inf\{\mu_S(k) : k \in \mathbf{R}^m \text{ and } u'k = 1, v'k = 0\}$$
$$= \begin{cases} \infty & if\ \text{rank}[u\,v] \neq \text{rank}\begin{bmatrix} u & v \\ 1 & 0 \end{bmatrix} \\ \frac{1}{\mu_{S^\circ}(u)} & if\ u \neq 0\ and\ v = 0 \\ \sup_{\alpha\in\mathbf{R}} \frac{1}{\mu_{S^\circ}(u + \alpha v)} & if\ \text{rank}[u\,v] = 2. \end{cases} \quad (17)$$

It is apparent that if μ_S and μ_{S° can be easily evaluated (which is the case in most practical applications), the computation of the right hand side of (16), which contains an optimization problem in a scaler variable α, is much simpler than the direct computation of the left hand side of (16), which contains an optimization problem in a vector variable k. For obvious reasons, the critical problem in the computation of (16) is the computation of $\sup_{\alpha\in\mathbf{R}} \frac{1}{\mu_{S^\circ}(u+\alpha v)}$ for any given $u, v \in \mathbf{R}^m$ with rank$[u\ v] = 2$. Therefore, in the following when we refer to the computation of (16), we always mean the computation of $\sup_{\alpha\in\mathbf{R}} \frac{1}{\mu_{S^\circ}(u+\alpha v)}$, or equivalently $\inf_{\alpha\in\mathbf{R}} \mu_{S^\circ}(u+\alpha v)$, for any given $u, v \in \mathbf{R}^m$.

If μ_{S° can be easily evaluated, the computation of $\inf_{\alpha\in\mathbf{R}} \mu_{S^\circ}(u + \alpha v)$ is actually very straightforward. It follows from Proposition 1 that $\mu_{S^\circ}(u+\alpha v)$, when considered as a function of α, is a continuous convex function on $\mathbf{R}$. As α goes to $\pm\infty$, $\mu_{S^\circ}(u+\alpha v)$ goes to ∞. Consequently, $\inf_{\alpha\in\mathbf{R}} \mu_{S^\circ}(u+\alpha v)$ is achieved at a finite point and any technique for a convex one-dimensional optimization, such as the Fibonacci search or the golden section search, can be used to find $\inf_{\alpha\in\mathbf{R}} \mu_{S^\circ}(u+\alpha v)$. It will be shown in the following sections that if μ_S is one of the commonly used Hölder p-norms, or if S is a polytope with known vertices, then the computation of $\inf_{\alpha\in\mathbf{R}} \mu_{S^\circ}(u + \alpha v)$ becomes very simple and no one-dimensional optimization is required.

In some applications, however, the evaluation of μ_{S° may not be easy; some techniques which may simplify the evaluation of μ_{S° are developed in [12].

5 Important Special Cases

This section deals with the computation problem of (16) when the convex set S is an ellipsoid, a parallelotope or a crosspolytope. An ellipsoid is just an ellipsoidal set which is defined in Definition 1. The definitions of parallelotopes and crosspolytopes are given as follows. For simplicity, we denote the convex hull of two points $x, y \in \mathbf{R}^m$ by $[x, y]$.

Definition 5 *Let $x_1, x_2, \ldots, x_m \in \mathbf{R}^m$ be linearly independent. A parallelotope in $\mathbf{R}^m$ is a set of the form*

$$[x_1, -x_1] + [x_2, -x_2] + \cdots + [x_m, -x_m].$$

The set of vectors $\{x_1, x_2, \ldots, x_m\}$ is called the basis of the parallelotope.

Definition 6 *Let $x_1, x_2, \ldots, x_m \in \mathbf{R}^m$ be linearly independent. A crosspolytope in $\mathbf{R}^m$ is a set of the form*

$$\mathrm{co}\{[x_1, -x_1], [x_2, -x_2], \ldots, [x_m, -x_m]\}.$$

The set of vectors $\{x_1, x_2, \ldots, x_m\}$ is called the basis of the crosspolytope.

The parallelotopes and the crosspolytopes defined here are centered at the origin. The definitions given in many textbooks also cover the shifted version of the ones defined above. It is a trivial fact that the Hölder ∞-norm and 1-norm are respectively the gauges of the parallelotope and the crosspolytope with basis $\{e_1, e_2, \ldots, e_m\}$, where $e_i, i = 1, 2, \ldots, m$, is a vector with 1 in the i-th coordinate and zero elsewhere. Conversely, it is shown in this section that the gauge of any parallelotope is a weighted version of the Hölder ∞-norm, and the gauge of any crosspolytope is a weighted version of the Hölder 1-norm. It is also taken as a trivial fact here that the images of a parallelotope and a crosspolytope with basis $\{x_1, x_2, \ldots, x_m\}$ under a nonsingular linear map H on $\mathbf{R}^m$ are respectively a parallelotope and a crosspolytope with basis $\{Hx_1, Hx_2, \ldots, Hx_m\}$.

Case I S is an ellipsoid

The following proposition shows that if S is an ellipsoid, then μ_S and μ_{S° are both weighted Hölder 2-norms. We denote the unique positive definite square root of a positive definite matrix M by $M^{\frac{1}{2}}$.

Proposition 5 *If $S \subset \mathbf{R}^m$ is an ellipsoid of the form $S = \{y \in \mathbf{R}^m : \langle y, My\rangle \le 1\}$, then $\mu_S(x) = \|M^{\frac{1}{2}}x\|_2$ and $\mu_{S^\circ}(x) = \|M^{-\frac{1}{2}}x\|_2$.*

Given $S = \{y \in \mathbf{R}^m : \langle y, My\rangle \le 1\}$, we obtain $\mu_{S^\circ}(u + \alpha v) = \|M^{-\frac{1}{2}}u + \alpha M^{-\frac{1}{2}}v\|_2$. The infimum of $\mu_{S^\circ}(u + \alpha v)$ is then achieved at the least square solution of the linear equation $\alpha M^{-\frac{1}{2}}v = -M^{-\frac{1}{2}}u$, which is given by $\alpha = -v'M^{-1}vv'M^{-1}u$. Thus

$$\sup_{\alpha\in\mathbf{R}} \frac{1}{\mu_{S^\circ}(u + \alpha v)} = \frac{(v'M^{-1}v)^{\frac{1}{2}}}{[u'M^{-1}uv'M^{-1}v - (u'M^{-1}v)^2]^{\frac{1}{2}}}. \quad (18)$$

Case II S is a parallelotope

The following proposition shows that if S is a parallelotope, then μ_S is a weighted Hölder ∞-norm and μ_{S° is a weighted Hölder 1-norm.

Proposition 6 *Assume that $S \subset \mathbf{R}^m$ is a parallelotope with basis $\{x_1, x_2, \ldots, x_m\}$. Let $H := [x_1\, x_2 \cdots x_m] \in \mathbf{R}^{m\times m}$ and $[y_1\, y_2 \cdots y_m] := H'^{-1}$. Then S° is a crosspolytope with basis $\{y_1, y_2, \ldots, y_m\}$. Moreover, $\mu_S(x) = \|H^{-1}x\|_\infty$ and $\mu_{S^\circ}(x) = \|H'x\|_1$.*

Given a parallelotope $S \subset \mathbf{R}^m$ with basis $\{x_1, x_2, \ldots, x_m\}$, we obtain $\mu_{S^\circ}(u + \alpha v) = \|H'u + \alpha H'v\|_1$, where $H = [x_1\, x_2 \cdots x_m]$. Let $H'u = [\zeta_1\, \zeta_2 \cdots \zeta_m]'$ and $H'v = [\eta_1\, \eta_2 \cdots \eta_m]'$; then

$$\|H'u + \alpha H'v\|_1 = |\zeta_1 + \alpha\eta_1| + |\zeta_2 + \alpha\eta_2| + \cdots + |\zeta_m + \alpha\eta_m|. \quad (19)$$

A continuous function on $\mathbf{R}$ is called polygonal (or piecewise linear) if there exist finite points $\alpha_1, \alpha_2, \ldots, \alpha_p \in \mathbf{R}$ with $\alpha_1 < \alpha_2 < \cdots < \alpha_l$ such that the function is linear on $(-\infty, \alpha_1]$, $[\alpha_l, \infty)$ and $[\alpha_i, \alpha_{i+1}]$, $i = 1, 2, \ldots, l-1$. In this case the points $\alpha_1, \alpha_2, \ldots, \alpha_l$ are called division points. If (19) is considered to be a function of α, then it is a polygonal function with at most m division points. The set of division points is just $\{-\frac{\zeta_i}{\eta_i} : i = 1, 2, \ldots, m \text{ and } \eta_i \neq 0\}$. The supremum and infimum of a polygonal function on $\mathbf{R}$ can only happen at ∞, $-\infty$ or one of its division points. Since $\|H'u + \alpha H'v\|_1$ goes to infinity as α goes to ∞ or $-\infty$, its infimum can only be achieved at one of its division points. This proves that

$$\begin{aligned}\sup_{\alpha\in\mathbf{R}} \frac{1}{\mu_{S^\circ}(u + \alpha v)} &= \sup_{\alpha\in\mathbf{R}} \frac{1}{\|H'u + \alpha H'v\|_1} \\ &= \max\{\frac{1}{\|H'u + \alpha H'v\|_1} : \alpha \in \{-\frac{\zeta_i}{\eta_i} : i = 1, \ldots, m \text{ and } \eta_i \neq 0\}\} \quad (20)\end{aligned}$$

To compute (20), we only need to compute the 1-norm of at most m vectors in $\mathbf{R}^m$.

Case III S is a crosspolytope

The following proposition shows that if S is a crosspolytope, then μ_S is a weighted Hölder 1-norm and μ_{S° is a weighted Hölder ∞-norm.

Proposition 7 *Assume that $S \subset \mathbf{R}^m$ is a crosspolytope with basis $\{x_1, x_2, \ldots, x_m\}$. Let $H := [x_1\, x_2 \cdots x_m] \in \mathbf{R}^{m\times m}$ and $[y_1\, y_2 \cdots y_m] := H'^{-1}$. Then S° is a parallelotope with basis $\{y_1, y_2, \ldots, y_m\}$. Moreover, $\mu_S(x) = \|H^{-1}x\|_1$ and $\mu_{S^\circ} = \|H'x\|_\infty$.*

Given a crosspolytope $S \subset \mathbf{R}^m$ with basis $\{x_1, x_2, \ldots, x_m\}$, we obtain $\mu_{S^\circ}(u + \alpha v) = \|H'u + \alpha H'v\|_\infty$, where $H = [x_1\, x_2 \cdots x_m]$. Let $H'u = [\zeta_1\, \zeta_2 \cdots \zeta_m]'$ and $H'v = [\eta_1\, \eta_2 \cdots \eta_m]'$; then

$$\|H'u + \alpha H'v\|_\infty = \max\{|\zeta_1 + \alpha\eta_1|, |\zeta_2 + \alpha\eta_2|, \ldots, |\zeta_m + \alpha\eta_m|\}. \quad (21)$$

It is easy to see that $\|H'u + \alpha H'v\|_\infty$ is also a polygonal function of α which goes to ∞ as α goes to ∞ or $-\infty$, so its infimum is achieved at one of its division points. However it appears that the division points of $\|H'u + \alpha H'v\|_\infty$ can not be obtained as easily as those of $\|H'u + \alpha H'v\|_1$. Note that at any division point of $\|H'u + \alpha H'v\|_\infty$, we must have $|\zeta_i + \alpha\eta_i| = |\zeta_j + \alpha\eta_j|$ for some $i, j = 1, 2, \ldots, m$ and $i \neq j$. So the set of division point is contained in the following set

$$\Lambda \;=\; \{\alpha : \zeta_i + \alpha\eta_i = \zeta_j + \alpha\eta_j,\ 1 \le i < j \le m\}$$

$$\cup\{\alpha : \zeta_i + \alpha\eta_i = -\zeta_j - \alpha\eta_j,\ 1 \le i < j \le m\}$$
$$= \{-\frac{\zeta_i - \zeta_j}{\zeta_i - \zeta_j} : 1 \le i < j \le m \text{ and } \eta_i - \eta_j \ne 0\}$$
$$\cup\{-\frac{\zeta_i + \zeta_j}{\eta_i + \eta_j} : 1 \le i < j \le m \text{ and } \eta_i + \eta_j \ne 0\}.$$

This proves that if S is a crosspolytope, then

$$\sup_{\alpha\in\mathbf{R}} \frac{1}{\mu_{S^\circ}(u+\alpha v)} = \sup_{\alpha\in\mathbf{R}} \frac{1}{\|H'u + \alpha H'v\|_\infty} = \max\{\frac{1}{\|H'u + \alpha H'v\|_\infty} : \alpha \in \Lambda\}. \quad (22)$$

In the worst case, Λ has $m(m-1)$ elements, while the number of the division points of $\|H'u+\alpha H'v\|_\infty$ may be much less than $m(m-1)$. It is possible to devise a search scheme to find the division points, but this requires extra computational effort. Thus formula (22) should be used, at least in the case when m is not too large.

6 Polytopes

The problem considered in this section is to compute (16) in the case when S is a absorbing polytope, i.e. a polytope which contains the origin as an interior point.

A polytope is usually represented either as the convex hull of a finite nonempty set of points in $\mathbf{R}^m$, which is known as the *internal representation*, or as the intersection of a finite number of closed halfspaces, which is known as the *external representation*. The convex hull of an arbitrary finite set of points is not necessarily an absorbing polytope. A necessary and sufficient condition for $\text{co}\{x_1, x_2, \ldots, x_l\}$ to be an absorbing polytope is that some of $\langle x, x_i\rangle$ are strictly positive and some of $\langle x, x_i\rangle$ are strictly negative for any nonzero $x \in \mathbf{R}^m$. Similarly, the intersection of an arbitrary finite number of halfspaces, which is a polyhedral set, is not necessarily an absorbing polytope. A necessary and sufficient condition for $\bigcap_{i=1}^l\{x \in \mathbf{R}^m : \langle x, y_i\rangle \le \alpha_i\}$ to be an absorbing polytope is that it must be possible for each of the halfspaces $\{x \in \mathbf{R}^m : \langle x, y_i\rangle \le \alpha_i\}$ to be rewritten as $\{x \in \mathbf{R}^m : \langle x, z_i\rangle \le 1\}$ and $\text{co}\{z_1, z_2, \ldots, z_l\}$ is an absorbing polytope.

Given an absorbing polytope S with either internal or external representation, it is not a trivial task to find its other representation. However, it is very easy to find its polar in the other representation.

Assume that

$$S = \text{co}\{x_1, x_2, \ldots, x_l\},$$

where $x_1, x_2, \ldots, x_l \in \mathbf{R}^m$. Then by definition

$$\begin{aligned} S^\circ &= \{y \in \mathbf{R}^m : \langle y, x\rangle \le 1, \forall x \in S\} \\ &\subset \{y \in \mathbf{R}^m : \langle y, x_i\rangle \le 1, i = 1, 2, \ldots l\} \\ &= \bigcap_{i=1}^l \{y \in \mathbf{R}^m : \langle y, x_i\rangle \le 1\}. \end{aligned}$$

On the other hand, if $\langle y, x_i\rangle \le 1$ for all $i = 1, 2, \ldots, l$, then

$$\sum_{i=1}^l \lambda_i\langle y, x_i\rangle = \langle y, \sum_{i=1}^l \lambda_i x_i\rangle \le 1$$

for all $\lambda_i \ge 0, i = 1, 2, \ldots, l$, with $\sum_{i=1}^l \lambda_i$. This implies $\langle y, x\rangle \le 1$ for all $x \in S$. Therefore, it follows that

$$S^\circ = \bigcap_{i=1}^l \{y \in \mathbf{R}^m : \langle y, x_i\rangle \le 1\}.$$

Assume now that

$$S = \bigcap_{i=1}^l \{x \in \mathbf{R}^m : \langle x, y_i\rangle \le 1\}$$

where $y_1, y_2, \ldots y_l \in \mathbf{R}^m$. Let $T = \text{co}\{y_1, y_2, \ldots, y_l\}$. Then $T^\circ = S$. Since T is also an absorbing polytope, it follows that $S^\circ = T^{\circ\circ} = T$.

Case I The internal representation of S is given

Let

$$S = \text{co}\{x_1, x_2, \ldots, x_l\}.$$

Then

$$S^\circ = \bigcap_{i=1}^l \{y \in \mathbf{R}^m : \langle y, x_i\rangle \le 1\}$$

and

$$\begin{aligned} \mu_{S^\circ}(u+\alpha v) &= \inf\{\beta > 0 : \langle \beta^{-1}(u+\alpha v), x_i\rangle \le 1, i = 1, 2, \ldots, l\} \\ &= \inf\{\beta > 0 : \langle u+\alpha v, x_i\rangle \le \beta, i = 1, 2, \ldots, l\} \\ &= \max\{\langle u+\alpha v, x_i\rangle : i = 1, 2, \ldots, l\} \\ &= \max\{\langle u, x_i\rangle + \alpha\langle v, x_i\rangle : i = 1, 2, \ldots l\}. \end{aligned}$$

Again we see that $\mu_{S^\circ}(u+\alpha v)$ is a polygonal function of α and its infimum can only happen at its division points or at $\pm\infty$. The fact that S is absorbing implies that some of $\langle v, x_i\rangle$ are strictly positive and some of $\langle v, x_i\rangle$ are strictly negative. Hence μ_{S° goes to infinity as α goes to $\pm\infty$. Consequently, $\mu_{S^\circ}(u+\alpha v)$ achieves its infimum at one of its division points. The division points of $\mu_{S^\circ}(u+\alpha v)$ are contained in the set

$$\begin{aligned} \Lambda &= \{\alpha : \langle u, x_i\rangle + \alpha\langle v, x_i\rangle = \langle u, x_j\rangle + \alpha\langle v, x_j\rangle, 1 \le i < j \le l\} \\ &= \{-\frac{\langle u, x_i\rangle - \langle u, x_j\rangle}{\langle v, x_i\rangle - \langle v, x_j\rangle} : 1 \le i < j \le l \text{ and } \langle v, x_i\rangle - \langle v, x_j\rangle \ne 0\}. \end{aligned}$$

This shows that if S is a polytope with an internal representation, then

$$\sup_{\alpha\in\mathbf{R}} \frac{1}{\mu_{S^\circ}(u+\alpha v)} = [\min_{\alpha\in\Lambda}\{\max\{\langle u, x_i\rangle + \alpha\langle v, x_i\rangle : i = 1, 2, \ldots, l\}\}]^{-1}. \quad (23)$$

Case II The external representation of S is given

Let

$$S = \bigcap_{i=1}^l \{x \in \mathbf{R}^m : \langle x, y_i\rangle \le 1\}.$$

Then

$$S^\circ = \text{co}\{y_1, y_2, \ldots, y_l\}$$

and

$$\begin{aligned} \mu_{S^\circ}(u+\alpha v) &= \inf\{\beta > 0 : \sum_{i=1}^l \lambda_i y_i = \beta^{-1}(u+\alpha v), \\ &\qquad \lambda_i \ge 0, i = 1, 2, \ldots, l, \text{ and } \sum_{i=1}^l \lambda_i = 1\} \\ &= \inf\{\beta > 0 : \sum_{i=1}^l \lambda_i y_i = u + \alpha v, \\ &\qquad \lambda_i \ge 0, i = 1, 2, \ldots, l, \text{ and } \sum_{i=1}^l \lambda_i = \beta\} \\ &= \inf\{\sum_{i=1}^l \lambda_i : \sum_{i=1}^l \lambda_i y_i = u + \alpha v \\ &\qquad \text{and } \lambda_i \ge 0, i = 1, 2, \ldots, l\}. \end{aligned}$$

It is easy to see that "inf" in above equation can be replaced by "min". Therefore $\inf_{\alpha\in\mathbf{R}} \mu_{S^\circ}(u+\alpha v)$ can be obtained by solving the following linear programming problem:

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^l \lambda_i \\ \text{subject to} & -\alpha v + \sum_{i=1}^l \lambda_i y_i = u \\ \text{and} & \lambda_i \ge 0 \end{array} \quad (24)$$

with respect to variables α and λ_i, $i = 1, 2, \ldots, l$.

The solution in this case is not as simple as in case I where no linear programming is needed. It is not obvious now whether a simpler solution exists in this case. We have mentioned in Section 3 that the problem to find quantity (15), which is the problem we are trying to solve, can be directly reduced to a linear programming problem (cf. [17]). The linear programming problem (24) has an advantage over the one which can be obtained directly from (15), in that it is almost in the standard form which can be solved by the simplex method ([16]) and it results a standard linear programming problem with fewer variables and constraints.

7 Search Problem

Recall that our major problem of this paper is to compute the robustness measure ρ defined as a function of a 4-tuple $(F, g, \mathbf{C}_g, S)$ in (7)-(10). It is shown in Section 3 that $\rho = \inf_{s\in\partial\mathbf{C}_g} \tau(s)$, where $\tau(s)$ is a function from $\partial\mathbf{C}_g$ to $\overline{\mathbf{R}}^+$ which can be evaluated at any fixed $s \in \partial\mathbf{C}_g$ by using the various methods developed in Section 4-6, if S is an bounded absorbing convex set. The quantity ρ can then be obtained by a "brute force" search over $\partial\mathbf{C}_g$. i.e. choose a finite subset of $\partial\mathbf{C}_g$ which is sufficiently "dense" in $\partial\mathbf{C}_g$ and find the minimum of $\tau(s)$ over all points in this finite subset. Since $\partial\mathbf{C}_g$ is usually an one-dimensional curve in $\overline{\mathbf{C}}$, this "brute force" search is considered to be numerically permissible in engineering applications. However, two numerical difficulties may occur during the search. The first difficulty is that when $\partial\mathbf{C}_g$ is unbounded (or virtually unbounded, i.e. part of $\partial\mathbf{C}_g$ is to big), it may be impossible (or unrealistic) to have a finite and sufficiently "dense" subset in $\partial\mathbf{C}_g$. The second difficulty occurs due to the fact that $\tau(s)$ is generally not a continuous function on $\partial\mathbf{C}_g$, so it may be easy to miss the true infimum when a "brute force" search over $\partial\mathbf{C}_g$ is carried out. This section discusses the methods to deal with these difficulties. The convex set S is assumed to be bounded and absorbing throughout this section.

The first difficulty can be overcome completely using a "tricky" but simple method. If $(F, g, \mathbf{C}_g, S)$ is given, then $\partial\mathbf{C}_g$, the boundary of the stability region, and $p(s,k)$, the affine map from $\mathbf{R}^m$ to $\mathcal{P}$, are also given, and the function τ is defined in (11). Let

$$J = \begin{bmatrix} & & & 1 \\ & 0 & 1 & \\ & & \cdot & \\ & \cdot & & \\ & \cdot & 0 & \\ 1 & & & \end{bmatrix}$$

and $\hat{F} = JF$, $\hat{g} = Jg$. Define a new map $\hat{p}(s,k)$ from $\mathbf{R}^m$ to $\mathcal{P}$ by

$$\hat{p}(s,k) = [s^n\ s^{n-1} \cdots 1](\hat{F}k + \hat{g}),$$

a reciprocal set $\partial\hat{\mathbf{C}}_g$ of $\partial\mathbf{C}_g$ by

$$\partial\hat{\mathbf{C}}_g = \{s \in \overline{\mathbf{C}} : \frac{1}{s} \in \partial\mathbf{C}_g\}$$

and a new function $\hat{\tau}$: $\partial\hat{\mathbf{C}}_g \to \overline{\mathbf{R}}^+$ by

$$\hat{\tau}(s) = \inf\{\mu_S(k) : k \in \mathcal{K} \text{ and } \hat{p}(s,k) = 0\}.$$

Since $\hat{p}(s,k) = 0$ if and only if $p(\frac{1}{s}, k)$, we obtain

$$\tau(\frac{1}{s}) = \hat{\tau}(s)$$

Therefore

$$\begin{aligned} \rho &= \inf_{s\in\partial\mathbf{C}_g} \tau(s) \\ &= \min\{\inf_{s\in\partial\mathbf{C}_g\cap\mathbf{U}} \tau(s), \inf_{s\in\partial\mathbf{C}_g\cap\mathbf{U}^c} \tau(s)\} \\ &= \min\{\inf_{s\in\partial\mathbf{C}_g\cap\mathbf{U}} \tau(s), \inf_{s\in\partial\hat{\mathbf{C}}_g\cap\mathbf{U}} \tau(\frac{1}{s})\} \\ &= \min\{\inf_{s\in\partial\mathbf{C}_g\cap\mathbf{U}} \tau(s), \inf_{s\in\partial\hat{\mathbf{C}}_g\cap\mathbf{U}} \hat{\tau}(s)\}. \end{aligned}$$

where $\mathbf{U}$ is the closed unit disk in $\mathbf{C}$ and $\mathbf{U}^c$ is the complement of $\mathbf{U}$ in $\overline{\mathbf{C}}$.

The function $\hat{\tau}$ has exactly the same form as τ except that F and g are replaced by $\hat{F}$ and $\hat{g}$; it can be evaluated at each $s \in \partial\hat{\mathbf{C}}_g$ using the same way as τ is at each point $s \in \partial\mathbf{C}_g$. This shows that ρ can be obtained by two independent searches and each of these searches is carried out over a "small" bounded set.

Some thought about the treatment of the second difficulty is given as follows.

Recall the expression of $\tau(s)$ given by (14) and the definition of vectors $u(s)$, $v(s)$. Define

$$\begin{aligned} \partial\mathbf{C}_{g1} &= \{s \in \partial\mathbf{C}_g : u(s) \neq 0 \text{ and } v(s) = 0\} \\ \partial\mathbf{C}_{g2} &= \{s \in \partial\mathbf{C}_g : \operatorname{rank}[u(s)\ v(s)] = 2\}. \end{aligned}$$

It is known from Theorem 1 that $\tau(s) = \infty$ for all $s \in \partial\mathbf{C}_g \backslash \{\partial\mathbf{C}_{g1} \cup \partial\mathbf{C}_{g2}\}$. Hence

$$\inf_{s\in\partial\mathbf{C}_g} \tau(s) = \inf_{s\in\partial\mathbf{C}_{g1}\cup\partial\mathbf{C}_{g2}} \tau(s) = \min\{\inf_{s\in\partial\mathbf{C}_{g1}} \tau(s), \inf_{s\in\partial\mathbf{C}_{g2}} \tau(s)\}.$$

Denote by $\tau|_{\partial\mathbf{C}_{gi}}$, $i = 1,2$, the restrictions of τ to $\partial\mathbf{C}_{gi}$, $i = 1,2$, respectively.

Lemma 2 *$\tau|_{\partial\mathbf{C}_{gi}}$, $i = 1,2$, are continuous functions.*

Lemma 4 implies that if we can identify $\partial\mathbf{C}_{g1}$ and $\partial\mathbf{C}_{g2}$ from ∂C_g and carry out the "brute force" searches to find $\inf_{s\in\partial\mathbf{C}_{g1}} \tau(s)$ and $\inf_{s\in\partial\mathbf{C}_{g2}} \tau(s)$ separately, then we will have very little chance to miss the true infimum of $\tau(s)$ over $\partial\mathbf{C}_g$.

The set $\partial\mathbf{C}_{g1}$ is contained in the intersection of $\partial\mathbf{C}_g$ and

$$\{s \in \overline{\mathbf{C}} : \Im(\frac{F'[s^n\ s^{n-1} \cdots 1]'}{[s^n\ s^{n-1} \cdots 1]g}) = 0\}.$$

The latter set contains $\mathbf{R} \cup \{\infty\}$ and usually consists of some curves and points in $\overline{\mathbf{C}}$. Hence the set $\partial\mathbf{C}_{g1}$ usually contains discrete points in $\partial\mathbf{C}_g$. If these points can be sorted out and special care is taken when the "brute force" search is carried out, then the search should reliably reach the real infimum. There are straightforward but tedious ways to sort out these points from $\partial\mathbf{C}_g$. In most cases when m, the dimension of k, is greater that 1, these points are just the elements of the intersection of $\partial\mathbf{C}_g$ and $\mathbf{R} \cup \{\infty\}$.

8 An Example

The following polynomial is considered in [7]:

$$\begin{aligned} p(s,k) &= a_0(k)s^4 + a_1(k)s^3 + a_2(k)s^2 + a_3(k)s + a_4(k) \\ &= [s^4\ s^3\ s^2\ s\ 1]\left(\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 10.75 & 0.75 & 7 & 0.25 \\ 32.5 & 7.5 & 12 & 0.5 \\ 18.75 & 18.75 & 10 & 0.5 \end{bmatrix} k + \begin{bmatrix} 1 \\ 12 \\ 47 \\ 70 \\ 50 \end{bmatrix}\right), \end{aligned}$$

where $k \in \mathbf{R}^4$ is uncertain.

The roots of $p(s,0)$ are -5, -5, $-1 \pm j$. If the desired stability region is assumed to be the open left half part of the complex plane, then the polynomial is nominally stable and the stability robustness measure is given by

$$\rho = \inf_{\omega\in\mathbf{R}} \tau(j\omega) = \min\{\inf_{\omega\in[0,1]} \tau(j\omega), \inf_{\omega\in[0,1)} \hat{\tau}(j\omega)\}$$

where

$$\begin{aligned} \tau(j\omega) &= \inf\{\|k\| : k \in \mathcal{K} \text{ and } p(j\omega,k) = 0\} \\ \hat{\tau}(j\omega) &= \inf\{\|k\| : k \in \mathcal{K} \text{ and } \hat{p}(j\omega,k) = 0\} \end{aligned}$$

and

$$\hat{p}(s,k) = a_4(k)s^4 + a_3(k)s^3 + a_2(k)s^2 + a_1(k)s + a_0(k)$$

$$= [s^4\ s^3\ s^2\ s\ 1]\left(\begin{bmatrix} 18.75 & 18.75 & 10 & 0.5 \\ 32.5 & 7.5 & 12 & 0.5 \\ 10.75 & 0.75 & 7 & 0.25 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} k + \begin{bmatrix} 50 \\ 70 \\ 47 \\ 12 \\ 1 \end{bmatrix}\right).$$

Case I S is the unit ball of the Hölder 2-norm

In this case, $\tau(j\omega)$ and $\hat{\tau}(j\omega)$ can be evaluated by using (17) and (18). We obtain

$$\inf_{\omega\in[0,1]} \tau(j\omega) = \tau(j\omega)|_{\omega=0} = 1.76$$
$$\inf_{\omega\in[0,1)} \hat{\tau}(j\omega) = \lim_{\omega\to 1} \hat{\tau}(j\omega) = 2.01.$$

Therefore $\rho = 1.76$.

Case II S is the unit ball of the Hölder ∞-norm

In this case, $\tau(j\omega)$ and $\hat{\tau}(j\omega)$ can be evaluated by using (17) and (20). We obtain

$$\inf_{\omega\in[0,1]} \tau(j\omega) = \tau(j\omega)|_{\omega=0} = 1.04$$
$$\inf_{\omega\in[0,1)} \hat{\tau}(j\omega) = \hat{\tau}(j\omega)|_{\omega=0.77} = 2.01.$$

Therefore $\rho = 1.04$.

Case III S is the unit ball of the Hölder 1-norm

In this case, $\tau(j\omega)$ and $\hat{\tau}(j\omega)$ can be evaluated by using (17) and (22). We obtain

$$\inf_{\omega\in[0,1]} \tau(j\omega) = \tau(j\omega)|_{\omega=0.71} = 2.00$$
$$\inf_{\omega\in[0,1)} \hat{\tau}(j\omega) = \lim_{\omega\to 1} \hat{\tau}(j\omega) = 2.67.$$

Therefore $\rho = 2.00$.

Case IV S is a simplex in $\mathbf{R}^4$

A polytope in $\mathbf{R}^m$ is called a *simplex* if it has exactly $m+1$ vertices. If S is an absorbing simplex in $\mathbf{R}^4$ given by

$$S = \mathrm{co}\left\{\begin{bmatrix} 4 \\ -1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 4 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 4 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \end{bmatrix}\right\},$$

then $\tau(j\omega)$ and $\hat{\tau}(j\omega)$ can be evaluated by using (16) and (23). We obtain

$$\inf_{\omega\in[0,1]} \tau(j\omega) = \tau(j\omega)|_{\omega=0.97} = 0.549$$
$$\inf_{\omega\in[0,1)} \hat{\tau}(j\omega) = \lim_{\omega\to 1} \hat{\tau}(j\omega) = 0.550.$$

Therefore $\rho = 0.549$.

9 Conclusion

An approach based on the framework of convex analysis is developed in this paper to study the stability robustness of polynomials. This approach unifies and improves some recent results in the area of the stability robustness of polynomials, motivates and solves new related problems, and provides theoretical soundness and completeness to this area which is full of *ad hoc* methods. The approach is quite general; it allows the stability region in the complex plane to be an arbitrary open set and allows the convex set which contains the uncertain parameters to have an arbitrary shape. The power of this approach is not limited to the material presented in this paper; for example, the remarkable Kharitonov's theorem and the edge theorem can also be proved within the framework of this approach.

References

[1] A.C. Barlett, C.V. Hollot and H. Lin, "Root locations for an entire polytope of polynomials: it suffices to check the edges", *Math. Control Signals and Systems*, vol. 1, pp. 61-71, 1988.

[2] B.R. Barmish, "A generalization of Kharitonov's four-polynomial concept for the robust stability problems with linearly dependent coefficient perturbations", *IEEE Trans. Automat. Contr.*, vol. AC-34, pp. 157-165, 1989.

[3] B.R. Barmish, "New tools for robustness analysis", *Proc. IEEE CDC*, pp. 1-6, 1988.

[4] A. Brøndsted, *An Introduction to Convex Polytopes*, Springer-Verlag, New York, 1983.

[5] R.M. Biernacki, H. Hwang and S.P. Bhattacharyya, "Robust stability with structured real parameter perturbations", *IEEE Trans. Automat. Contr.*, vol AC-32, pp. 495-505, 1987.

[6] H. Chapellat and S.P. Bhattacharyya, "A generalization of Kharitonov's theorem: robust stability of interval plants", *IEEE Trans. Automat. Contr.*, vol. AC-34, pp. 306-311, 1989.

[7] M. Fu, "Polytopes of polynomials with zeros in a prescribed region: new criteria and algorithms", To appear in *Robustness in Identification and Control*, 1988.

[8] D. Hinrichsen and A.J. Prichard, "New robustness results for linear systems under real perturbations", *Proc. IEEE CDC*, pp. 1375-1379, 1988.

[9] M. Fu and B.R. Barmish, "Maximal unidirectional perturbation bounds for stability of polynomials and matrices", *Systems & Control Letters*, vol. 11, pp. 173-179, 1988.

[10] V.L. Kharitonov, "Asymptotic stability of an equilibrium position of a family of systems of linear differential equations", *Differencjalnyje Uravnenija*, vol. 14, pp. 2086-2088, 1978.

[11] L. Qiu and E.J. Davison, "A simple procedure for the exact stability robustness computation of polynomials with affine coefficient perturbations", *Proc. IEEE International Conference on Systems Engineering*, Dayton, Ohio, 1989, To appear in *Systems & Control Letters.*

[12] L. Qiu and E.J. Davison, "A unified approach for the stability robustness of polynomials in a convex set", Systems Control Group Report, No. 8912, Dept. of Electrical Eng., Univ. of Toronto, 1989.

[13] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.

[14] W. Rudin, *Functional Analysis*, McGraw-Hill Book Company, New York, 1973.

[15] H.L. Royden, *Real Analysis*, MacMillan Publishing Co., Inc., New York, 1968.

[16] D.M. Simmons, *Linear Programming for Operations Research*, Holden-Day, Inc., San Francisco, 1972.

[17] A. Tesi and A. Vicino, "Robustness analysis for uncertain dynamical systems with structured perturbations", *Proc. IEEE CDC*, pp. 519-525, 1988.