Automatica, Vol. 28, No. 5, pp. 945–959, 1992
Printed in Great Britain.

0005–1098/92 \$5.00 + 0.00
Pergamon Press Ltd

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

L. QIU‡ and E. J. DAVISON‡

A unified approach to compute the stability robustness measure of a convex set of polynomials is obtained using convex analysis techniques.

Key Words—Stability-robustness; polynomials; gauge functions; Hahn–Banach theorem; stability; robustness.

Abstract—Consider a polynomial $p(s, k)$ which is affine in the parameter k; assume that the vector k is uncertain but belongs to a bounded closed convex set which contains the origin as an interior point, 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 using the framework of convex analysis. 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. An example is included to illustrate the type of results which may be obtained. The approach developed in this paper, when specialized, leads to some of the available results in the literature; moreover, it generates many new and interesting results. Another important feature of this approach is that it provides a rich mathematical insight to the stability robustness problem of polynomials.

1. INTRODUCTION

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

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

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

$$a(k) := [a_0(k)\, a_1(k) \cdots a_n(k)]' = Fk + q. \quad (2)$$

In many applications, the parameter k is somewhat uncertain but is known to be contained in a given set $S \subset \mathbb{R}^m$. 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.

Polynomial models of the form (1)–(2) are encountered in many circumstances in control problems. For example, if a matrix $A \in \mathbb{R}^{n\times n}$ is subject to a unity-rank perturbation of the form

$$A + kb',$$

where $b \in \mathbb{R}^n$ is a known vector and $k \in \mathbb{R}^n$ is the uncertain perturbation, then the characteristic polynomial of the perturbed matrix has the form of (1)–(2). Another example is given by a closed loop MISO (similarly SIMO) system (Chapellat and Bhattacharyya, 1989); assume that a MISO system is described by the transfer function

$$\left[\frac{n_1(s)}{d(s)}\frac{n_2(s)}{d(s)}\cdots\frac{n_r(s)}{d(s)}\right],$$

where

$$d(s) := (d_0 + \delta d_0)s^p + (d_1 + \delta d_1)s^{p-1} + \cdots + (d_p + \delta d_p)$$

$$n_i(s) := (n_{i0} + \delta n_{i0})s^p + (n_{i1} + \delta n_{i1})s^{p-1} + \cdots + (n_{ip} + \delta n_{ip}), \quad i = 1, 2, \ldots, r,$$

and the entries of the parameter vector

$$k = [\delta d_0\, \delta d_1 \cdots \delta d_p]' \times [\delta n_{10}\, \delta n_{11} \cdots \delta n_{1p}]' \times \cdots \times [\delta n_{r0}\, \delta n_{r1} \cdots \delta n_{rp}]',$$

(where "×" means the Cartesian product) are uncertain and independent. If the system is controlled by a fixed linear time-invariant proper

* Received 5 December 1990; revised 5 June 1991; revised 4 October 1991; received in final form 22 February 1992. The original version of this paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor R. V. Patel under the direction of Editor H. Kwakernaak.

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

‡ Department of Electrical Engineering, University of Toronto, Toronto, Ontario, Canada M5S1A4; L. Qiu is currently with the Fields Institute for Research in Mathematical Sciences, University of Waterloo, Waterloo, Ontario, Canada N2L5Z5.

controller, the closed loop characteristic polynomial will be of the form (1)–(2).

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 $\mu_S:\mathbb{R}^m \rightarrow [0, \infty]$ such that $\mu_S(k) = \inf\{\alpha > 0 : k \in \alpha S\}$, where αS is the set $\{\alpha x : x \in S\}$, 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 hyper-rectangle. In this case, the polynomial $p(s, k)$ is called an interval polynomial. The remarkable theorem of Kharitonov (1978) 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 (Bialas and Garloff, 1985; Fu and Barmish, 1988).

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 Bartlett *et al.* (1988), Barmish (1989) and Chapellat and Bhattacharyya (1989), 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. Bartlett *et al.* (1988) using the concept of the "root space", and Barmish (1989), using the concept of the "value set", considered the robustness checking problem for the case when the set S is a polytope. Since Bartlett *et al.* (1988) need to check all the edges of S and Barmish (1989) needs to use all the vertices of S, the computation required is combinatorially explosive when these methods are applied to an important class of polytopes—parallelotopes. This is because a parallelotope in $\mathbb{R}^k$ in general has $k2^{k-1}$ edges and 2^k vertices. The maximal robustness problem is solved in Biernacki *et al.* (1987) and Hinrichsen and Pritchard (1988) 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 Fu (1989) by using the geometry of the value sets and in Tesi and Vicino (1990) by using a linear programming method. Fu (1989) also considered parallelotopes. It is shown that the method given in Fu (1989) for general polytopes can be simplified for the case when S is a parallelotope. Sarideréli and Kern (1987) studied the maximal robustness problem when the size of the parameter uncertainty is measured by the Hölder ∞-norm. The method in Fu (1989), when applied to the parallelotope case, and the method in Sarideréli and Kern (1987) completely eliminate 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 Qiu and Davison (1989a) to solve the maximal robustness problem for the case when the stability robustness is measured by norms, which are just special kinds of gauges. The same problem is also studied in Hinrichsen and Pritchard (1989) independently using a similar method. In this paper, the method used in Qiu and Davison (1989a) is extended to the case when S is any bounded closed convex set containing the origin as an interior point. This method unifies all different cases studied before and simplifies the computation in many cases. It also generates many new and interesting results.

One of the motivations in studying general convex sets lies in the case when S is a polytope; this case has been studied extensively in the literature and arises in applications where the parameters have to meet certain linear constraints. On applying the proposed general approach, a method for the polytope case is obtained, which is distinct from the ones in the literature with certain advantages. Another motivation in studying the general convex set comes from the practical situation where the uncertain parameters enter the coefficients of the polynomial in a complicated non-affine way. This will in general lead to a non-convex set of polynomials. Since there is no general exact analysis method available for this situation, one usually uses a convex set to overbound the non-convex set so that a sufficient condition on robust stability can be obtained. Clearly, a rich family of exploitable convex sets can help to reduce the conservatism of this analysis.

The extension of the theory in Qiu and Davison (1989a) to general convex sets is not trivial. Since we want to accommodate convex sets which are not necessarily given by the unit balls of norms, many natural notions related to a norm, such as the dual norm, have to be extended. This is done in this paper by introducing some powerful tools and concepts in convex analysis, such as gauge, polar, dual gauge, as well as the Hahn–Banach theorem. In fact, we believe that the successful application of these concepts to the robust stability problem is one of the major contributions of this paper.

The structure of this paper is as follows. Section 2 provides some background of convex analysis required in the development of this paper. Section 3 formulates the problem considered and shows that the problem can be solved by finding the infimum of a real valued function over the complex variable on the boundary of the stability region. The main result concerning the evaluation of this function at each point on the boundary of the stability region is given in Section 4. Applications of this main result in some important special cases are considered in Sections 5–6. Some problems involved in the one-dimensional search to find the infimum of the real valued function are discussed in Section 7. Section 8 contains a numerical example and Section 9 is the Conclusion.

The following notation is used throughout this paper:

$\mathbb{R}$ (or $\mathbb{C}$)	is the field of real (or complex) numbers,
$\Re(s)$ (or $\Im(s)$)	is the real (or complex) part of $s \in \mathbb{C}$,
cl(S)	is closure of the set S,
co(S)	is convex hull of the set S.

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

2. PRELIMINARIES

Assume that the vector space $\mathbb{R}^m$ is equipped with the usual inner product $\langle x, y \rangle = x'y$ for $x, y \in \mathbb{R}^n$. This inner product induces a norm (the Hölder 2-norm) $\|x\|_2 = \langle x, x \rangle^{1/2}$ for $x \in \mathbb{R}^m$. Let S, T be subsets of $\mathbb{R}^m$ and let α, β be scalars in $\mathbb{R}$. We use αS to denote the set $\{\alpha x : x \in S\}$ and $S + T$ to denote the set $\{x + y : x \in S, y \in T\}$. As is well-known, a set S is said to be *convex* if $\alpha S + (1 - \alpha)S \subset S$ for all $\alpha \in [0, 1]$. This section introduces some results in convex analysis which will be used in the development to follow. The proofs of the results presented can be found in many standard references, (e.g. Brøndsted, 1983; Rockafellar, 1970; Rudin, 1973). Reference Qiu and Davison (1989b) also provides a complete and concise source of these proofs.

Definition 1. A set $S \subset \mathbb{R}^m$ 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 $x \in \mathbb{R}^m$, 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 $S = \{x \in \mathbb{R}^m : \langle x, Mx \rangle \leq 1\}$ for some positive definite matrix $M \in \mathbb{R}^{m \times m}$,
(e) polyhedral if S is the intersection of a finite number of sets of the form $\{x \in \mathbb{R}^m : \langle x, y \rangle \leq \alpha\}$ for some $y \in \mathbb{R}^m$ and some $\alpha \in \mathbb{R}$.

A set containing the origin as an interior point must be absorbing, but the converse is not necessarily true. However, if a set is convex and absorbing, it must contain the origin as an interior point.

If $y \neq 0$, a set of the form $\{x \in \mathbb{R}^m : \langle x, y \rangle \leq \alpha\}$ is usually called a *closed halfspace*; its boundary $\{x \in \mathbb{R}^m : \langle x, y \rangle = \alpha\}$ is called a *hyperplane*. In the case when $y = 0$, the set

$\{x \in \mathbb{R}^m : \langle x, y \rangle \leq \alpha\}$ is either $\emptyset$ or $\mathbb{R}^m$. 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 $S \subset \mathbb{R}^m$ be a closed convex set containing the origin.

Definition 2. The gauge (or Minkowski functional) of S is the function $\mu_S : \mathbb{R}^m \to [0, \infty]$, defined by

$$\mu_S(x) = \inf\{\alpha > 0 : x \in \alpha S\}.$$

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

Proposition 1. Suppose that $S \subset \mathbb{R}^m$ 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 $x \in \mathbb{R}^m$ if S is absorbing,
(e) μ_S is a norm if S is bounded, absorbing and balanced,
(f) $S = \{x \in \mathbb{R}^m : \mu_S(x) \leq 1\}$,
(g) μ_S is a continuous function 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, the Hölder p-norm $\|x\|_p = \left(\sum_{i=1}^m |x_i|^p\right)^{1/p}$ is the gauge of the set $S = \left\{x : \sum_{i=1}^m |x_i|^p = 1\right\}$. If $p = 1$ or ∞, the set S is polyhedral (actually a polytope) and if $p = 2$, the set S is ellipsoidal.

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

Definition 3. The polar of S, denoted by S^0, is defined by

$$S^0 = \{y \in \mathbb{R}^m : \langle x, y \rangle \leq 1, \forall x \in S\}.$$

Proposition 2. The polar S^0 of any set $S \subset \mathbb{R}^m$ is a closed convex set and contains the origin. If S is absorbing, then S^0 is bounded. If S is bounded, then S^0 is absorbing. If S is balanced, ellipsoidal or polyhedral, so is S^0.

Since S^0 is also a set in $\mathbb{R}^m$, it has a polar which is written as S^{00} instead of $(S^0)^0$. The set S^{00} is sometimes called the *bipolar* of S.

Proposition 3. For any set $S \subset \mathbb{R}^m$, $S^{00} = \text{cl}[\text{co}(\{0\} \cup S)]$.

The following corollary is immediately obtained from Proposition 3.

Corollary 1. If $S \subset \mathbb{R}^m$ is a closed convex set containing the origin, then $S^{00} = S$.

Since S^0 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 $S \subset \mathbb{R}^m$ containing the origin. Then the gauge μ_{S^0} of S^0 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 $S \subset \mathbb{R}^m$ containing the origin. Then its dual gauge μ_{S^0} satisfies

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

If μ_S is a norm, the right hand side of the above equality provides a 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 Hölder p-norm, then μ_{S^0} is the Hölder 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 (Royden, 1968).

Theorem 1. (Hahn–Banach Theorem.) Suppose

(a) $\phi : \mathbb{R}^m \to \mathbb{R}$ is a function satisfying $\phi(x+y) \leq \phi(x) + \phi(y)$ and $\phi(\alpha x) = \alpha \phi(x)$ for all $x, y \in \mathbb{R}^m$, $\alpha \geq 0$,
(b) f is a linear functional on a subspace V of $\mathbb{R}^m$ and $f(x) \leq \phi(x)$ for all $x \in V$.

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

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

3. PROBLEM FORMULATION

Let $\bar{\mathbb{C}}$ be the one point compactification of $\mathbb{C}$. It is known that $\bar{\mathbb{C}}$ is homeomorphic to the Riemann sphere. Partition $\bar{\mathbb{C}}$ into two disjoint subsets $\mathbb{C}_g$ and $\mathbb{C}_b$, i.e. $\bar{\mathbb{C}} = \mathbb{C}_g \dot{\cup} \mathbb{C}_b$, such that $\mathbb{C}_g$ is open. The reason why we consider $\bar{\mathbb{C}}$ instead of $\mathbb{C}$ is to avoid ambiguity at ∞; we want ∞ to belong to either $\mathbb{C}_g$ or $\mathbb{C}_b$ but not both. Let $\mathbb{P}$ be the space of all real polynomials with degree bounded by n. The number of roots of each polynomial in $\mathbb{P}$ is made to be n by supplementing an appropriate number of ∞'s as its roots. It follows from the results in Stewart (1975) that the roots of $p(s, k)$ are continuous in k. A polynomial in $\mathbb{P}$ is said to be *stable* if its roots are contained in $\mathbb{C}_g$. We use $p(s, k)$ to denote the image of $k \in \mathbb{R}^m$ under a fixed affine map from $\mathbb{R}^m$ to $\mathbb{P}$ which is completely characterized by a matrix $F \in \mathbb{R}^{(n+1)\times m}$ and a vector $g \in \mathbb{R}^{n+1}$ as

$$p(s, k) = [s^n s^{n-1} \cdots 1](Fk + g). \quad (3)$$

For any bounded absorbing closed convex set $S \subset \mathbb{R}^m$, let us define

$$\rho := \inf \{\mu_S(k) : k \in \mathbb{R}^m \text{ and } p(s, k) \text{ is unstable}\}. \quad (4)$$

The purpose of this paper is to find a procedure to compute ρ when F, g, $\mathbb{C}_g$ 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

$$\rho = \inf \{\mu_S(k) : k \in \mathbb{R}^m \text{ and } \exists s \in \mathbb{C}_b \text{ such that } p(s, k) = 0\}, \quad (5)$$

where $p(\infty, k) = 0$ means that $p(s, k)$ has a root at infinity.

Denote the boundary of $\mathbb{C}_g$ by $\partial\mathbb{C}_g$, i.e. $\partial\mathbb{C}_g = \mathbb{C}_b \cap \text{cl}(\mathbb{C}_g)$. If $p(s, k) = 0$ for some $s \in \mathbb{C}_b$, then the root locus of $p(s, \alpha k)$ as $\alpha \in [0, 1]$ contains a connected path which intersects both $\mathbb{C}_g$ and $\mathbb{C}_b$. The connectedness of this path assures that it must intersect $\partial\mathbb{C}_g$ as well. This shows that

$$\rho = \inf \{\mu_S(k) : k \in \mathbb{R}^m \text{ and } \exists s \in \partial\mathbb{C}_g \text{ such that } p(s, k) = 0\}, \quad (6)$$

which can be rewritten as

$$\rho = \inf_{s \in \partial\mathbb{C}_g} \{\inf \{\mu_S(k) : k \in \mathbb{R}^m \text{ and } p(s, k) = 0\}\}. \quad (7)$$

Define a function $\tau(s) : \partial\mathbb{C}_g \to [0, \infty]$ by

$$\tau(s) = \inf \{\mu_S(k) : k \in \mathbb{R}^m \text{ and } p(s, k) = 0\}. \quad (8)$$

It is seen from (7) that the computation of ρ can be accomplished in two phases. The first phase is to find $\tau(s)$ for any fixed $s \in \partial\mathbb{C}_g$. The second phase is to carry out a search over all points in $\partial\mathbb{C}_g$, which is usually a one-dimensional curve in $\bar{\mathbb{C}}$, to find $\inf_{s \in \partial\mathbb{C}_g} \tau(s)$.

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

$$[s^n s^{n-1} \cdots 1]Fk = -[s^n s^{n-1} \cdots 1]g, \quad (9)$$

for $s \neq \infty$ and

$$[1\, 0 \cdots 0]Fk = -[1\, 0 \cdots 0]g, \quad (10)$$

for $s = \infty$. The assumption on the stability of $p(s, 0)$ implies that the right hand side of (9) and (10) are nonzero. Let

$$w(s) := \begin{cases} -\dfrac{F'[s^n s^{n-1} \cdots 1]'}{[s^n s^{n-1} \cdots 1]g} & \text{if } s \neq \infty, \\[2ex] -\dfrac{F'[1\, 0 \cdots 0]'}{[1\, 0 \cdots 0]g}, & \text{if } s = \infty, \end{cases}$$

and let $u(s) := \Re[w(s)]$, $v(s) := \Im[w(s)]$. Then $u(s)$, $v(s) \in \mathbb{R}^m$ and the equation $p(s, k) = 0$ is equivalent to

$$u(s)'k = 1 \quad \text{and} \quad v(s)'k = 0. \quad (11)$$

Consequently, we have

$$\tau(s) = \inf \{\mu_S(k) : k \in \mathbb{R}^m \text{ and } u(s)'k = 1,\ v(s)'k = 0\}. \quad (12)$$

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

$$\inf \{\mu_S(k) : k \in \mathbb{R}^m \text{ and } u'k = 1,\ v'k = 0\}, \quad (13)$$

for any bounded absorbing closed convex set $S \subset \mathbb{R}^m$ and any $u, v \in \mathbb{R}^m$. 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 Tesi and Vicino (1990). 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 $\partial\mathbb{C}_g$, namely, choose a finite subset of $\partial\mathbb{C}_g$ which is sufficiently "dense" in $\partial\mathbb{C}_g$ and find the minimum of $\tau(s)$ over all s in this finite subset of $\partial\mathbb{C}_g$. Two numerical difficulties may occur here. The first difficulty is that when $\partial\mathbb{C}_g$ is unbounded it may be hard to define a finite subset of $\partial\mathbb{C}_g$ which is sufficiently "dense" in $\partial\mathbb{C}_g$. The second numerical difficulty occurs due to the fact that $\tau(s)$ in general is not a continuous function on $\partial\mathbb{C}_g$, so it may be easy to miss the true infimum when a brute force search over $\partial\mathbb{C}_g$ is carried out. Procedures to handle these two difficulties will be discussed in Section 8. Finally, we note that in most applications $\partial\mathbb{C}_g$ is symmetric to the real axis. Since $p(s, k) = 0$ if and only if $p(\bar{s}, k) = 0$, where $\bar{s}$ is the conjugate of s, it is sufficient in this case to carry out the search over the intersection of $\partial\mathbb{C}_g$ with the closed upper half of the complex plane.

4. THE MAIN RESULT

The purpose of this section is to analyze the quantity (13) which is essential in the computation of the function $\tau(\omega)$ defined in the preceding section. The main theorem given in this section simplifies the quantity (13) to a form more readily computable.

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

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

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

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

Then $\phi(x)$ satisfies the conditions in the Hahn–Banach Theorem. Consequently, there exists a linear functional $\bar{f}$ on $\mathbb{R}^m$ such that $\bar{f}(x) = f(x)$ for all $x \in V$, and $\bar{f}(x) \leq \phi(x)$ for all $x \in \mathbb{R}^m$. This means that there exists $k \in \mathbb{R}^m$ such that $k'u = 1$, $k'v = 0$, and $k'x \leq \phi(x)$ for all $x \in \mathbb{R}^m$. Furthermore, the inequality becomes an equality for some $x \in V$. By Proposition 4, we obtain

$$\begin{aligned} \mu_S(k) &= \mu_{S^{00}}(k) \\ &= \sup_{\substack{\alpha,\beta \in \mathbb{R} \\ \alpha u + \beta v \neq 0}} \frac{\alpha}{\mu_{S^0}(\alpha u + \beta v)}, \\ &= \max\left\{ \sup_{\substack{\alpha \in \mathbb{R} \\ u + \alpha v \neq 0}} \frac{1}{\mu_{S^0}(u + \alpha v)}, \right. \\ &\qquad \left. \sup_{\substack{\alpha \in \mathbb{R} \\ -u + \alpha v \neq 0}} \frac{1}{\mu_{S^0}(-u + \alpha v)} \right\}. \end{aligned}$$

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

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

and

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

The following theorem summarizes the above development.

Theorem 2. For any bounded absorbing closed convex set $S \subset \mathbb{R}^m$ and any $u, v \in \mathbb{R}^m$,

$$\inf\{\mu_S(k) : k \in \mathbb{R}^m \text{ and } u'k = 1,\ v'k = 0\}$$
$$= \begin{cases} \infty & \text{if } \operatorname{rank}[u\, v] \neq \operatorname{rank}\begin{bmatrix} u & v \\ 1 & 0 \end{bmatrix} \\ \max\left\{\dfrac{1}{\mu_{S^0}(u)}, \dfrac{1}{\mu_{S^0}(-u)}\right\} & \text{if } u \neq 0 \text{ and } v = 0 \\ \max\left\{\sup\limits_{\alpha \in \mathbb{R}} \dfrac{1}{\mu_{S^0}(u + \alpha v)}, \sup\limits_{\alpha \in \mathbb{R}} \dfrac{1}{\mu_{S^0}(-u + \alpha v)}\right\} & \\ & \text{if } \operatorname{rank}[u\, v] = 2. \end{cases} \tag{14}$$

In the special case when S is balanced, the gauges μ_S and μ_{S^0} are actually norms. Theorem 2 can be simplified as follows in this case, resulting in Theorem 1 of Qiu and Davison (1989a).

Corollary 2. If S is a balanced bounded absorbing closed convex set in $\mathbb{R}^m$, then for any $u, v \in \mathbb{R}^m$,

$$\inf\{\mu_S(k) : k \in \mathbb{R}^m \text{ and } u'k = 1,\ v'k = 0\}$$
$$= \begin{cases} \infty & \text{if } \operatorname{rank}[u\, v] \neq \operatorname{rank}\begin{bmatrix} u & v \\ 1 & 0 \end{bmatrix} \\ \dfrac{1}{\mu_{S^0}(u)} & \text{if } u \neq 0 \text{ and } v = 0 \\ \sup\limits_{\alpha \in \mathbb{R}} \dfrac{1}{\mu_{S^0}(u + \alpha v)} & \text{if } \operatorname{rank}[u\, v] = 2. \end{cases} \tag{15}$$

It is apparent that if μ_S and μ_{S^0} can be easily evaluated (which is the case in most practical applications), the computation of the right hand side of (14), which contains an optimization problem in a scalar variable α, is much simpler than the direct computation of the left hand side of (14), which contains an optimization problem in a vector variable k. It follows obviously that the critical problem in the computation of (14) is in the determination of $\sup\limits_{\alpha \in \mathbb{R}} \dfrac{1}{\mu_{S^0}(u + \alpha v)}$ for any $u, v \in \mathbb{R}^m$ with $\operatorname{rank}[u\, v] = 2$. Therefore, in the following when we refer to the computation of (14), we always mean the computation of

$$\sup_{\alpha \in \mathbb{R}} \frac{1}{\mu_{S^0}(u + \alpha v)},$$

or equivalently $\inf_{\alpha\in\mathbb{R}} \mu_{S^0}(u + \alpha v)$, for any $u,v \in \mathbb{R}^m$ with rank $[u\ v] = 2$.

If μ_{S^0} can be easily evaluated, the computation of $\inf_{\alpha\in\mathbb{R}} \mu_{S^0}(u + \alpha v)$ is actually very straightforward. It follows from Proposition 1 that $\mu_{S^0}(u + \alpha v)$, when considered as a function of α, is a continuous convex function on $\mathbb{R}$ and it goes to ∞ as α goes to $\pm\infty$. Consequently, $\inf_{\alpha\in\mathbb{R}} \mu_{S^0}(u + \alpha v) = \mu_{S^0}(u + \bar{\alpha} v)$ for a finite $\bar{\alpha}$ and any technique for a convex one-dimensional optimization, such as the Fibonacci search or the golden section search, can be used to find $\bar{\alpha}$ and $\inf_{\alpha\in\mathbb{R}} \mu_{S^0}(u + \alpha v)$. It will be shown in the following sections that if μ_S is the weighted version of one of the commonly used Hölder p-norms, or if S is a polytope with known vertices, then the computation of $\inf_{\alpha\in\mathbb{R}} \mu_{S^0}(u + \alpha v)$ becomes very simple, and no one-dimensional optimization is required.

In some applications, however, the evaluation of μ_{S^0} may not be easy; techniques which can be used to simplify the evaluation of μ_{S^0} are developed in Qiu and Davison (1989b).

5. ELLIPSOIDS, PARALLELOTOPES AND CROSSPOLYTOPES

This section deals with the computational problem of (14) 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 $x, y \in \mathbb{R}^m$ by $[x, y]$.

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

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

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

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 ith coordinate and 0 elsewhere. It will be 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 $\mathbb{R}^m$ are, respectively a parallelotope and a crosspolytope with basis $\{Hx_1, Hx_2, \ldots, Hx_m\}$.

Case I. *Ellipsoids*

The following proposition shows that if S is an ellipsoid, then μ_S and μ_{S^0} are both weighted Hölder 2-norms. We denote the unique positive definite square root of a positive definite matrix M by $M^{1/2}$.

Proposition 5. If $S \subset \mathbb{R}^m$ is an ellipsoid of the form $S = \{y \in \mathbb{R}^m : \langle y, My\rangle \le 1\}$, then $\mu_S(x) = \|M^{1/2}x\|_2$ and $\mu_{S^0}(x) = \|M^{-1/2}x\|_2$.

Proof. Since $\langle y, My\rangle \le 1$ if and only if $\langle M^{1/2}y, M^{1/2}y\rangle^{1/2} \le 1$, the ellipsoid S can also be expressed as

$$S = \{y \in \mathbb{R}^m : \langle M^{1/2}y, M^{1/2}y\rangle^{1/2} \le 1\} = \{y \in \mathbb{R}^m : \|M^{1/2}y\|_2 \le 1\}.$$

This implies that the gauge $\mu_S(x)$ is equal to $\|M^{1/2}x\|_2$. By Proposition 2, we know that S^0 is also an ellipsoid. Consider the set $T = \{y \in \mathbb{R}^m : \langle y, M^{-1}y\rangle \le 1\}$. It is clear that T is ellipsoidal and for $x \in S$ and $y \in T$,

$$\langle x, y\rangle = \langle M^{1/2}x, M^{-1/2}y\rangle \le \|M^{1/2}x\|_2\, \|M^{-1/2}y\|_2 = \langle x, Mx\rangle\langle y, M^{-1}y\rangle \le 1.$$

This implies $T \subset S^0$. On the other hand, if $y \ne T$, then $\langle y, M^{-1}y\rangle > 1$. Let $x = M^{-1}y/\langle y, M^{-1}y\rangle^{1/2}$. Then it is easy to check that $x \in S$ and $\langle x, y\rangle > 1$. This proves $T = S^0$. Hence, the gauge $\mu_{S^0}(x) = \|(M^{-1})^{1/2}x\|_2 = \|M^{-1/2}x\|_2$.

Given $S = \{y \in \mathbb{R}^m : \langle y, My\rangle \le 1\}$, we obtain $\mu_{S^0}(u + \alpha v) = \|M^{-1/2}u + \alpha M^{-1/2}v\|_2$. The infimum of $\mu_{S^0}(u + \alpha v)$ is then achieved at the least square solution of the linear equation $\alpha M^{-1/2}v = -M^{-1/2}u$, which is given by $\alpha = -(v'M^{-1}v)^{-1}u'M^{-1}u$. Thus

$$\inf_{\alpha\in\mathbb{R}} \mu_{S^0}(u + \alpha v) = \frac{[u'M^{-1}uv'M^{-1}v - (u'M^{-1}v)^2]^{1/2}}{(v'M^{-1}v)^{1/2}}. \quad (16)$$

Similar results for this case are obtained in Biernacki *et al.* (1987) and Hinrichsen and Pritchard (1988).

Case II. *Parallelotopes*

The following proposition shows that if S is a parallelotope, then μ_S is a weighted Hölder ∞-norm and μ_{S^0} is a weighted Hölder 1-norm.

Proposition 6. Assume that $S \subset \mathbb{R}^m$ is a parallelotope with basis $\{x_1, x_2, \ldots, x_m\}$. Let $H = [x_1 x_2 \cdots x_m] \in \mathbb{R}^{m\times m}$ and $[y_1 y_2 \cdots y_m] = H'^{-1}$. Then S^0 is a crosspolytope with basis $\{y_1, y_2, \ldots, y_m\}$. Moreover, $\mu_S(x) = \|H^{-1}x\|_\infty$ and $\mu_{S^0}(x) = \|H'x\|_1$.

Proof. Consider H as a linear map on $\mathbb{R}^m$ and let T be the parallelotope with basis $\{e_1, e_2, \ldots, e_m\}$, where e_i is a vector with 1 in the ith coordinate and 0 elsewhere. Then S is just the image of T under the map H. Since H is nonsingular, it follows that $x \in S$ if and only if $H^{-1}x \in T$. By noticing that the gauge of T is the Hölder ∞-norm, we obtain $\mu_S(x) = \mu_T(H^{-1}x) = \|H^{-1}x\|_\infty$.

By Proposition 4, we obtain

$$\begin{aligned}\mu_{S^0}(x) &= \inf\{\alpha \ge 0 : \langle y, x\rangle \le \|H^{-1}y\|_\infty, \forall y \in \mathbb{R}^m\}\\ &= \inf\{\alpha \ge 0 : \langle Hz, x\rangle \le \alpha\,\|z\|_\infty, \forall z \in \mathbb{R}^m\}\\ &\quad \text{(by letting } z = H^{-1}y)\\ &= \inf\{\alpha \ge 0 : \langle z, H'x\rangle \le \alpha\,\|z\|_\infty, \forall z \in \mathbb{R}^m\}\\ &= \|H'x\|_1.\end{aligned}$$

It remains to show that $S^0 = \{x \in \mathbb{R}^m : \|H'x\|_1 < 1\}$ is a crosspolytope. Let $y = H'x$. Then $x \in S^0$ if and only if $\|y\|_1 \le 1$. It follows that S^0 is the image of $R = \{y \in \mathbb{R}^m : \|y\|_1 \le 1\}$ under the linear map H'^{-1}. Since R is a crosspolytope, so is S^0.

Given a parallelotope $S \subset \mathbb{R}^m$ with basis $\{x_1, x_2, \ldots, x_m\}$, we obtain $\mu_{S^0}(u + \alpha v) = \|H'u + \alpha H'v\|_1$, 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

$$\|H'u + \alpha H'v\|_1 = |\zeta_1 + \alpha\eta_1| + |\zeta_2 + \alpha\eta_2| + \cdots + |\zeta_m + \alpha\eta_m|. \quad (17)$$

A continuous function on $\mathbb{R}$ is called *polygonal* (or *piecewise linear*) if there exists finite points $\alpha_1, \alpha_2, \ldots, \alpha_p \in \mathbb{R}$ with $\alpha_1 < \alpha_2 < \cdots < \alpha_l$ such that the function is affine on $(-\infty, \alpha_1]$, $[\alpha_l, \infty)$ and $[\alpha_i, \alpha_{i+1}]$, $i = 1, 2, \ldots, l-1$. If such points $\alpha_1, \alpha_2, \ldots, \alpha_l$ are chosen so that they have least number, then they are called vertices of the polygonal function. If (17) is considered to be a function of α, then it is a polygonal function with at most m vertices. The set of vertices is just $\left\{-\frac{\zeta_i}{\eta_i} : i = 1, 2, \ldots, m \text{ and } \eta_i \ne 0\right\}$. The supremum and infimum of a polygonal function on $\mathbb{R}$ can only occur at ∞, $-\infty$ or one of its vertices. Since $\|H'u + \alpha H'v\|_1$ goes to infinity as α goes to ∞ or $-\infty$, its infimum can only occur at one of its vertices. This shows that

$$\begin{aligned}\inf_{\alpha\in\mathbb{R}} \mu_{S^0}(u + \alpha v) &= \inf_{\alpha\in\mathbb{R}} \|H'u + \alpha H'v\|_1\\ &= \min\Big\{\|H'u + \alpha H'v\|_1 :\\ &\qquad \alpha \in \Big\{-\frac{\zeta_i}{\eta_i} : i = 1, 2, \ldots, m \text{ and } \eta_i \ne 0\Big\}\Big\}. \quad (18)\end{aligned}$$

To compute (18), we only need to compute the 1-norm of at most m vectors in $\mathbb{R}^m$.

The parallelotope case is also considered in Fu (1989), whose method is different from ours. A method given in Saridereli and Kern (1987), for the case when $\mu_s = \|\cdot\|_\infty$, has certain similarity with our method here although the derivation is completely different. It is to be noted that all of these three methods completely eliminate the "combinatorial explosion" possessed by some early methods which require the use of all of the vertices of the parallelotopes.

Case III. *Crosspolytopes*

The following proposition shows that if S is a crosspolytope, then μ_S is a weighted Hölder 1-norm and μ_{S^0} is a weighted Hölder ∞-norm.

Proposition 7. Assume that $S \subset \mathbb{R}^m$ is a crosspolytope with basis $\{x_1, x_2, \ldots, x_m\}$. Let $H = [x_1 x_2 \cdots x_m] \in \mathbb{R}^{m\times m}$ and $[y_1 y_2 \cdots y_m] = H'^{-1}$. Then S^0 is a parallelotope with basis $\{y_1, y_2, \ldots, y_m\}$. Moreover, $\mu_S(x) = \|H^{-1}x\|_1$ and $\mu_{S^0} = \|H'x\|_\infty$.

Proof. The proof follows the same line as the proof of Proposition 6 and hence is omitted.

Given a crosspolytope $S \subset \mathbb{R}^m$ with basis $\{x_1, x_2, \ldots, x_m\}$, we obtain $\mu_{S^0}(u + \alpha v) = \|H'u + \alpha H'v\|_\infty$, 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

$$\begin{aligned}&\|H'u + \alpha H'v\|_\infty\\ &\quad = \max\{|\zeta_1 + \alpha\eta_1|, |\zeta_2 + \alpha\eta_2|, \ldots,\\ &\qquad |\zeta_m + \alpha\eta_m|\}. \quad (19)\end{aligned}$$

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 occurs at one of its vertices. However it appears that the vertices of $\|H'u + \alpha H'v\|_\infty$ cannot be obtained as easily as those of $\|H'u + \alpha H'v\|_1$. Note that at any vertex of $\|H'u + \alpha H'v\|_\infty$, we must have $|\zeta_i + \alpha\eta_i| = |\zeta_i + \alpha\eta_j|$ for some $i, j = 1, 2, \ldots, m$

and $i \neq j$. So the set of vertices is contained in the following set

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

This shows that if S is a crosspolytope, then

$$\inf_{\alpha \in \mathbb{R}} \mu_{S^0}(u + \alpha v) = \inf_{\alpha \in \mathbb{R}} \|H'u + \alpha H'v\|_\infty = \min\{\|H'u + \alpha H'v\|_\infty : \alpha \in V\}. \quad (20)$$

In the worst case, V has $m(m-1)$ elements, while the number of the vertices 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 vertices, but this requires extra computational effort. Thus formula (20) 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 (14) in the case when S is an absorbing polytope, i.e. a polytope which contains the origin as an interior point. We first present some theoretical results on the facial structure of polytopes and their polars. These results give some insight into the relationship between a polytope and its polar. All the results are from (Brøndsted, 1983); hence the proofs are not included here. We then develop methods to compute (14) when the polytope S is represented either as the convex hull of a finite number of points or as the intersection of a finite number of halfspaces.

The *faces* of a polytope $S \subset \mathbb{R}^m$ are defined to be the intersections of S and the hyperplanes which do not intersect the interior of S. The set $\emptyset$ and S are also considered to be faces of S (called improper faces). The dimension of each face is defined to be the dimension of the affine space it spans; the dimensions of $\emptyset$ is assumed to be -1. The zero-dimensional faces of S are called the *vertices* of S, the one-dimensional faces of S are called the *edges* of S, and the $(\dim(S) - 1)$-dimensional faces are also called the *facets* of S. Denote the set of all faces of S (including $\emptyset$ and S) by $\mathbb{F}(S)$. Then $\mathbb{F}(S)$ becomes a partially ordered set if the set inclusion "$\subset$" is defined as the partial order in $\mathbb{F}(S)$. Using intuition, we may conjecture that $\mathbb{F}(S)$ is a finite set and that the intersection of two faces in $\mathbb{F}(S)$ is also a face in $\mathbb{F}(S)$. These two statements can be proved to be true and they imply that $\mathbb{F}(S)$ is a complete lattice (Szász, 1963), since the meet of two faces can be defined to be their intersection and the join of two faces defined to be the intersection of all (finite many) faces containing them. This lattice is called the *face-lattice* of S.

If S is an absorbing polytope, Proposition 2 shows that S^0, the polar of S, is also an absorbing polytope. The face-lattice of S^0 is denoted by $\mathbb{F}(S^0)$.

Recall (Szász, 1963) that a bijective map from one lattice to another is called an *dual isomorphism* if it reverses the partial order. Two lattices are said to be *dual isomorphic* to each other if there exists a lattice dual isomorphism between them.

Proposition 8.
(a) $\mathbb{F}(S)$ and $\mathbb{F}(S^0)$ are dual isomorphic.
(b) A dual isomorphism from $\mathbb{F}(S)$ to $\mathbb{F}(S^0)$ is given by

$$\phi(P) = \{y \in S^0 : \langle x, y \rangle = 1\ \forall x \in P\}.$$

(c) Let ψ be any dual isomorphism from $\mathbb{F}(S)$ onto $\mathbb{F}(S^0)$ and P be any face in $\mathbb{F}(S)$. Then $\dim(P) + \dim[\psi(P)] = n - 1$.

An advantage of the dual isomorphism given in Proposition 8(b) over any other possible dual isomorphism is that† $\phi(\phi(P)) = P$ for each $P \in \mathbb{F}(S)$. Hence $(P, \phi(P))$ forms a so-called mutually conjugate pair.

A polytope is usually represented either as the convex hull of a finite nonempty set of points in $\mathbb{R}^m$, 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 internal representation is also called the *vertex representation* since the vertices of polytope $\mathrm{co}\{x_1, x_2, \ldots, x_l\}$ are contained in the set $\{x_1, x_2, \ldots, x_l\}$. The external representation is also called the *facet representation* since the facets of polytope $\bigcap_{i=1}^{l} \{x \in \mathbb{R}^m : \langle x, y_i \rangle \le \alpha_i\}$ are contained in the union of hyperplanes $\{x \in \mathbb{R}^m : \langle x, y_i \rangle = \alpha_i\}$, $i = 1, 2, \ldots, l$. The convex hull of an arbitrary finite set of points is not necessarily an absorbing polytope. A necessary and sufficient condition for $\mathrm{co}\{x_1, x_2, \ldots, x_l\}$ 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 each nonzero $x \in \mathbb{R}^m$. 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

† Strictly speaking, the two ϕs in this expression are different; one is $\mathbb{F}(S) \to \mathbb{F}(S^0)$ and the other is $\mathbb{F}(S^0) \to \mathbb{F}(S)$. They share the same notation since they are defined using the same rule.

condition for $\bigcap_{i=1}^{l} \{x \in \mathbb{R}^m : \langle x, y_i \rangle \le \alpha_i\}$ to be an absorbing polytope is that it must be possible for each of the halfspaces $\{x \in \mathbb{R}^m : \langle x, y_i \rangle \le \alpha_i\}$ to be rewritten as $\{x \in \mathbb{R}^m : \langle x, z_i \rangle \le 1\}$ 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. This fact is actually implied by Proposition 8, for the facets (or vertices) of S^0 are just the image of the vertices (or facets) under any dual isomorphism from $\mathbb{F}(S)$ onto $\mathbb{F}(S^0)$.

Assume that

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

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

$$\begin{aligned} S^0 &= \{y \in \mathbb{R}^m : \langle y, x \rangle \le 1, \forall x \in S\} \\ &\subset \{y \in \mathbb{R}^m : \langle y, x_i \rangle \le 1, i = 1, 2, \ldots, l\} \\ &= \bigcap_{i=1}^{l} \{y \in \mathbb{R}^m : \langle y, x_i \rangle \le 1\}. \end{aligned}$$

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

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

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

$$S^0 = \bigcap_{i=1}^{l} \{y \in \mathbb{R}^m : \langle y, x_i \rangle \le 1\}.$$

Assume now that

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

where $y_1, y_2, \ldots y_l \in \mathbb{R}^m$. Let $T = \text{co}\{y_1, y_2, \ldots, y_l\}$. Then $T^0 = S$. Since T is also an absorbing polytope, it follows that $S^0 = T^{00} = T$.

A representation of a polytope

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

or

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

is said to be *irreducible* if

$$S \ne \text{co}\{\{x_1, x_2, \ldots, x_l\} \backslash x_j\},$$

or

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

for each $j = 1, 2, \ldots, l$. It is not assumed in this section that any representation of a polytope is irreducible. It is natural to believe, however, that if the representation is irreducible, then the computational complexity of the following proposed procedure to compute (14) would be minimized.

In the remaining part of this section, we develop methods for the computation of (14) when the polytope is represented either internally or externally.

Case I. *The internal representation of* S *is given*

Let

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

Then

$$S^0 = \bigcap_{i=1}^{l} \{y \in \mathbb{R}^m : \langle y, x_i \rangle \le 1\},$$

and

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

Again we see that $\mu_{S^0}(u + \alpha v)$ is a polygonal function of α and its infimum can only happen at its vertices 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 $\mu_{S^0}(u + \alpha v)$ goes to infinity as α goes to $\pm\infty$. Consequently, $\mu_{S^0}(u + \alpha v)$ achieves its infimum at one of its vertices. At each vertex of $\mu_{S^0}(u + \alpha v)$, there must be i,j with $1 \le i, j \le l$ and $i \ne j$ such that $\langle u, x_i \rangle + \alpha \langle v, x_i \rangle = \langle u, x_j \rangle + \alpha \langle v, x_j \rangle$. Therefore, the vertices of $\mu_{S^0}(u + \alpha v)$ are contained in the set

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

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

$$\inf_{\alpha \in \mathbb{R}} \mu_{S^0}(u + \alpha v) = \min_{\alpha \in V} \{\max\{\langle u, x_i \rangle + \alpha \langle v, x_i \rangle : i = 1, 2, \ldots, l\}\}. \quad (21)$$

The set V generally contains redundant points which are not the vertices of $\mu_{S^0}(u+\alpha v)$. The following result however can be used to reduce such redundancy. Two distinct vertices of a polytope are said to be *adjacent* if their convex hull is an edge of the polytope.

Proposition 9. The vertices of μ_{S^0} are contained in the set

$$V_1 = \left\{ -\frac{\langle u, x_i\rangle - \langle u, x_j\rangle}{\langle v, x_i\rangle - \langle v, x_j\rangle} : 1 \le i < j \le l,\ \langle v, x_i\rangle - \langle v, x_j\rangle \ne 0 \text{ and } x_i, x_j \text{ are adjacent vertices of } S \right\}.$$

Proof. We can assume without loss of generality that the representation

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

is irreducible, i.e. $\{x_i, i = 1, 2, \ldots, l\}$ is the set of all vertices of S. Let ϕ be the dual isomorphism given in Proposition 8(b). Then $\{\phi(x_i): i = 1, 2, \ldots, l\}$ is the set of all facets of S^0. The boundary of S^0, denoted by ∂S^0, is $\bigcup_{i=1}^{l} \phi(x_i)$. Let T be the union of all $(m-2)$-dimensional faces of S^0. Then $\partial S^0 \setminus T$ consists of the relative interior of all $\phi(x_i)$, $i = 1, 2, \ldots, l$. For each α, the ray starting at the origin and passing through $u + \alpha v$ intersects ∂S^0 at a point y_α. The point y_α is either in $\partial S^0 \setminus T$ or T. If y_α is in $\partial S^0 \setminus T$, then y_α is in the relative interior of $\phi(x_i)$ for some $1 \le i \le l$. This not only means that $\mu_{S^0}(u + \alpha v)$ is equal to $\langle u + \alpha v, x_i\rangle = \langle u, x_i\rangle + \alpha\langle v, x_i\rangle$, but also means that $\mu_{S^0}(u + \beta v)$ is equal to $\langle u, x_i\rangle + \beta\langle v, x_i\rangle$ for all β in a neighborhood of α. This implies that α is not a vertex of $\mu_{S^0}(u + \alpha v)$. Accordingly, if α is a vertice of $\mu_{S^0}(u + \alpha v)$, then $y_\alpha \in T$, i.e. $y_\alpha \in \phi(x_i) \cap \phi(x_j)$ for some $1 \le i, j \le l$ with $\phi(x_i) \cap \phi(x_j)$ being an $(m-2)$-dimensional face. Since ϕ is a dual isomorphism, the fact that $\phi(x_i) \cap \phi(x_j)$ is a $(m-2)$-dimensional face implies that the convex hull of x_i and x_j is a one-dimensional face, i.e. an edge.

Proposition 9 implies that V in (21) can be replaced by its subset V_1. We obtain

$$\inf_{\alpha\in\mathbb{R}} \mu_{S^0}(u + \alpha v) = \min_{\alpha\in V_1} \{\max\{\langle u, x_i\rangle + \alpha\langle v, x_i\rangle : i = 1, 2, \ldots, l\}\}. \tag{22}$$

Although Proposition 9 is theoretically interesting, extra computational effort is needed to test whether or not x_i, x_j are adjacent vertices, when we are given $S = \text{co}\,\{x_1, x_2, \ldots, x_l\}$. Hence, (21) may be more convenient to use in actual computation than (22).

The results for this case are new and have certain advantages over the results in Barmish (1989) and Fu (1989), e.g. we do not need a sweep over an auxiliary variable as is required in Barmish (1989), and we obtain closed form formulas which are easy to compute. It is also of interest to note that the information given by the edges of S plays an important role in the analysis, which gives an analogy between the edge theorem in Bartlett *et al.* (1988) and our results.

Case II. *The external representation of* S *is given*
Let

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

Then

$$S^0 = \text{co}\,\{y_1, y_2, \ldots, y_l\}$$

and

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

It is easy to see that "inf" in above equation can be replaced by "min". Therefore $\inf_{\alpha\in\mathbb{R}} \mu_{S^0}(u + \alpha v)$ can be obtained by solving the following linear programming problem:

$$\begin{aligned} &\text{minimize} && \sum_{i=1}^{l} \lambda_i \\ &\text{subject to} && -\alpha v + \sum_{i=1}^{l} \lambda_i y_i = u \\ &\text{and} && \lambda_i \ge 0 \end{aligned} \tag{23}$$

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 whether a simpler solution exists in this case. We have mentioned in Section 3 that the problem of finding quantity (13), which is the problem we are trying to solve, can be directly reduced to a linear programming problem. In the case when $S = \bigcap_{i=1}^{l} \{x \in \mathbb{R}^m : \langle x, y_i \rangle \le 1\}$, the linear programming problem becomes

$$\begin{array}{lll} \text{minimize} & \lambda & \\ \text{subject to} & k'y_i \le \lambda, \quad i = 1, 2, \ldots l & (24) \\ \text{and} & k'u = 1,\ k'v = 0, & \end{array}$$

(cf. Tesi and Vicino, 1990). To transform the linear programming problem (23) to the standard form which is assumed by the simplex method, we only need to replace α by $\alpha_1 - \alpha_2$, where α_1 and α_2 are positive; this results in a standard linear programming problem with $l + 2$ variables and m constraints. The corresponding standard form of the linear programming problem (24) has $2m + l + 2$ variables and $l + 2$ constraints. It is to be noted that to make S an absorbing polytope, l has to be greater than m. The advantage of (23) over (24) now becomes apparent.

7. SEARCH PROBLEM

Recall that our major problem of this paper is to compute the robustness measure ρ defined as a function of a four-tuple $(F, g, \mathbb{C}_g, S)$ in (4)–(7). It is shown in Section 3 that $\rho = \inf_{s \in \partial\mathbb{C}_g} \tau(s)$, where τ is a function from $\partial\mathbb{C}_g$ to $[0, \infty]$ which can be evaluated at any fixed $s \in \partial\mathbb{C}_g$ by using the various methods developed in Section 4 through Section 6, if S is an bounded absorbing convex set. The quantity ρ can then be obtained by a "brute force" search over $\partial\mathbb{C}_g$, i.e. choose a finite subset of $\partial\mathbb{C}_g$ which is sufficiently "dense" in $\partial\mathbb{C}_g$ and find the minimum of $\tau(s)$ over all points in this finite subset. Since $\partial\mathbb{C}_g$ is usually a one-dimensional curve in $\bar{\mathbb{C}}$, this "brute force" search is numerically feasible to do in engineering applications. However as we mentioned earlier, two numerical difficulties may occur during the search. The first difficulty is that when $\partial\mathbb{C}_g$ is unbounded (or virtually unbounded, i.e. part of $\partial\mathbb{C}_g$ is large), it may be impossible (or unrealistic) to find a finite and sufficiently "dense" subset in $\partial\mathbb{C}_g$. The second difficulty occurs due to the fact that $\tau(s)$ is generally not a continuous function on $\partial\mathbb{C}_g$, so it may be easy to miss the true infimum when a "brute force" search over $\partial\mathbb{C}_g$ is carried out. This section discusses methods to deal with these difficulties.

The first difficulty can be overcome completely by using a "tricky" but simple method which generalizes a similar method in Chapellat *et al.* (1988) to our setup. If $(F, g, \mathbb{C}_g, S)$ is given, then $\partial\mathbb{C}_g$, the boundary of the stability region, and $p(s, k)$, the affine map from $\mathbb{R}^m$ to $\mathbb{P}$, are also specified, and the function τ is defined in (8). Let

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

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

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

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

$$\partial\hat{\mathbb{C}}_g = \left\{ s \in \bar{\mathbb{C}} : \frac{1}{s} \in \partial\mathbb{C}_g \right\},$$

and a new function $\hat{\tau} : \partial\hat{\mathbb{C}}_g \to [0, \infty]$ by

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

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

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

Therefore

$$\begin{aligned} \rho &= \inf_{s \in \partial\mathbb{C}_g} \tau(s) \\ &= \min\left\{ \inf_{s \in \partial\mathbb{C}_g \cap \mathbb{U}} \tau(s),\ \inf_{s \in \partial\mathbb{C}_g \cap \mathbb{U}^c} \tau(s) \right\} \\ &= \min\left\{ \inf_{s \in \partial\mathbb{C}_g \cap \mathbb{U}} \tau(s),\ \inf_{s \in \partial\hat{\mathbb{C}}_g \cap \mathbb{U}} \tau\left(\frac{1}{s}\right) \right\} \\ &= \min\left\{ \inf_{s \in \partial\mathbb{C}_g \cap \mathbb{U}} \tau(s),\ \inf_{s \in \partial\hat{\mathbb{C}}_g \cap \mathbb{U}} \hat{\tau}(s) \right\}, \qquad (25) \end{aligned}$$

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

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 $s \in \partial\hat{\mathbb{C}}_g$ by using the same method as carried out for τ at each point $s \in \partial\mathbb{C}_g$. This shows that ρ can be obtained by two independent searches and that each of these searches is carried out over a "small" bounded set.

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

Recall the expression of $\tau(s)$ given by (12) and the definition of vectors $u(s)$, $v(s)$. Define

$$\partial\mathbb{C}_{g1} = \{s \in \partial\mathbb{C}_g : u(s) \ne 0 \text{ and } v(s) = 0\},$$

$$\partial\mathbb{C}_{g2} = \{s \in \partial\mathbb{C}_g : \operatorname{rank}[u(s)\ v(s)] = 2\}.$$

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

$$\inf_{s\in\partial\mathbb{C}_g} \tau(s) = \min\left\{\inf_{s\in\partial\mathbb{C}_{g1}} \tau(s), \inf_{s\in\partial\mathbb{C}_g\setminus\partial\mathbb{C}_{g1}} \tau(s)\right\} \quad (26)$$

$$= \min\left\{\inf_{s\in\partial\mathbb{C}_g\setminus\partial\mathbb{C}_{g2}} \tau(s), \inf_{s\in\partial\mathbb{C}_{g2}} \tau(s)\right\}. \quad (27)$$

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

Lemma 1. $\tau|_{\partial\mathbb{C}_{gi}}$ is a continuous function for $i=1,2$.

Lemma 1 implies that if we can identify either $\partial\mathbb{C}_{g1}$ or $\partial\mathbb{C}_{g2}$ from $\partial\mathbb{C}_g$ and carry out the "brute force" searches according to either (26) or (27), respectively, then we will have very little chance to miss the true infimum of $\tau(s)$ over $\partial\mathbb{C}_g$.

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

$$\left\{s \in \bar{\mathbb{C}} : \Im\left(\frac{F'[s^n\, s^{n-1} \cdots 1]'}{[s^n\, s^{n-1} \cdots 1]g}\right) = 0\right\}.$$

The latter set contains $\mathbb{R} \cup \{\infty\}$ and usually consists of some curves and points in $\bar{\mathbb{C}}$. Hence the set $\partial\mathbb{C}_{g1}$ usually contains discrete points in $\partial\mathbb{C}_g$. 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 $\partial\mathbb{C}_g$.) In most cases when m, the dimension of k, is greater than 1, these points are just the elements of the intersection of $\partial\mathbb{C}_g$ and $\mathbb{R} \cup \{\infty\}$.

Proof. (Lemma 1.) First note that $u(s)$ and $v(s)$ are continuous functions of $s \in \partial\mathbb{C}_g$.

It follows from Theorem 2 that

$$\tau|_{\partial\mathbb{C}_{g1}}(s) = \max\left\{\frac{1}{\mu_{S^0}[u(s)]}, \frac{1}{\mu_{S^0}[-u(s)]}\right\}.$$

Since $\mu_{S^0}(u)$ and $\mu_{S^0}(-u)$ are continuous functions of u if S^0 is absorbing, or equivalently if S is bounded, the continuity of $\tau|_{\partial\mathbb{C}_{g1}}$ follows.

Define a function $\theta : \partial\mathbb{C}_{g2} \times \mathbb{R} \to [0, \infty)$ by $\theta(s, \alpha) = \mu_{S^0}[u(s) + \alpha v(s)]$ and another function

$$\xi : \partial\mathbb{C}_{g2} \to [0, \infty) \text{ by } \xi(s) = \inf_{\alpha\in\mathbb{R}} \theta(s, \alpha).$$

Proposition 1(g) implies that the function θ is continuous in both variables and it is shown in Section 4 that the infimum in the definition of ξ occurs at a finite point for each $s \in \partial\mathbb{C}_{g2}$. Therefore, we actually have $\xi(s) = \min_{\alpha\in\mathbb{R}} \theta(s, \alpha)$.

To show that $\tau|_{\alpha\mathbb{C}_{g2}}$ is continuous, it is enough to show that ξ is continouus, or equivalently, to show that $\lim_{i\to\infty} \xi(s_i) = \xi(\bar{s})$ for any $\bar{s} \in \partial\mathbb{C}_{g2}$ and any sequence $\{s_i\} \subset \partial\mathbb{C}_{g2}$ such that $s_i \to \bar{s}$.

Let $\bar{s} \in \partial\mathbb{C}_{g2}$ and let $\{s_i\}$ be a sequence in $\partial\mathbb{C}_{g2}$ such that $s_i \to \bar{s}$. Since $v(\bar{s}) \neq 0$, there exists $\varepsilon > 0$ such that $\mu_{S^0}[v(s_i)] \geq \varepsilon$ for all large enough i. Since $\{u(s_i)\}$ is a convergent sequence in $\mathbb{R}^m$, there exists $\gamma > 0$ such that $\mu_{S^0}[u(s_i)] \leq \gamma$ for all i. The sequence $\{\xi(s_i)\}$ is bounded since $\xi(s_i) \leq \mu_{S^0}[u(s_i)] \leq \gamma$ for each i. Now suppose that we do not have $\lim_{i\to\infty} \xi(s_i) = \xi(\bar{s})$. This means that $\{\xi(s_i)\}$ has a subsequence $\{\xi(s_{i_j})\}$ such that $\xi(s_{i_j}) \to \sigma \neq \xi(\bar{s})$. Let $\bar{s}_j = s_{i_j}$ and α_j be a real number satisfying $\xi(\bar{s}_j) = \theta(\bar{s}_j, \alpha_j)$. The sequence $\{\alpha_j\}$ must be contained in a compact subset of $\mathbb{R}$, since each α_j must satisfy $|\alpha_j| \leq 2\frac{\gamma}{\varepsilon}$; this implies that $\{\alpha_j\}$ has a convergent subsequence $\{\alpha_{j_\nu}\}$. Let $\alpha_{j_\nu} \to \bar{\alpha}$. Then

$$\begin{aligned}\theta(\bar{s}, \bar{\alpha}) &= \lim_{\nu\to\infty} \theta(\bar{s}_{j_\nu}, \alpha_{j_\nu})\\ &= \lim_{j\to\infty} \xi(\bar{s}_j)\\ &= \sigma\\ &\neq \xi(\bar{s}).\end{aligned}$$

Let $\bar{\beta}$ be chosen so that $\xi(\bar{s}) = \min_{\alpha\in\mathbb{R}} \theta(\bar{s}, \alpha) = \theta(\bar{s}, \bar{\beta})$. If $\bar{\beta} = \bar{\alpha}$, then $\xi(\bar{s}) = \theta(\bar{s}, \bar{\beta}) = \theta(\bar{s}, \bar{\alpha}) = \sigma$, which contradicts our hypothesis. Now assume $\bar{\beta} \neq \bar{\alpha}$. In this case $\xi(\bar{s}) = \theta(\bar{s}, \bar{\beta}) \leq \theta(\bar{s}, \bar{\sigma}) = \sigma$. Hence we must have

$$\theta(\bar{s}, \bar{\beta}) < \theta(\bar{s}, \bar{\alpha}).$$

This also means that there exists $\delta > 0$ such that

$$\theta(s, \beta) < \theta(s, \alpha),$$

for all $s \in \partial\mathbb{C}_{g1}$ with $|s - \bar{s}| < \delta$, all α with $|\alpha - \bar{\alpha}| < \delta$ and all β with $|\beta - \bar{\beta}| < \delta$. Choose $\bar{s}_{j_\nu}$ such that $|\bar{s}_{j_\nu} - \bar{s}| < \delta$ and $|\alpha_{j_\nu} - \bar{\alpha}| < \delta$. Then

$$\theta(\bar{s}_{j_\nu}, \bar{\beta}) < \theta(\bar{s}_{j_\nu}, \alpha_{j_\nu}).$$

This contradicts the fact that $\min_{\alpha\in\mathbb{R}} \theta(\bar{s}_{j_\nu}, \alpha) = \theta(\bar{s}_{j_\nu}, \alpha_{j_\nu})$. Therefore, we conclude that there is no subsequence of $\{\xi(s_i)\}$ which does not converge to $\xi(\bar{s})$. This means that $\{\xi(s_i)\}$ must converge to $\xi(\bar{s})$.

8. A FLEXIBLE BEAM EXAMPLE

A possible model with parameter uncertainty for a flexible beam studied in MacLean (1990) is given by the following transfer function

$$P(s)=\frac{-8.1876(1+k_5)s^2+0.7895(1+k_6)s+219.57(1+k_7)}{5(1+k_1)s^4+0.8707(1+k_2)s^3+139.61(1+k_3)s^2+0.0933(1+k_4)s}$$

where $k=[k_1\,k_2\cdots k_7]$ represents the uncertainties. A feedback controller, which has a transfer function

$$C(s)=-\frac{0.8745s^2+4.0787s^2+2.4574s+0.6105}{s^3+3.7897s^2+5.9143s},$$

is designed for P to solve the robust servomechanism problem with respect to constant signals, i.e. to make the closed loop system track constant reference signals and reject constant disturbance signals.

After a straightforward but tedious manipulation, we obtain the characteristic polynomial of the closed loop system as

$$p(s,k)=[s^7\,s^6\,s^5\,s^4\,s^3\,s^2\,s\;1](Fk+g),$$

where

$$F=\begin{bmatrix} 5 & 0 & 0 & 0 & 0 & 0 & 0\\ 18.948 & 0.8707 & 0 & 0 & 0 & 0 & 0\\ 29.571 & 3.2996 & 139.61 & 0 & -7.1597 & 0 & 0\\ 0 & 5.1494 & 529.07 & 0.0933 & -33.395 & 0.6904 & 0\\ 0 & 0 & 825.68 & 0.3535 & -20.120 & 3.2202 & 192.01\\ 0 & 0 & 0 & 0.5517 & -4.9988 & 1.9402 & 895.56\\ 0 & 0 & 0 & 0 & 0 & 0.4820 & 539.58\\ 0 & 0 & 0 & 0 & 0 & 0 & 134.06 \end{bmatrix},$$

$$g=\begin{bmatrix} 5\\ 19.819\\ 165.32\\ 501.61\\ 1001.1\\ 893.06\\ 540.07\\ 134.06 \end{bmatrix}.$$

The polynomial is nominally stable and from (25) the stability robustness measure is given by

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

where

$$\tau(j\omega)=\inf\{\mu_S(k):k\in\mathbb{R}^m \text{ and } p(j\omega,k)=0\},$$
$$\hat{\tau}(j\omega)=\inf\{\mu_S(k):k\in\mathbb{R}^m \text{ and } \hat{p}(j\omega,k)=0\},$$

and

$$\hat{p}(s,k)=[s^7\,s^6\,s^5\,s^4\,s^3\,s^2\,s\;1](\hat{F}k+\hat{g}),$$

where

$$\hat{F}=\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 134.06\\ 0 & 0 & 0 & 0 & 0 & 0.4820 & 539.58\\ 0 & 0 & 0 & 0.5517 & -4.9988 & 1.9402 & 895.56\\ 0 & 0 & 825.68 & 0.3535 & -20.120 & 3.2202 & 192.01\\ 0 & 5.1494 & 529.07 & 0.0933 & -33.395 & 0.6904 & 0\\ 29.571 & 3.2996 & 139.61 & 0 & -7.1597 & 0 & 0\\ 18.948 & 0.8707 & 0 & 0 & 0 & 0 & 0\\ 5 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\hat{g}=\begin{bmatrix} 134.06\\ 540.07\\ 893.06\\ 1001.1\\ 501.61\\ 165.32\\ 19.819\\ 5 \end{bmatrix}.$$

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 (15) and (16). We obtain

$$\inf_{\omega\in[0,1]}\tau(j\omega)=\tau(j\omega)|_{\omega=0.4856}=0.5924,$$

$$\inf_{\omega\in[0,1)}\hat{\tau}(j\omega)=\hat{\tau}(j\omega)|_{\omega=0.2876}=0.3268.$$

Therefore $\rho=0.3268$.

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 (15) and (18). We obtain

$$\inf_{\omega\in[0,1]}\tau(j\omega)=\tau(j\omega)|_{\omega=0.4871}=0.4534,$$

$$\inf_{\omega\in[0,1)}\hat{\tau}(j\omega)=\hat{\tau}(j\omega)|_{\omega=0.2748}=0.1977.$$

Therefore $\rho=0.1977$.

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 (15) and (20). We obtain

$$\inf_{\omega\in[0,1]} \tau(j\omega)=\tau(j\omega)|_{\omega=0.4854}=0.6340,$$

$$\inf_{\omega\in[0,1)} \hat{\tau}(j\omega)=\hat{\tau}(j\omega)|_{\omega=0.2906}=0.3638.$$

Therefore $\rho=0.3638$.

Case IV. *S is a simplex in* $\mathbb{R}^7$

A polytope with nonempty interior in $\mathbb{R}^m$ is called a *simplex* if it has exactly $m+1$ vertices. For a simplex, the convex hull of every pair of vertices is an edge. If S is an absorbing simplex in $\mathbb{R}^7$ given by

$$S=\mathrm{co}\left\{\begin{bmatrix}7\\-1\\-1\\-1\\-1\\-1\\-1\end{bmatrix},\begin{bmatrix}-1\\7\\-1\\-1\\-1\\-1\\-1\end{bmatrix},\begin{bmatrix}-1\\-1\\7\\-1\\-1\\-1\\-1\end{bmatrix},\begin{bmatrix}-1\\-1\\-1\\7\\-1\\-1\\-1\end{bmatrix},\begin{bmatrix}-1\\-1\\-1\\-1\\7\\-1\\-1\end{bmatrix},\begin{bmatrix}-1\\-1\\-1\\-1\\-1\\7\\-1\end{bmatrix},\begin{bmatrix}-1\\-1\\-1\\-1\\-1\\-1\\7\end{bmatrix},\begin{bmatrix}-1\\-1\\-1\\-1\\-1\\-1\\-1\end{bmatrix}\right\}, \quad (28)$$

then $\tau(j\omega)$ and $\hat{\tau}(j\omega)$ can be evaluated by using (14) and (21). We obtain

$$\inf_{\omega\in[0,1]} \tau(j\omega)=\tau(j\omega)|_{\omega=0.4863}=0.1730,$$

$$\inf_{\omega\in[0,1)} \hat{\tau}(j\omega)=\hat{\tau}(j\omega)|_{\omega=0.2906}=0.0968.$$

Therefore $\rho=0.0968$.

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 a rich mathematical insight to this area. 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.

Acknowledgement—The authors are grateful to Dr D. Miller for helpful discussions.

REFERENCES

Barmish, B. R. (1989). A generalization of Kharitonov's four-polynomial concept for the robust stability problems with linearly dependent coefficient perturbations. *IEEE Trans. Aut. Contr.*, **AC-34**, 157–165.

Bartlett, A. C., C. V. Hollot and Lin, H. (1988). Root locations for an entire polytope of polynomials: it suffices to check the edges. *Math. Control Signals and Systems*, **1**, 61–71.

Bialas, S. and J. Garloff (1985). Stability of polynomials under coefficient perturbation. *IEEE Trans. Aut. Contr.*, **AC-30**, 310–312.

Biernacki, R. M., H. Hwang and S. P. Bhattacharyya (1987). Robust stability with structured real parameter perturbations. *IEEE Trans. Aut. Contr.*, **AC-32**, 495–505.

Brøndsted, A. (1983). *An Introduction to Convex Polytopes.* Springer-Verlag, New York.

Chapellat, H. and S. P. Bhattacharyya (1989). A generalization of Kharitinov's theorem: robust stability of interval plants. *IEEE Trans. Aut. Contr.*, **AC-34**, 306–311.

Chapellat, H., S. P. Bhattacharyya and L. H. Keel (1988). Stability margins for Hurwitz polynomials. *Proc. IEEE Conf. Decision and Control*, pp. 1392–1398.

Fu, M. (1989). Polytopes of polynomials with zeros in a prescribed region: new criteria and algorithms. In M. Milanese, R. Tempo and A. Vicino (Eds), *Robustness in Identification and Control*, pp. 125–145. Plenum Press, New York.

Fu, M. and B. R. Barmish (1988). Maximal unidirectional perturbation bounds for stability of polynomials and matrices. *Systems & Control Letters*, **11**, 173–179.

Hinrichsen, D. and A. J. Pritchard (1988). New robustness results for linear systems under real perturbations. *Proc. IEEE Conf. Decision and Control*, p. 1375–1379.

Hinrichsen, D. and A. J. Pritchard (1989). An application of state space methods to obtain explicit formulae for robustness measures of polynomials. M. Milanese, R. Tempo and A. Vicino (Eds), *Robustness in Identification and Control*, pp. 183–206. Plenum Press, New York.

Kharitonov, V. L. (1978). Asymptotic stability of an equilibrium position of a family of systems of linear differential equations. *Differenncjalnyje Uravnenija*, **14**, 2086–2088.

MacLean, J. (1990). Modelling and control of a flexible beam. Master's thesis, Dept. of Electrical Eng., University of Toronto, Toronto, Canada.

Qiu, L. and E. J. Davison (1989a). A simple procedure for the exact stability robustness computation of polynomials with affine coefficient perturbations. *Systems & Control Letters*, **13**, 413–420.

Qiu, L. and E. J. Davison (1989b). A unified approach for the stability robustness of polynomials in a convex set. Technical Report No. 8912, System Control Group, University of Toronto.

Rockafellar, R. T. (1970). *Convex Analysis.* Princeton University Press, Princeton, NJ.

Royden, H. L. (1968). *Real Analysis.* MacMillan, New York.

Rudin, W. (1973). *Functional Analysis.* McGraw-Hill, New York.

Sarídereli, M. K. and F. J. Kern (1987). The stability of polynomials under correlated coefficient perturbations. *Proc. IEEE Conf. Decision and Control*, pp. 1618–1621.

Stewart, G. W. (1975). Gershgorin theory for the generalized eigenvalue problem $Ax=\lambda Bx$. *Math. of Comp.*, **29**, 600–606.

Szász, G. (1963). *Introduction to Lattice Theory.* Academic Press, New York.

Tesi, A. and A. Vicino (1990). Robustness analysis for linear dynamical systems with linearly correlated parametric uncertainties. *IEEE Trans. Aut. Contr.*, **AC-35**, 186–191.