# Condition and Complexity Measures for Infeasibility Certificates of Systems of Linear Inequalities and Their Sensitivity Analysis*

Hugo J. Lara †     Levent Tunçel‡

July 2002

**Abstract**

We begin with a study of the infeasibility measures for linear programming problems. For this purpose, we consider feasibility problems in Karmarkar's standard form. Our main focus is on the complexity measures which can be used to bound the amount of computational effort required to solve systems of linear inequalities and related problems in certain ways. We propose a new complexity measure that is particularly well-suited for the generalized Tardos' scheme for the real number data model. We prove that the new measure is between Ye's (smallest large variable) measure and $\bar{\chi}$. We present geometric interpretations of the complexity measures and then turn to the sensitivity analyses and the computation of the directional derivatives of the complexity measures. For this purpose, various sets of allowed perturbations are identified (depending on the complexity measure) using the minimal and maximal sign vectors of the subspaces involved. Finally, we consider the generalization of the infeasibility certificates to convex optimization problems in conic form. We present a geometric generalization of a condition measure proposed by Cheung-Cucker. We derive various new relationships amongst the existing and new complexity measures in this context.

**Keywords:** linear inequality systems, linear programming, convex optimization, computational complexity, condition numbers, complexity measures, interior-point methods

**AMS Subject Classification:** 90C05, 90C25, 90C60, 52A41, 49K40, 90C31, 90C51

---

*Results of this manuscript were presented at the Fields Institute and at the SIAM Conference on Optimization in Toronto during May 2002.

†(hugol@ucla.edu.ve) Decanato de Ciencias. Universidad CentroOccidental Lisandro Alvarado (UCLA). Apdo 400. Barquisimeto, 3001, Venezuela. Research suported in part by CDCHT-UCLA, Venezuela. Also supported by a research grant from NSERC and a PREA of the second author.

‡(ltuncel@math.uwaterloo.ca) Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada. Research of this author was supported in part by research grants from NSERC and a PREA from Ontario, Canada.

# 1 Introduction, Definitions and Notation

Let $A \in \mathbb{R}^{m\times n}$ such that $\mathrm{rank}(A) = m$ (this will be assumed throughout the paper) be given.

Consider the feasibility problems:

$$\text{Does there exist } x \in \mathbb{R}^n \quad \text{such that } Ax = b,\ x \geq 0? \tag{1}$$

$$\text{Does there exist } y \in \mathbb{R}^m \quad \text{such that } A^T y \leq c? \tag{2}$$

In the above, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$ are given. The feasibility problems (1) and (2) correspond to the primal-dual linear optimization problems:

$$\begin{array}{llrcl} (P) & \text{minimize} & c^T x & & \\ & & Ax & = & b, \\ & & x & \geq & 0, \end{array}$$

$$\begin{array}{llrcl} (D) & \text{maximize} & b^T y & & \\ & & A^T y & \leq & c. \end{array}$$

By Farkas' lemma, (1) has no solution iff there exist $\bar{y} \in \mathbb{R}^m$, $\bar{s} \in \mathbb{R}^n$ such that

$$A^T\bar{y} + \bar{s} = 0, e^T\bar{s} = 1, \bar{s} \geq 0, \text{ and } b^T\bar{y} > 0.$$

Similarly, (2) has no solution iff there exists $\bar{x} \in \mathbb{R}^n$ such that

$$A\bar{x} = 0, e^T\bar{x} = 1, \bar{x} \geq 0, \text{ and } c^T\bar{x} < 0.$$

Since $\mathrm{rank}(A) = m$, the system $Ax = b$ always has solution(s) and we can find $l \in \mathbb{R}^n$ such that $Al = b$ (e.g., $l := A^T(AA^T)^{-1}b$).

Let $\mathcal{N}(\cdot)$, $\mathcal{R}(\cdot)$ denote the nullspace and the range (respectively) of the matrix argument. We can write: (1) is infeasible iff

$$\exists \bar{s} \in \mathcal{R}(A^T) \text{ such that } e^T\bar{s} = 1, \bar{s} \geq 0, \text{ and } l^T\bar{s} < 0.$$

The above provides some motivation for studying complexity and condition measures for the feasibility problems in Karmarkar's standard form:

$$\left\{x :\ Ax = 0,\ e^T x = 1,\ x \geq 0\right\}.$$

For such a feasibility problem, we have a dual which exposes a beautiful structure (see Vavasis and Ye [40]):

$$(FP) \left\{ \begin{array}{rcl} x & \in & \mathcal{N}(A), \\ \|x\|_1 & = & 1, \\ x & \geq & 0, \end{array} \right.$$

$$(FD) \left\{ \begin{array}{rcl} s & \in & \mathcal{R}(A^T), \\ \|s\|_1 & = & 1, \\ s & \geq & 0. \end{array} \right.$$

The strict complementarity theorem of linear optimization translates to the following fact: For every $A \in \mathbb{R}^{m \times n}$, there exists $[B, N]$, a partition of $\{1, 2, \ldots, n\}$ ($B$ or $N$ may be empty) such that there exists $\bar{x}$ feasible in $(FP)$, $\bar{x}_B > 0$, $\bar{x}_N = 0$ and there exists $\bar{s}$ feasible in $(FD)$, $\bar{s}_B = 0$, $\bar{s}_N > 0$. In particular, $B = \emptyset$ iff $(FP)$ is infeasible and $B = \{1, 2, \ldots, n\}$ iff $(FD)$ is infeasible. Note that the partition $[B, N]$ as described above is always unique and is called *the strict complementarity partition determined by* $A$.

Since $(FP)$ and $(FD)$ are defined over any pair of orthogonally complementary linear subspaces in $\mathbb{R}^n$, we have the following definitions (for convenience, $S$ is the nullspace of $A$):

$$\sigma_P(A) := \sigma_P(S) := \min_{j \in B} \max \left\{ x_j : x \in S, \|x\|_1 = 1, x \geq 0 \right\},$$

$$\sigma_D(A) := \sigma_D(S) := \sigma_P(S^\perp) = \min_{j \in N} \max \left\{ s_j : s \in S^\perp, \|s\|_1 = 1, s \geq 0 \right\},$$

where $\sigma_P(S) := 1$ if $B = \emptyset$ (and therefore, $\sigma_D(S) := 1$ if $N = \emptyset$). The primal-dual complexity measure of Ye [42, 41] for the pair $(FP), (FD)$ is then defined by

$$\sigma(A) := \min \left\{ \sigma_P(A), \sigma_D(A) \right\}$$

(we define $\sigma(S)$ similarly).

Ye [42, 41] and Vavasis-Ye [40] show that the abovementioned strict complementarity partition $[B, N]$ can be computed in $O\left(\sqrt{n} \ln\left(\frac{n}{\sigma(A)}\right)\right)$ interior-point iterations.

Many of the concepts in our presentation become more apparent when we focus on those characterizations of the complexity measures involving sign patterns of vectors in certain linear subspaces. For $x \in \mathbb{R}^n$, $\operatorname{sign}(x) \in \{-, 0, +\}^n$ encodes the signs of the entries of $x$. Let $S \subseteq \mathbb{R}^n$ be a linear subspace. We denote by $\operatorname{sign}(S) \subseteq \{-, 0, +\}^n$ the set of sign vectors of the elements of $S$.

Note that if $A \in \mathbb{R}^{m \times n}$ such that $\mathcal{N}(A) = S$ then every nonzero vector in $S$ represents a linear dependence amongst the columns of $A$. Minimal linear dependencies play a particularly important role in what follows.

We denote the set of sign patterns of those minimal elements in $S$ by $\underline{\operatorname{sign}}(S)$. That is, $\underline{\operatorname{sign}}(S) \subset \operatorname{sign}(S)$ denotes those nonzero sign patterns in $\operatorname{sign}(S)$ such that setting any number of $+$'s and $-$'s to zero (without changing the others) does not give another nonzero element of $\operatorname{sign}(S)$. For $x \in \mathbb{R}^n$, let

$$J_-(x) := \left\{ j \in \{1, 2, \ldots, n\} : x_j < 0 \right\},$$

$$J_0(x) := \left\{ j \in \{1, 2, \ldots, n\} : x_j = 0 \right\},$$

$$J_+(x) := \{j \in \{1, 2, \ldots, n\} : x_j > 0\}.$$

Then, $\bar{x} \in S \backslash \{0\}$ is *minimal* if for all $\hat{x} \in S \backslash \{0\}$ satisfying $J_-(\hat{x}) \subseteq J_-(\bar{x})$, $J_+(\hat{x}) \subseteq J_+(\bar{x})$, $J_0(\hat{x}) \supseteq J_0(\bar{x})$ we have $\text{sign}(\hat{x}) = \text{sign}(\bar{x})$. So, $\bar{x} \in S$ is minimal iff $\text{sign}(\bar{x}) \in \underline{\text{sign}}(S)$.

It is well-known that if we identify the elements of $\underline{\text{sign}}(S)$ as the *circuits* on the ground set $\{-, 0, +\}^n$, we obtain an *oriented matroid* of rank $[n - \dim(S)]$. Nonnegative sign patterns are particularly important to us:

$$\text{sign}_+(S) := \{\text{sign}(x) : x \in S, x \geq 0\},$$

$$\underline{\text{sign}}_+(S) := \underline{\text{sign}}(S) \cap \text{sign}_+(S).$$

Also relevant to our study are the *maximal* elements of a subspace. Analogously, we say $\bar{x} \in S$ is *maximal* if for all $\hat{x} \in S$ satisfying $J_-(\hat{x}) \supseteq J_-(\bar{x})$, $J_+(\hat{x}) \supseteq J_+(\bar{x})$, $J_0(\hat{x}) \subseteq J_0(\bar{x})$ we have $\text{sign}(\hat{x}) = \text{sign}(\bar{x})$. We denote by $\overline{\text{sign}}(S)$ the set of sign vectors of all maximal elements of $S$. We also define

$$\overline{\text{sign}}_+(S) := \overline{\text{sign}}(S) \cap \text{sign}_+(S).$$

Note that for a given subspace $S$, $\overline{\text{sign}}_+(S)$ is either empty or is a singleton. Moreover, if $S = \mathcal{N}(A)$ then $\overline{\text{sign}}_+(S)$ identifies the elements of $B$ and $\overline{\text{sign}}_+(S^\perp)$ identifies the elements of $N$.

The following very elementary lemma expresses the minimal elements in $\mathcal{N}(A)$ as minimal linear dependencies amongst the columns of $A$.

**Lemma 1.1** *Consider $\bar{x} \in \mathcal{N}(A)$ such that $\|\bar{x}\|_1 = 1$. Then, $\bar{x}$ is minimal in $\mathcal{N}(A)$ if and only if for $J := J_+(\bar{x}) \cup J_-(\bar{x})$ the system of equations*

$$A_J x_J = 0, \|x_J\|_1 = 1 \tag{3}$$

*has $\bar{x}_J$ as the unique solution.*

**Proof.** Take an arbitrary solution $\hat{x}_J$ of (3). Suppose for a contradiction that $\hat{x}_J \neq \bar{x}_J$. Define $h_J(\alpha) = \bar{x}_J + \alpha \hat{x}_J$. Now choose $\bar{\alpha}$ such that $|\bar{\alpha}|$ is the smallest positive number such that some of the components of $h_J(\alpha)$ become zero while the rest of the components preserve their signs in $\bar{x}_J$ (not all the components can be driven to zero at the same time, because $\hat{x}_J \neq \bar{x}_J$). Define $h(\bar{\alpha})$ in $\mathbb{R}^n$ by completing with zeros. Clearly $0 \neq h(\bar{\alpha}) \in \mathcal{N}(A)$, $J_+(h(\bar{\alpha})) \subseteq J_+(\bar{x})$, $J_-(h(\bar{\alpha})) \subseteq J_-(\bar{x})$ and $J_0(h(\bar{\alpha})) \supset J_0(\bar{x})$. This contradicts the minimality of $\bar{x}$ since $\text{sign}(\bar{x}) \neq \text{sign}(h(\bar{\alpha}))$. Conversely, suppose that $\bar{x}_J$ is the unique solution of (3) and that it is not minimal. Then there exists $\hat{x} \in \mathcal{N}(A)$ different from $\bar{x}$ satisfying $\|\hat{x}\|_1 = 1$, $J_+(\hat{x}) \subseteq J_+(\bar{x})$, $J_-(\hat{x}) \subseteq J_-(\bar{x})$ and $J_0(\hat{x}) \supseteq J_0(\bar{x})$. Then we have that $(J_+(\hat{x}) \cup J_-(\hat{x})) \subseteq J$. Therefore, $\hat{x}_J \in \mathcal{N}(A_J)$ and $\|\hat{x}_J\|_1 = 1$. This provides a contradiction because $\bar{x}_J$ is the unique point satisfying these last relations.

□

Next we define a new complexity measure

$$\xi_P(A) := \xi_P(S) := \min\left\{x_j \neq 0 : x \in S, \|x\|_1 = 1, \mathrm{sign}(x) \in \underline{\mathrm{sign}}_+(S)\right\},$$

$$\xi_D(A) := \xi_D(S) := \xi_P(S^\perp) = \min\left\{s_j \neq 0 : s \in S^\perp, \|s\|_1 = 1, \mathrm{sign}(s) \in \underline{\mathrm{sign}}_+(S^\perp)\right\},$$

$$\xi(A) := \min\left\{\xi_P(A), \xi_D(A)\right\},$$

where $\xi_P(S)$ is defined to be 1 if the set over which the minimum is computed is empty. We define $\xi(S)$ accordingly. These complexity measures provide our starting point in this paper. We study the properties of infeasibility certificates for systems of linear inequalities and their generalizations in the more general context of convex optimization.

In the next section, we give a dual characterization of $\xi(A)$ using LP duality theory. Then in Section 3, we review characterizations of $\bar{\chi}(A)$ and extend and/or sharpen some of the known results. In Section 4, we show that the new complexity measure is particularly well-suited to Tardos' scheme in solving LP feasibility problems. In this section we improve a result of Ho and Tunçel [20] and refine a theorem of Todd, Tunçel and Ye [36]. Section 5 concerns the basic setup for a variety of conditions on the sign vectors of subspaces and the study of how $\sigma(A)$ and $\xi(A)$ behave under perturbations. Many of the complexity measures can be expressed as the maximum (or the minimum) of the optimal values of a family of LP problems. For perturbation theory of linear programming problems see, for instance, first Robinson [32, 31, 30, 29, 28], then Hirabayashi-Jongen-Shida [19] and Renegar [27]. The complexity measures, we are concerned with here, are motivated by the complexity analyses of interior-point methods. There are many papers which discuss perturbation and sensitivity analysis from an interior-point method point of view; see Adler-Monteiro [1], Greenberg [15, 16], Mehrotra-Monteiro [24] and Yıldırım-Todd [43]. None of these works is concerned with analyzing the sensitivity of $\sigma(A)$, $\xi(A)$ or $\bar{\chi}(A)$. Luo and Tseng [22] obtained perturbation results for Hoffman constants for systems of linear inequalities (see also Deng [7] and Azé and Corvellec [2] for more recent studies of related issues in more general settings). These approaches are based on rank conditions on certain submatrices of $A$. Our approach is more geometric and is based on sign vectors defined by the subspace partition determined by $A$, as suggested in [37]. Our approach easily applies to $\sigma(A)$, $\xi(A)$ and $\bar{\chi}(A)$ in a unified manner. We do not analyze the Hoffman constants in this paper; however, our approach can also be used to obtain results similar to those of Luo and Tseng [22].

In Section 6, we begin extending our study to a convex optimization setting. Again we take a geometric viewpoint. We consider the complexity measure sym$(\cdot)$ studied by Epelman and Freund [8] (also see the measure $g(\cdot)$ proposed by Freund [11] and a related geometric approach to primal-dual level sets in convex optimization [12]). We also consider the condition measure of Renegar (see [27, 26]). In the context of *strong infeasibility certificates*, we generalize a condition measure of Cheung-Cucker [5] (originally proposed for systems of linear inequalities) to convex optimization. We present various results establishing some new relationships amongst these condition and complexity measures in addition to results involving the *width* and the *norm approximation coefficients* of convex cones.

# 2 Dual Characterization of $\xi$

We use linear programming duality theory repeatedly throughout the paper. To warm the reader up to the notation, we present the following dual characterization of $\xi$.

**Theorem 2.1** *For all linear subspaces $S \subseteq \mathbb{R}^n$, we have the following dual characterization of $\xi$:*

$$\xi_P(S) = \min_{\substack{\emptyset \neq J \subset B \\ i \in (B \backslash J)}} \quad \max_{\substack{s \in S^\perp \\ \mathrm{sign}(s_j) = +, \forall j \in J \\ \mathrm{sign}(s_i) = -}} \quad \min\left\{\min_{j \in J}\{s_j\}, (1+s_i)\right\}.$$

When $B = \emptyset$, the last minimum above is over the empty set and for convenience, it is defined to be 1. Similarly, when $|B| = 1$, we have $\xi_P(S) = 1$ and the minimum above is again over the empty set and is defined to be 1. Finally, and again similarly, if the maximum above is over the empty set then it is also defined to be 1. Note that if $S$ contains unit vectors, this is a rather trivial situation too, similar to what happens with other complexity measures $\sigma$ (see [37]) and $\rho$ (defined in Section 3, see [36]). If $e_j$ is in $S$, we simply remove the $j$th component and work with the restriction of $S$ (and $S^\perp$) to $\mathbb{R}^{n-1}$.

**Proof.** As we noted above, we can assume $|B| \geq 2$ and that $S$ contains no unit vectors. Let $J^*, i^*, s^*$ determine the right-hand-side above (the existence of such a solution follows from the LP duality theorem applied to a set $J \cup \{i\}$ identifying a nonnegative, minimal element in $S$—also see below). Then the value of the right-hand-side is the optimal value of the linear optimization problem

$$\begin{array}{ll} \text{maximize} & t \\ \text{subject to} & A_{J^*}^T w + et \leq 0 \\ & A_{i^*}^T w + t \leq 1. \end{array}$$

Let $w^* \in \mathbb{R}^m$ such that $s^* = -A^T w^*$, and define $t^* := \min\left\{\min_{j \in J^*}\{s_j^*\}, (1+s_{i^*}^*)\right\} > 0$. Then $(w^*, t^*)$ is an optimal solution of the linear optimization problem above. Conversely, an optimal solution $(w^*, t^*)$ of the above linear optimization problem determines the optimal $s^*$ (for the fixed pair $(J^*, i^*)$) in the right-hand-side above. The dual of the linear optimization problem is

$$\begin{array}{ll} \text{minimize} & x_{i^*} \\ \text{subject to} & A_{J^*} x_{J^*} + A_{i^*} x_{i^*} = 0 \\ & e^T x_{J^*} + x_{i^*} = 1 \\ & x_{J^*}, x_{i^*} \geq 0. \end{array}$$

Choose an extreme point solution $\bar{x}_{J^*}, \bar{x}_{i^*}$ which is optimal in this problem. Obviously, $\bar{x}_{i^*} > 0$ by the duality theorem of linear programming. Define $\bar{x} \in \mathbb{R}^n$ by completing with zeros. This is a minimal linear dependence verifying $A\bar{x} = 0, e^T\bar{x} = 1, \bar{x} \geq 0$; so, $\xi_P(A)$ is at most the right-hand-side.

Now we take $\bar{d} \in \mathbb{R}^n$ defining $\xi_P(A)$. Identify $i$ such that $\xi_P(A) = \bar{d}_i$. Take $J \cup \{i\}$ the set of nonzero components of $\bar{d}$. The system of equations $Ad = 0$ can be written as $A_J d_J + A_i d_i = 0$ where the columns of $A_J$ are linearly independent, because $\bar{d}$ is a minimal linear dependence. The problem

$$\begin{array}{ll} \text{minimize} & d_i \\ \text{subject to} & A_J d_J + A_i d_i = 0 \\ & e^T d_J + d_i = 1 \\ & d_J, d_i \geq 0 \end{array}$$

has a unique solution given by $\bar{d}_J, \bar{d}_i$. The dual of this problem is

$$\begin{array}{ll} \text{maximize} & t \\ \text{subject to} & A_J^T w + et \leq 0 \\ & A_i^T w + t \leq 1. \end{array}$$

Now take an optimal solution $(\bar{w}, \bar{t})$ with optimal value $\bar{t} = \xi_P(A)$ which is at least the right-hand-side. $\square$

Since the above theorem applies to all subspaces and our definition of $\xi$ can be written in terms of $\xi_P(S)$ and $\xi_P(S^\perp)$, we have

**Corollary 2.1** *Let $A \in \mathbb{R}^{m\times n}$ be given. Also let $S := \mathcal{N}(A)$. Then,*

$$\xi(A) = \min \left\{ \begin{array}{c} \begin{array}{ccc} \min & \max & \min\left\{\min\limits_{j\in J}\{s_j\}, (1+s_i)\right\}, \\ \emptyset \neq J \subset B & s \in S^\perp & \\ i \in (B\backslash J) & \text{sign}(s_j) = +, \forall j \in J & \\ & \text{sign}(s_i) = - & \end{array} \\ \\ \begin{array}{ccc} \min & \max & \min\left\{\min\limits_{j\in J}\{x_j\}, (1+x_i)\right\} \\ \emptyset \neq J \subset N & x \in S & \\ i \in (N\backslash J) & \text{sign}(x_j) = +, \forall j \in J & \\ & \text{sign}(x_i) = - & \end{array} \end{array} \right\}.$$

Even though the above algebraic description seems long-winded, the geometric interpretation of this dual characterization is quite nice and is included in Section 5 together with the geometric characterizations of other condition and complexity measures.

# 3 Characterizations of $\bar{\chi}$

Let $\mathcal{D}$ denote the set of $n \times n$ positive definite, diagonal matrices. We define

$$\bar{\chi}_p(A) := \sup\{\|A^T(ADA^T)^{-1}AD\|_p \ : \ D \in \mathcal{D}\}, \ p \in [1, +\infty].$$

We easily have

$$\bar{\chi}_p(A) = \sup\left\{\frac{\|A^T y\|_p}{\|c\|_p} : y \in \operatorname{argmin}\|D^{\frac{1}{2}}(A^T y - c)\|_2, c \in \mathbb{R}^n, D \in \mathcal{D}\right\}. \tag{4}$$

The complexity measure $\bar{\chi}(A) := \bar{\chi}_2(A)$ has been studied by many researchers (see for instance Vavasis [38], Vavasis-Ye [40, 39] and Ho-Tunçel [20]; also see Forsgren [10] for a historical account up to early 1990s). The next result was stated in [36] for the 2-norm only.

**Lemma 3.1** *For every $p \in [1, +\infty]$, we have $\bar{\chi}_p(A) = \max\{\|A^T A_B^{-T}\|_p \ : \ B \in \mathcal{B}(A)\}$.*

**Proof.** Proofs of Lemma 1 (also see Todd [35]) and Lemma 2 of [36] (also see Vavasis-Ye [39]) go through with an arbitrary $p$-norm in place of the 2-norm. □

In [23], Megiddo and Shub studied the behavior of large variables of the least square solutions. For a sequence $\{d^{(k)}\}$ in $\mathbb{R}^n_{++}$ converging to $\bar{d} \in \mathbb{R}^n_+$, they demonstrate among other things the following fact.

**Lemma 3.2** *(Megiddo and Shub [23]) Let $J := J_+(\bar{d}) \neq \emptyset$. Then the sequence $\{y(d^{(k)})\}$ defined by*

$$y(d^{(k)}) := \operatorname{argmin}\{\|D_k^{\frac{1}{2}}(A^T y - c)\|_2\},$$

*where $D_k = \operatorname{Diag}(d^{(k)})$, converges to $y(\bar{d}) = \operatorname{argmin}\{\|\bar{D}_J^{\frac{1}{2}}(A_J^T y - c_J)\|_2\}$.*

Let $X := \{Du \ : \ u \in \mathcal{N}(A), D \in \operatorname{cl}(\mathcal{D})\}$, $Y_p := \{v \in \mathcal{R}(A^T) \ : \ \|v\|_p = 1\}$,

$$\rho_{p,q}(A) := \inf\{\|u - v\|_q \ : \ u \in X, v \in Y_p\}.$$

We also define

$$\rho_p(A) := \rho_{p,p}(A).$$

The fact that $\bar{\chi}_2(A) = 1/\rho_2(A)$ was established by Stewart and O'Leary.

**Proposition 3.1** *(O'Leary [25]) Denoting the variables by $u$, $v$ and $J$ we have*

$$\begin{array}{rc} \rho_{p,q}(A) = & \min \quad \|v_J\|_q \\ & \operatorname{sign}(v_j) = \operatorname{sign}(u_j), \forall j \notin J \\ & \|v\|_p = 1 \\ & v \in \mathcal{R}(A^T) \\ & u \in \mathcal{N}(A) \\ & \emptyset \neq J \subset \{1, \ldots, n\}. \end{array}$$

**Proposition 3.2** *In the above proposition, there always exists a maximal $u \in \mathcal{N}(A)$ which attains the minimum.*

**Proof.** Let $u \in \mathcal{N}(A), v \in \mathcal{R}(A), J \subset \{1, 2, \ldots, n\}$ attain the minimum. If $u$ is not maximal then there exists $\bar{u} \in \mathcal{N}(A)$ maximal such that $J_+(\bar{u}) \supseteq J_+(u)$ and $J_-(\bar{u}) \supseteq J_-(u)$. Then $(\bar{u}, v, J \cup (J_0(u) \backslash J_0(\bar{u})))$ is also an optimal solution. □

**Theorem 3.1** *For every $p \in [1, +\infty]$, we have $\bar{\chi}_p(A) = \frac{1}{\rho_p(A)}$.*

**Proof.** Take $v = A^T y$ and $c$ satisfying the maximum in (4). Then there exists a sequence $\{d^{(k)}\}$ in $\mathbb{R}^n_{++}$ such that the sequence $\{y(d^{(k)})\}$ of least squares solutions associated with $d^{(k)}$ converges to $y$. So, the sequence $\{v^{(k)}\}$ defined by $v^{(k)} := A^T y(d^{(k)})$ converges to $v$. Define the sequence $\{u^{(k)}\}$ by $u^{(k)} := v^{(k)} - c$. We claim that for each $k$, $u^{(k)} \in \mathcal{N}(AD_k)$. In fact, $AD_k u^{(k)} = AD_k(A^T(AD_kA^T)^{-1}AD_kc - c) = 0$. By definition of $\{u^{(k)}\}$, $\|c\|_p = \|v^{(k)} - u^{(k)}\|_p$ for all $k$, so

$$\left\| \frac{v^{(k)}}{\|v^{(k)}\|_p} - \frac{u^{(k)}}{\|v^{(k)}\|_p} \right\|_p = \frac{\|c\|_p}{\|v^{(k)}\|_p} = \frac{\|c\|_p}{\|A^T y(d^{(k)})\|_p}.$$

Taking limits, we obtain

$$\frac{1}{\bar{\chi}_p(A)} = \left\| \frac{v}{\|v\|_p} - \frac{u}{\|v\|_p} \right\|_p.$$

Since $v/\|v\|_p \in Y_p$ and $u/\|v\|_p \in X$ we obtain $\frac{1}{\bar{\chi}_p(A)} \geq \rho_p(A)$.

Now, we take $v, u$ and $J$ optimal in Proposition 3.1. Let $\hat{y} \in \mathbb{R}^m$ such that $v = A^T \hat{y}$. We have $\rho_p(A) = \|v_J\|_p$. Consider a sequence $\{\epsilon_k\}$ in $\mathbb{R}_+$ converging to zero. Define the sequence of diagonal matrices $\{D_k\}$ by $(D_k)_{ii} := \epsilon_k$ if $i \notin J$; $(D_k)_{ii} = 1$ if $i \in J$. Consider $\hat{v} \in \mathbb{R}^n$ defined as $\hat{v}_J := v_J$ and $\hat{v}_j := 0$ for all other $j$. Also consider the sequences $\{y^{(k)}\}$ defined by $y^{(k)} := \operatorname{argmin}\{\|D_k^{\frac{1}{2}}(A^T y - \hat{v})\|_2\}$, and $\{v^{(k)}\}$ defined by $v^{(k)} := A^T y^{(k)}$. By Lemma 3.2, the limit of $y^{(k)}$ is given by $\text{argmin}\{\|A_J^T y - v_J\|_2\}$. We claim that this limit coincides with $\hat{y}$. To see it, note that the sequence $\{\tilde{y}^{(k)}\}$ defined by $\tilde{y}^{(k)} := \text{argmin}\{\|D_k^{1/2}(A^T y - v)\|_2\}$ has the same limit; moreover, for each $k$, $A^T\tilde{y}^{(k)}$ is the oblique projection of $v$ onto $\mathcal{R}(A^T)$. Since $v \in \mathcal{R}(A^T)$ then the projection must be itself. This proves that $y^k \to \hat{y}$ and so $\{v^{(k)}\} \to v$. By hypothesis, $\|v\|_p = \|A^T\hat{y}\|_p = 1$ and $\rho_p(A) = \|v_J\|_p = \|A_J^T\hat{y}\|_p$. Then

$$\frac{1}{\rho_p(A)} = \frac{\|A^T\hat{y}\|_p}{\|v_J\|_p} = \frac{\|A^T\hat{y}\|_p}{\|\hat{v}\|_p} \leq \bar{\chi}_p(A).$$

The inequality above follows from the characterization (4). □

Gonzaga and Lara [14] proved that for $p = 2$, $\bar{\chi}_p(S)$ and $\bar{\chi}_p(S^\perp)$ coincide. For general norms this behavior is not preserved, even if we consider dual norm type relations: Consider the matrix $A := \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$. Here, $\mathcal{N}(A) = S$ and $\mathcal{R}(A^T) = S^\perp$. Then the matrix $Z := \begin{pmatrix} -2 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix}$ is a null space matrix of $A$. Using Lemma 3.1, we have $\bar{\chi}_1(S) = \max\left\{\|A^T A_B^{-T}\|_1 : B \in \mathcal{B}(A)\right\} = 6$, $\bar{\chi}_2(S) = \sqrt{14}$, and $\bar{\chi}_\infty(S) = 3$, while $\bar{\chi}_1(S^\perp) = \max\left\{\|Z^T Z_B^{-T}\|_1 : B \in \mathcal{B}(Z)\right\} = 4$, $\bar{\chi}_2(S^\perp) = \sqrt{14}$ and $\bar{\chi}_\infty(S^\perp) = 5$. This example shows that in general $\bar{\chi}_p(S) \neq \bar{\chi}_p(S^\perp)$, $\bar{\chi}_p(S) \neq \bar{\chi}_{p/(p-1)}(S^\perp)$, $\rho_{p,q}(S) \neq \rho_{p,q}(S^\perp)$ and $\rho_{p,q}(S) \neq \rho_{q,p}(S^\perp)$.

# 4 Infeasibility Detection via a Constructive Proof of a Helly-type Theorem

Consider the feasibility problem (2). In this case, Helly's Theorem (see for instance [6]) implies that "$A^T y \leq c$ is infeasible iff there exists $J \subseteq \{1, 2, \ldots, n\}$, $|J| \leq (m+1)$ such that $A_J^T y \leq c_J$ is infeasible."

Tardos' scheme to solve the feasibility problem (2), is a constructive proof of a Helly-type theorem (see [20]). The scheme solves "easier" related systems, recursively. In each iteration, the scheme identifies at least one constraint to drop. We will outline such a scheme after the next lemma.

**Lemma 4.1** *Let $c \in \mathcal{N}(A)$. Then for every $s \in \mathbb{R}^n$ such that $A^T y + s = c$ (for some $y \in \mathbb{R}^m$) we have $\|s\|_2 \geq \|c\|_2$.*

**Proof.** Note that $A^T y + s = c$ implies $c^T A^T y + c^T s = c^T c$ which in turn implies $c^T s = c^T c$ (since $c \in \mathcal{N}(A)$). Now, by Cauchy-Schwarz inequality, we obtain

$$c^T c = c^T s \leq \|c\|_2 \|s\|_2.$$

If $c = 0$ then the claimed inequality is clearly true; otherwise, we divide both sides by $\|c\|_2$ and we obtain $\|s\|_2 \geq \|c\|_2$, as desired. $\square$

This lemma implies that in Section 7.2 of [20], it suffices to choose $p \geq \sqrt{n}\left(\frac{1}{\xi_P(A)} + 1\right)$ rather than $p \geq 2n^{3/2}(\bar{\chi}(A))^2$.

In answering the question posed by (2), first, we can replace $c$ by its orthogonal projection onto $\mathcal{N}(A)$. Secondly, if $c = 0$ then $\bar{y} := 0$ solves (2). Therefore, we can assume $\|c\|_\infty = 1$. We choose $p \geq \sqrt{n}\left(\frac{1}{\xi_P(A)} + 1\right)$ and solve the system

$$A^T y + s = \lceil pc \rceil, \;\; s \geq 0. \tag{5}$$

If the system (5) is infeasible then so is the system (2). Otherwise, every solution $(\bar{y}, \bar{s})$ of the system (5) satisfies

$$A^T \bar{y} + \bar{s} + pc - \lceil pc \rceil = pc.$$

Thus, by Lemma 4.1,

$$\sqrt{n}\,\|\bar{s} + pc - \lceil pc \rceil\|_\infty \geq \|\bar{s} + pc - \lceil pc \rceil\|_2 \geq \|pc\|_2 \geq p.$$

Therefore,

$$\|\bar{s}\|_\infty \geq \frac{p}{\sqrt{n}} - \|pc - \lceil pc \rceil\|_\infty > \frac{1}{\xi_P(A)}.$$

We have

**Theorem 4.1** *Let $J := \left\{ j \in \{1, 2, \ldots, n\} : \bar{s}_j < \frac{1}{\xi_P(A)} \right\}$. Then the system $A^T y \leq c$ has a feasible solution iff the system $A_J^T y \leq c_J$ does.*

**Proof.** Essentially the same as the proof of Lemma 7.3 of [20]. $\square$

To appreciate this improvement, we offer the following refinements of a theorem of Todd, Tunçel and Ye [36]. First we need a definition. Let $G \subseteq \{1, 2, \ldots, n\}$. If we multiply by $(-1)$ the columns of $A$ indexed by $G$, then the resulting matrix (this was called the *signing* of $A$ by $G$ and denoted $A_{-G}$ in [36]) has the same $\bar{\chi}$ value as $A$. We write $\underline{\sigma}(A)$ for $\min_{G \subseteq \{1,2,\ldots,n\}} \sigma(A_{-G})$. Analogously, we write $\underline{\xi}(A)$ for $\min_{G \subseteq \{1,2,\ldots,n\}} \xi(A_{-G})$ and $\underline{\xi}_P(A)$ for $\min_{G \subseteq \{1,2,\ldots,n\}} \xi_P(A_{-G})$.

**Theorem 4.2** *For all $A \in \mathbb{R}^{m \times n}$ with* $\operatorname{rank}(A) = m$, *we have*

$$\underline{\xi}_P(A) = \frac{1}{\bar{\chi}_\infty(A) + 1} = \frac{\rho_\infty(A)}{1 + \rho_\infty(A)}.$$

**Proof.** First we prove "$\geq$". Since $\bar{\chi}_\infty(A)$ is invariant under signings of $A$, for this part of the proof we can assume $\underline{\xi}_P(A) = \xi_P(A)$. Let $\bar{x} \in \mathbb{R}^n$ determine $\xi_P(A)$ such that $\bar{x} \in \mathcal{N}(A)$, $e^T\bar{x} = 1$, $\text{sign}(\bar{x}) \in \underline{\text{sign}}_+(\mathcal{N}(A))$ and there exist $B \in \mathcal{B}(A)$, $k \notin B$ such that $\bar{x}_k = \xi_P(A)$ and all the nonzero components of $\bar{x}$ are determined by the unique solution of the linear system

$$A_B x_B = -\xi_P(A) A_k.$$

Then

$$x_B = -\xi_P(A) A_B^{-1} A_k \geq 0$$

and

$$e^T\bar{x}_B = \xi_P(A) \|A_B^{-1}A_k\|_1 \leq \xi_P(A)\bar{\chi}_\infty(A).$$

Since $e^T\bar{x}_B + \xi_P(A) = 1$, we get $1 \leq \xi_P(A)\left(\bar{\chi}_\infty(A) + 1\right)$. Thus we obtain,

$$\underline{\xi}_P(A) \geq \frac{1}{\bar{\chi}_\infty(A) + 1}$$

as desired.

To prove the reverse inequality, let $B \in \mathcal{B}(A)$, and $k \notin B$ determine $\bar{\chi}_\infty(A)$. Then pick a signing of $A$ corresponding to the diagonal $m \times m$ matrix $G$ such that $\text{diag}(G) \in \{-1, 1\}^m$ and

$$GA_B^{-1}A_k \leq 0.$$

(Note that for $j \notin B$, $A_j$ is not signed.) For the same basis $B$ of the signed matrix $A$, we have the linear system

$$A_B G x_B = -A_k$$

which has a unique solution (hence it determines a minimal linear dependence amongst the columns of $A$). Upon normalizing by $(e^T x_B + 1)$ and focusing on the component corresponding to $A_k$, we have

$$\underline{\xi}_P(A) \leq \frac{1}{e^T x_B + 1} = \frac{1}{\bar{\chi}_\infty(A) + 1}.$$

The rest of the statement of the theorem follows from Theorem 3.1. $\square$

**Theorem 4.3** *Let $A \in \mathbb{R}^{m\times n}$, $\text{rank}(A) = m$. Then*

$$\tfrac{1}{\sqrt{n}}\underline{\xi}(A) \;=\; \tfrac{1}{\sqrt{n}}\underline{\sigma}(A) \;\leq\; \rho_2(A) \;\leq\; \underline{\xi}(A) \;=\; \underline{\sigma}(A) \;\leq\; \xi(A) \leq \sigma(A).$$

A proof of the above theorem is included in Appendix A. It follows from the above theorem that the probabilistic analysis of [36] also applies to the complexity measures $\xi(A)$ and $\underline{\xi}(A)$. We note that there are instances $A(\epsilon)$ such that as $\epsilon \to 0$, $\xi(A(\epsilon)) \to 0$, even though $\sigma(A(\epsilon)) \to \frac{1}{2}$ (this can be easily seen using the geometric characterizations of $\sigma$ and $\xi$ presented in Section 6). Also, there are instances $\tilde{A}(\epsilon)$ such that as $\epsilon \to 0$, $\underline{\sigma}(\tilde{A}(\epsilon)) \to 0$ even though $\xi(\tilde{A}(\epsilon))$ stays $\Omega(1)$.

The above theorems and the discussion also imply that even though $\xi_P(A)$ and $\xi_D(A)$ can take very different values, $\underline{\xi}_P(A)$ and $\underline{\xi}_D(A)$ must always be "close" to each other. (Use, for instance, the fact that $\rho_2(S) = \rho_2(S^\perp)$.) Similarly, Theorem 2 of [36] and the discussion above imply that even though $\sigma_P(A)$ and $\sigma_D(A)$ can take very different values, $\underline{\sigma}_P(A)$ and $\underline{\sigma}_D(A)$ must always be "close" to each other.

To conclude this section, we note that Tardos' scheme is applied to $A$ in a way that the complexity measures of submatrices obtained from $A$ by column deletion are important. As it was shown in Proposition 2.4 of [20], $\bar{\chi}(\cdot)$ is monotone nonincreasing under this operation. Here, we note that the same property extends to $1/\xi_P(\cdot)$ (the property also extends to $1/\sigma_P(\cdot)$ whose proof is omitted).

**Theorem 4.4** *Let $A \in \mathbb{R}^{m\times n}$, $\operatorname{rank}(A) = m$. Suppose $\tilde{A}$ is obtained from $A$ by removing a column of $A$. Then we have the following facts.*
*(a) If $\operatorname{rank}(\tilde{A}) = m$ then $\xi_P(\tilde{A}) \geq \xi_P(A)$.*
*(b) If $\operatorname{rank}(\tilde{A}) \leq m-1$ then let $\bar{A}$ be obtained from $\tilde{A}$ by removing any linearly dependent row. Then $\operatorname{rank}(\bar{A}) = m-1$ and $\xi_P(\bar{A}) = \xi_P(A)$.*

**Proof.** If $\operatorname{rank}(\tilde{A}) = m$ then the set of minimal elements of $\mathcal{N}(\tilde{A})$ can be extended to a subset of the set of minimal elements of $\mathcal{N}(A)$ (by appending an appropriate zero). Therefore, $\xi_P(\tilde{A}) \geq \xi_P(A)$ trivially follows. If $\operatorname{rank}(\tilde{A}) \leq m-1$ then without loss of generality, assume that the deleted column was $A_n$. Then there exists an $m \times m$ nonsingular matrix $G$ such that

$$\xi_P(A) = \xi_P(GA) = \xi_P\begin{pmatrix} A' & 0 \\ 0^T & 1 \end{pmatrix},$$

where $\operatorname{rank}(A') = \operatorname{rank}(\tilde{A}) = m-1$. We easily see from the structure of $GA$ that every vector $x$ in the null space of $A$ must have $x_n = 0$. Hence

$$\xi_P(A') = \xi_P(\bar{A}) = \xi_P(A).$$

$\square$

The proof of Proposition 2.4 of [20] also trivially extends to $\bar{\chi}_p(\cdot)$ for all $p \in [1, +\infty]$ by utilizing our Lemma 4.1 and Theorem 4.1 from the current paper.

Note that the analogous results apply to $\xi_D(A)$. Therefore, $\xi(A)$ can be used in the above-mentioned manner as a complexity measure in Tardos' scheme to determine the feasibility of the systems (1) and (2).

# 5 Geometric Interpretations and Sensitivity Analysis

Epelman and Freund [8] showed that the complexity measure of Ye is very closely related to the symmetry measure interpreted as follows.

Let

$$\mathrm{sym}(A) := \max\left\{t : -tv \in \mathrm{conv}\{A_i : i \in B\},\ \text{for all } v \in \mathrm{conv}\{A_i : i \in B\}\right\}.$$

$\mathrm{sym}(A)$ measures the symmetry of $\mathrm{conv}\{A_i : i \in B\}$ about the origin in $\mathbb{R}^m$ (symmetry of a compact convex set containing the origin in its interior is defined analogously—this is used in Section 6). When $\mathrm{conv}\{A_i : i \in B\}$ is centrally symmetric about the origin, we have $\mathrm{sym}(A) = 1$.

**Theorem 5.1** *(Epelman and Freund [8]) Let $A \in \mathbb{R}^{m\times n}$ be such that the corresponding strict complementarity partition $[B, N]$ satisfies $B \neq \emptyset$. Then*

$$\sigma_P(A) = \frac{\mathrm{sym}(A)}{1 + \mathrm{sym}(A)}.$$

Note that the above theorem and the fact that $\sigma_P(A)$ only depends on the pair of subspaces $\mathcal{N}(A)$, $\mathcal{R}(A)$, imply that we can define $\mathrm{sym}(\cdot)$ for subspaces. Thus, we have $\mathrm{sym}(A) = \mathrm{sym}(S)$ for all $A$ such that $\mathcal{N}(A) = S$.

We denote by $r_P(A;Q)$ the radius of the largest ball with respect to the norm induced by $Q \in \Sigma^m_{++}$, contained in $\mathrm{conv}\{A_i : i \in B\}$ and centered at the origin. Similarly, $R_P(A;Q)$ is the radius of the smallest ball (with respect to the norm induced by $Q \in \Sigma^m_{++}$) containing $\mathrm{conv}\{A_i : i \in B\}$ and centered at the origin. The following are alternative descriptions:

$$r_P(A;Q) = \max\left\{r \in \mathbb{R} : \{u \in \mathbb{R}^m : \|u\|_Q \leq r\} \subseteq \mathrm{conv}\{A_i : i \in B\}\right\},$$

$$R_P(A;Q) = \min\left\{R \in \mathbb{R} : \{u \in \mathbb{R}^m : \|u\|_Q \leq R\} \supseteq \mathrm{conv}\{A_i : i \in B\}\right\},$$

where we define $\|u\|_Q := \left(u^T Q^{-1} u\right)^{\frac{1}{2}}$. We easily have

$$\mathrm{sym}(A) \quad \geq \quad \sup_{Q \in \Sigma^m_{++}} \frac{r_P(A;Q)}{R_P(A;Q)}. \tag{6}$$

This allows us to make immediate geometric statements about sensitivity analysis of $\mathrm{sym}(\cdot)$ and hence $\sigma(\cdot)$. Instead of the ellipsoidal norms above, we could use any norm induced by a compact convex set in $\mathbb{R}^m$ containing the origin in its interior such that the set is symmetric about the origin. Then if we take the supremum in (6) over all such convex bodies, the analogous statement to (6) would hold with equality (see Epelman and Freund [8]).

Any perturbation made to $A_N$ which does not change the strict complementarity partition $[B, N]$ does not change the value $\sigma(A)$ either. On the other hand, changes to $A_B$ can be analyzed geometrically using the above lower bound on $\mathrm{sym}(A)$. For instance, suppose each $A_i$ for $i \in B$ is perturbed along a direction $(\Delta A)_i$ such that $\|(\Delta A)_i\|_2 = 1$ (or zero) for every $i \in B$. Also assume that the over all perturbation (including the perturbations to $A_N$) has the property that there exists a positive $\bar{\alpha}$ such that $[A + \alpha(\Delta A)]$ determines the same partition $[B, N]$ as $A$ does, for every $\alpha \in [0, \bar{\alpha}]$. We would like to understand how $\mathrm{sym}(A(\alpha))$ behaves for small $\alpha$.

First, note that the supremum in (6) is attained. Also, we can restrict $Q$ to those positive definite matrices with the largest eigenvalue equal to 1. Secondly, let $\bar{Q}$ denote a symmetric positive definite matrix attaining that supremum and denote by $\lambda_i$ the $i$th largest eigenvalue of $\bar{Q}$.

**Proposition 5.1** *Let $A$, $\Delta A$, $\bar{Q}$ and $\lambda_m$ be as defined above. Then there exists $\bar{\alpha} \in (0,1)$ such that for every $\alpha \in [0, \bar{\alpha}]$,*

$$\mathrm{sym}(A(\alpha)) \geq \frac{1}{\sqrt{m}} \left( \frac{r_P(A;\bar{Q})\sqrt{\lambda_m} - \alpha}{R_P(A;\bar{Q})\sqrt{\lambda_m} + \alpha} \right). \tag{7}$$

**Proof.** Let $\bar{u} \in \left\{ u \in \mathbb{R}^m : u^T \bar{Q}^{-1} u \leq \left[R_P(A;\bar{Q})\right]^2 \right\}$. Then for every $v \in \mathbb{R}^m$ such that $\|v\|_2 \leq \alpha \leq \bar{\alpha}$, we have

$$\begin{aligned} (\bar{u}+v)^T \bar{Q}^{-1} (\bar{u}+v) &= \bar{u}^T \bar{Q}^{-1} \bar{u} + 2\bar{u}^T \bar{Q}^{-1} v + v^T \bar{Q}^{-1} v \\ &\leq \left[R_P(A;\bar{Q})\right]^2 + \frac{\alpha}{\lambda_m} + 2\|\bar{Q}^{-1/2} v\|_2 \|\bar{Q}^{-1/2}\bar{u}\|_2 \\ &\leq \left( R_P(A;\bar{Q}) + \frac{\alpha}{\sqrt{\lambda_m}} \right)^2 . \end{aligned}$$

This implies that the ellipsoid $\left\{ u \in \mathbb{R}^m : \|u\|_{\bar{Q}} \leq R_P(A;\bar{Q}) + \frac{\alpha}{\sqrt{\lambda_m}} \right\}$ contains $\mathrm{conv}\left\{[A(\alpha)]_i : i \in B\right\}$.

Let $u^{(j)}$ denote the eigenvectors of $\bar{Q}$. Then the vectors $\pm r_P(A;\bar{Q}) u^{(j)} \in \mathrm{conv}\{A_i : i \in B\}$, for every $j \in \{1, 2, \ldots, n\}$. Thus, $\pm \left[r_P(A;\bar{Q})\sqrt{\lambda_j} - \alpha\right] u^{(j)} \in \mathrm{conv}\left\{[A(\alpha)]_i : i \in B\right\}$, for every $j \in \{1, 2, \ldots, m\}$. Hence, by standard arguments, we conclude that the ellipsoid

$$\left\{ u \in \mathbb{R}^m : \|u\|_{\bar{Q}} \leq \frac{1}{\sqrt{m}} \left( r_P(A;\bar{Q}) - \frac{\alpha}{\sqrt{\lambda_m}} \right) \right\}$$

is contained in $\mathrm{conv}\left\{[A(\alpha)]_i : i \in B\right\}$.

Finally, using the above ellipsoids, we conclude the desired lowerbound. $\square$

The above result is simply meant to motivate the sensitivity analysis of $\sigma(A)$ (which is coming up) using the elegant geometric setting from Epelman-Freund [8]. In the same setting, we have the following geometric interpretations.

Now, we are ready to give a geometric interpretation of the definition of $\xi_P(A)$. Consider all simplices made from ext (conv$\{A_i : i \in B\}$), containing the origin in their relative interior. Let $\mathcal{S}(A)$ denote the set of all such simplices. For each $\Delta \in \mathcal{S}(A)$, consider the (unique) barycentric coordinates of 0 in $\Delta$. Assign the smallest of these coordinates to $\Delta$. The smallest number assigned to any such simplex is equal to $\xi_P(A)$.

Next, we give a geometric interpretation of the dual characterization of $\xi_P(A)$ provided by Theorem 2.1 and Corollary 2.1. Consider pairs $(J, i)$ such that $J \subset B$, $i \in B \backslash J$, and the point $A_i$ and the set conv$\{A_j : j \in J\}$ can be strictly separated by a hyperplane through the origin. For each such hyperplane, consider the "distances" from $A_i$ and conv$\{A_j : j \in J\}$ to the hyperplane. Choose the hyperplane which maximizes the ratio of the smaller to the larger of these distances, and assign this ratio to the pair $(J, i)$. Now, $\xi_P(A)$ is basically the minimum of all these ratios.

## 5.1 Perturbation matrices

Now, we begin studying the effect that the perturbations on the data of the problem $(FP)$ can have on some of the complexity measures for linear feasibility and linear programming problems. Since the definitions of these condition numbers can be made to depend on the sign patterns of the subspaces defined by $A$, we deal with perturbations that preserve the sign patterns of the original data. Consider the perturbation matrix $\Delta A \in \mathbb{R}^{m \times n}$. That is, in problem $(FP)$ the matrix $A$ is replaced by $A(\alpha) := (A + \alpha(\Delta A))$, for small positive $\alpha$. The following conditions on certain extreme elements of the subspaces defined by $A$ and $A(\alpha)$ respectively will be very important in our treatment.

Condition $C1$: The perturbation matrix $\Delta A$ is said to satisfy condition $C1$ for $A$ if there exists $\bar{\alpha} > 0$ such that for all $\alpha \in (0, \bar{\alpha})$,

$$\underline{\text{sign}}(\mathcal{N}(A)) = \underline{\text{sign}}(\mathcal{N}(A + \alpha(\Delta A))).$$

Condition $C2$: The perturbation matrix $\Delta A$ is said to satisfy condition $C2$ for $A$ if there exists $\bar{\alpha} > 0$ such that for all $\alpha \in (0, \bar{\alpha})$,

$$\overline{\text{sign}}(\mathcal{N}(A)) = \overline{\text{sign}}(\mathcal{N}(A + \alpha(\Delta A))).$$

Condition $C3$: The perturbation matrix $\Delta A$ is said to satisfy condition $C3$ for $A$ if there exists $\bar{\alpha} > 0$ such that for all $\alpha \in (0, \bar{\alpha})$,

$$\underline{\text{sign}}_+(\mathcal{N}(A)) = \underline{\text{sign}}_+(\mathcal{N}(A + \alpha(\Delta A))).$$

Condition $C4$: The perturbation matrix $\Delta A$ is said to satisfy condition $C4$ for $A$ if there exists $\bar{\alpha} > 0$ such that for all $\alpha \in (0, \bar{\alpha})$,

$$\overline{\text{sign}}_+(\mathcal{N}(A)) = \overline{\text{sign}}_+(\mathcal{N}(A + \alpha(\Delta A))).$$

Conditions $C1$ and $C2$ are relevant to perturbation results involving $\bar{\chi}$; condition $C3$ is relevant to $\xi$ and condition $C4$ (which says that the strict complementarity partition $[B, N]$ does not change) is relevant to $\sigma$.

Note that if $\Delta A$ satisfies $C1$ for $A$, then for every $\alpha \in [0, \bar{\alpha})$ the strict complementarity partition determined by $(A + \alpha(\Delta A))$ is the same as the one determined by $A$.

**Theorem 5.2** *Let $A, (\Delta A) \in \mathbb{R}^{m \times n}$. Then we have the following facts.*
*(a) For every $\alpha \in \mathbb{R}$, we have*

$$\underline{\mathrm{sign}}(\mathcal{N}(A)) = \underline{\mathrm{sign}}(\mathcal{N}(A + \alpha(\Delta A))) \text{ iff } \underline{\mathrm{sign}}(\mathcal{R}(A^T)) = \underline{\mathrm{sign}}(\mathcal{R}(A^T + \alpha(\Delta A)^T)).$$

*(b) For every $\alpha \in [0, \bar{\alpha})$, $\overline{\mathrm{sign}}_+(\mathcal{N}(A)) = \overline{\mathrm{sign}}_+(\mathcal{N}(A + \alpha(\Delta A)))$*
*if and only if*
*for every $\alpha \in [0, \bar{\alpha})$, $\overline{\mathrm{sign}}_+(\mathcal{R}(A^T)) = \overline{\mathrm{sign}}_+(\mathcal{R}(A^T + \alpha(\Delta A)^T))$.*
*(c) Condition $C1$ implies conditions $C2, C3$ and $C4$.*
*(d) Condition $C2$ implies condition $C4$.*
*(e) Condition $C3$ implies condition $C4$.*

**Proof.** (a) $\underline{\mathrm{sign}}(\mathcal{N}(A)) = \underline{\mathrm{sign}}(\mathcal{N}(A + \alpha(\Delta A)))$ implies that these two sets of minimal sign vectors determine the same oriented matroid on $\{1, 2, \ldots, n\}$. The dual of the oriented matroid determined by $\underline{\mathrm{sign}}(\mathcal{N}(A))$ is $\mathrm{sign}(\mathcal{R}(A^T))$ (for more general results related to this fact see Bland and LasVergnas [4]; also see Proposition 3.4.1 and Lemma 3.4.2 of [3]). The latter is also determined by $\underline{\mathrm{sign}}(\mathcal{R}(A^T))$. Therefore, we have the statement (a).
(b) This follows from the facts that $A$ determines a strict complementarity partition $[B, N]$ and (of course) $B$ stays the same iff $N$ does.
(c) Since the minimal sign vectors completely determine all sign vectors in the pair of orthogonal subspaces, $C2, C3$ and $C4$ are consequences of $C1$.
(d) and (e) follow easily from the definitions. □

Note that the other direct implication relations amongst the conditions $C1$-$C4$ are false in general.

The next lemma establishes a relationship between the minimal vectors in $\mathcal{N}(A)$ and those in $\mathcal{N}(A + \alpha(\Delta A))$ in the case that $\Delta A$ satisfies condition $C4$ for $A$.

**Lemma 5.1** *Assume that $\Delta A$ satisfies condition $C4$ for $A$ and that $\underline{\mathrm{sign}}_+(\mathcal{N}(A)) \neq \emptyset$. Then for every $x \in \mathcal{N}(A)$ such that $\mathrm{sign}(x) \in \underline{\mathrm{sign}}_+(\mathcal{N}(A))$ there exist $\hat{\alpha} > 0$ and a path of solutions $\{x(\alpha) : \alpha \in (0, \hat{\alpha}]\}$ such that $x(\alpha) \in \mathcal{N}(A + \alpha(\Delta A))$, $\mathrm{sign}(x(\alpha)) \in \underline{\mathrm{sign}}_+(\mathcal{N}(A + \alpha(\Delta A)))$ and is constant for every $\alpha \in (0, \hat{\alpha}]$ and $x(\alpha) \to x$ as $\alpha \to 0$.*

**Proof.** Consider $x$ as in the statement of the lemma. Let $J := J_+(x)$. Then there exists $\hat{\alpha} \in (0, \bar{\alpha}]$ such that the columns of $[A_J + \alpha(\Delta A)_J]$ are either minimally linearly dependent or are linearly independent.

In the first case, we focus on the unique solution (path) $x_J(\alpha)$ of the system

$$[A_J + \alpha(\Delta A)_J]\, x_J = 0, \|x_J\|_1 = 1. \tag{8}$$

It is clear that $x(\alpha)$ (the extension to $\mathbb{R}^n$ of $x_J(\alpha)$ by completing with zeros), satisfies $\mathrm{sign}(x(\alpha)) \in \underline{\mathrm{sign}}_+(\mathcal{N}(A + \alpha(\Delta A)))$ for all small enough $\alpha$ and that $x(\alpha) \to x$.

In the second case, we note that the strict complementarity partition $[B, N]$ defined by $A$ has the property that $J \subset B$. So, we extend $J$ by adding to it $j \in B \backslash J$ until we have $\hat{J} \supset J$ (and $\hat{J} \subseteq B$) such that the columns of $\left[A_{\hat{J}} + \alpha(\Delta A)_{\hat{J}}\right]$ are minimally linearly dependent and the unique solution (path) $x(\alpha)$ determined by

$$\left[A_{\hat{J}} + \alpha(\Delta A)_{\hat{J}}\right] x_{\hat{J}} = 0, \|x_{\hat{J}}\|_1 = 1$$

converges to $x$. Here, we redefine $\hat{\alpha}$ if necessary (by reducing it) to ensure that $\mathrm{sign}(x(\alpha))$ stays the same for all $\alpha \in (0, \hat{\alpha}]$. If $x(\alpha) \geq 0$ then we are done. Otherwise, there exists $j \in \hat{J}$ such that $x_j(\alpha) < 0$ for all $\alpha \in (0, \hat{\alpha}]$; moreover (since $x(\alpha)$ converges to a nonnegative vector), $x_j = 0$ for every such $j$. Let $\tilde{x}(\alpha)$ be a maximal element of $\mathcal{N}(A + \alpha(\Delta A))$ such that $\mathrm{sign}(\tilde{x}(\alpha)) \in \overline{\mathrm{sign}}_+(\mathcal{N}(A + \alpha(\Delta A)))$ and $\tilde{x}(\alpha) \to \tilde{x}$, where $\mathrm{sign}(\tilde{x}) \in \overline{\mathrm{sign}}_+(\mathcal{N}(A))$. By Condition $C4$, $\mathrm{sign}(\tilde{x}) = \mathrm{sign}(\tilde{x}(\alpha))$, for every $\alpha \in (0, \hat{\alpha}]$. For $\alpha \in (0, \hat{\alpha}]$, we define

$$\gamma(\alpha) := \min\left\{\gamma \in [0,1] : x(\alpha) + \gamma\left[\tilde{x}(\alpha) - x(\alpha)\right] \geq 0\right\}.$$

Such $\gamma$ exists since $\tilde{x}_{\hat{J}}(\alpha) > 0$, for every $\alpha \in (0, \hat{\alpha}]$. We also define

$$u(\alpha) := \frac{1}{e^T\left\{x(\alpha) + \gamma(\alpha)\left[\tilde{x}(\alpha) - x(\alpha)\right]\right\}}\left\{x(\alpha) + \gamma(\alpha)\left[\tilde{x}(\alpha) - x(\alpha)\right]\right\}.$$

We have $u(\alpha) \to x$ as $\alpha \to 0$. We redefine $\hat{\alpha}$ if necessary (by reducing it) to ensure that $\mathrm{sign}(u(\alpha))$ stays the same for all $\alpha \in (0, \hat{\alpha}]$. We constructed a nonnegative path $\{u(\alpha) : \alpha \in (0, \hat{\alpha}]\}$ such that if $\mathrm{sign}(u(\alpha)) \in \underline{\mathrm{sign}}_+(\mathcal{N}(A + \alpha(\Delta A)))$ then we are done. If $u(\alpha)$ is not minimal, then note that $|J_+(u(\alpha))| \leq |B| - 1$. We remove all columns $A_j, (\Delta A)_j$ for every $j \in J_0(u(\alpha))$, redefine $B := J_+(u(\alpha))$ and repeat the above. As we keep iterating, the procedure will have to stop after at most $(n-2)$ steps leaving the last $u(\alpha)$ as the desired path. $\square$

The above lemma can be interpreted as a statement about how the extreme points of a polytope behave (in terms of the extreme points of certain "nearby" polytopes) under a perturbation.

Consider a minimal $x$ and $x(\alpha)$ for some $\alpha \in (0, \hat{\alpha}]$ as in the lemma above. Denote $J_+(\alpha) := J_+(x(\alpha))$. Clearly $J_0(x) \supseteq J_0(x(\alpha))$. $x_{J(0)}$ is the unique solution of the system (3) for $J = J_+(0)$

and $x_{J(\alpha)}(\alpha)$ is the unique solution of (8) for $J = J_+(\alpha)$. Choose a basis $B \in \mathcal{B}\,(A + \alpha(\Delta A))$ and $j_B \in J_+(0)\backslash B$ such that $J(\alpha) \subseteq B \cup \{j_B\}$. Let

$$M(\alpha) := \begin{pmatrix} [A(\alpha)]_B & [A(\alpha)]_{j_B} \\ e^T & 1 \end{pmatrix},$$

where $A(\alpha) := A + \alpha(\Delta A)$. This matrix is nonsingular for every $\alpha \in [0, \hat{\alpha}]$. We denote $M := M(0)$. Consider an arbitrary index $i \in J_+(\alpha)$ and the dual equations

$$A_J^T w_J + te = e_i, \tag{9}$$

where $J := B \cup \{j_B\}$ as above, and the perturbed ones

$$(A_J + \alpha(\Delta A)_J)^T w + te = e_i. \tag{10}$$

Associated to each $i \in J_+(\alpha)$ and the basis $B$ defined above, there exists a unique solution $(w^{(i)}, t^{(i)})$ of the system of equations (9). It is easy to see that $t^{(i)} = x_i$ (to see this claim we can express $x_i$ as the optimal value for the problem $\min\{x_i \,:\, A_J x_J = 0, e^T x_J = 1\}$; then the claim easily follows from the duality theorem of linear programming).

**Lemma 5.2** *Consider $A, (\Delta A) \in \mathbb{R}^{m\times n}$ such that $(\Delta A)$ satisfies $C4$ for $A$. Also consider a minimal $x \in \mathcal{N}(A)$ and $x(\alpha) \in \mathcal{N}(A(\alpha))$ such that $x \geq 0$, $x(\alpha) \geq 0$, the path $\{x(\alpha) : \alpha \in (0, \hat{\alpha}]\}$ satisfies all the properties mentioned in Lemma 5.1, and $x(\alpha) \to x$. Finally consider the strict complementarity partition $[B, N]$ determined by $A$. Then for each $\alpha \in [0, \hat{\alpha}]$ we have $N \subseteq J_0(x(\alpha)) \subseteq J_0(x)$, and if $i \in J_+(\alpha)$ then*

$$x_i(\alpha) = x_i - \alpha \left\{ \left[ w^{(i)} \right]^T (\Delta A)_J \left[M(\alpha)\right]^{-1} e_i \right\},$$

*where $J_+(\alpha) \subseteq J := \tilde{B} \cup \{j_{\tilde{B}}\}$ for some $j_{\tilde{B}} \in J_+(0)$ and $\tilde{B} \in \mathcal{B}(A)$, and $(w^{(i)}, t^{(i)})$ is the unique solution of (9).*

**Proof.** Consider a fixed $\alpha \in [0, \hat{\alpha}]$, $x$ and $x(\alpha)$ as in the statement of the lemma. Then we have $J_0(x) \supseteq J_0(x(\alpha)) \supseteq N$ as a consequence of condition $C4$ and the hypothesis. Let $i \in J_+(\alpha)$ (note that $J_-(x(\alpha)) = \emptyset$). For each $\alpha \in [0, \bar{\alpha}]$, $x_{J(\alpha)}(\alpha)$ is the unique solution of (8) for $J = J_+(\alpha)$. Consider the dual system of equations (10) with $J = B \cup \{j_B\}$ for some $j_B \in J_+(0)$, $J_+(0) \subseteq J_+(\alpha) \subseteq J$ and $B \in \mathcal{B}(A)$. Express the unique solution $(w^{(i)}(\alpha), t^{(i)}(\alpha))$ for these equations as $(w^{(i)}(\alpha), t^{(i)}(\alpha)) = (w^{(i)}, t^{(i)}) + (v^{(i)}(\alpha), \eta^{(i)}(\alpha))$. By substituting this relation in (10) and noticing that $(w^{(i)}, t^{(i)})$ satisfies the system (9), we conclude that $(v^{(i)}(\alpha), \eta^{(i)}(\alpha))$ satisfies the equations $(A_J + \alpha(\Delta A)_J)^T v + \eta e = -\alpha(\Delta A)_J^T w^{(i)}$. By construction, the matrix $[M(\alpha)]^T$ which forms the left hand side coefficients is invertible. This means that

$$\begin{pmatrix} v^{(i)}(\alpha) \\ \eta^{(i)}(\alpha) \end{pmatrix} = -\alpha[M(\alpha)]^{-T}(\Delta A)_J^T w^{(i)}.$$

The last component is $\eta^{(i)}(\alpha) = -\alpha \left\{ \left[w^{(i)}\right]^T (\Delta A)_J \, [M(\alpha)]^{-1} \, e_i \right\}$ as desired. $\square$

If $i \in J_+(0)$ and there exists $B$ such that $J(\alpha) \subseteq B \cup \{i\}$ and $B \in \mathcal{B}(A)$ (that is $j_B = i$ in the construction above) then the inverse $[A(\alpha)]_B^{-1}$ exists for every $\alpha \in [0, \hat{\alpha}]$ and the inverse of $M(\alpha)$ can be expressed as the constant

$$\frac{1}{1 - e^T [A(\alpha)]_B^{-1} [A(\alpha)]_i}$$

multiple of the matrix

$$\begin{pmatrix} \left\{1 - e^T [A(\alpha)]_B^{-1}[A(\alpha)]_i\right\} [A(\alpha)]_B^{-1} + [A(\alpha)]_B^{-1}[A(\alpha)]_i e [A(\alpha)]_B^{-1} & -[A(\alpha)]_B^{-1} A_i \\ -e^T [A(\alpha)]_B^{-1} & 1 \end{pmatrix}.$$

We have $\lim_{\alpha \to 0} [A(\alpha)]_B^{-1} = A_B^{-1}$, and $\lim_{\alpha \to 0} [M(\alpha)]^{-1} = M^{-1}$. These facts help us specialize the lemma above to $i \in J_+(0)$.

**Corollary 5.1** *Consider the conditions of the lemma above, and suppose that $i \in J_+(0)$. Then*

$$x_i(\alpha) = x_i - \frac{\alpha}{1 - e^T [A(\alpha)]_B^{-1} [A(\alpha)]_i} \left[w^{(i)}\right]^T \left\{ (\Delta A)_i - (\Delta A)_B [A(\alpha)]_B^{-1} [A(\alpha)]_i \right\}.$$

**Proof.** Since $i \in J_+(0)$ and $J_+(0) \subseteq J(\alpha)$, we have that $x_i$ and $x_i(\alpha)$ are both positive. Using Lemma 5.2 and the explicit formula for the inverse of $M(\alpha)$ given above, we obtain the desired result. $\square$

We conclude this subsection by giving a sufficient condition for $C1$ to hold. Suppose that the perturbation matrix $\Delta A$ satisfies the following condition: For every minimal $x \in \mathcal{N}(A)$, and every $j \in (J_+(x) \cup J_-(x))$, the column $(\Delta A)_j \in \mathcal{R}(A_{J_+(x) \cup J_-(x)})$. Since the columns of $(\Delta A)_J$ are in $\mathcal{R}(A_J)$, for each $j \in J$ there exist $\gamma^{(j)} \in \mathbb{R}^J$ such that $\Delta A_j = A_J \gamma^{(j)}$. Therefore, we can write $A_J + \alpha(\Delta A)_J = A_J(I + \alpha\Gamma)$; where the columns of $\Gamma$ are the vectors $\gamma^{(j)}$ previously defined. Note that for $\alpha$ small enough, the matrix $(I + \alpha\Gamma)$ is invertible. So, there exists $\bar{\alpha} > 0$ such that for each $\alpha \in (0, \bar{\alpha})$, the system (8) has a unique solution, say $x_J(\alpha)$; and $(I + \alpha\Gamma) x_J(\alpha)$ is collinear with $\bar{x}_J$.

**Proposition 5.2** *Consider the perturbation matrix $\Delta A \in \mathbb{R}^{m \times n}$ and suppose that for each minimal point $x \in \mathcal{N}(A)$ we have*

$$j \in J_+(x) \cup J_-(x) \text{ implies } (\Delta A)_j \in \mathcal{R}(A_{J_+(x) \cup J_-(x)}).$$

*Then there exists $\bar{\alpha} > 0$ such that for all $\alpha \in (0, \bar{\alpha})$, $\underline{\text{sign}}(\mathcal{N}(A + \alpha\Delta A)) = \underline{\text{sign}}(\mathcal{N}(A))$.*

**Proof.** Consider arbitrary $x \in \mathcal{N}(A)$. Denote $J = J_+(x) \cup J_-(x)$. Note that for arbitrary $T \in \mathbb{R}^{|J|\times|J|}$ and small $\alpha > 0$ the equations

$$A_J(I+\alpha T)w = 0, \|w\|_1 = 1 \tag{11}$$

have a unique solution if and only if the equations (3) do, because the matrix $(I+\alpha T)$ is invertible for every $\alpha$ sufficiently small. Since for $T = \Gamma$ the systems (8) and (11) are equivalent, we have that (8) has a unique solution if and only if (3) has a unique solution. Therefore, $x \in \mathcal{N}(A)$ is minimal if and only if the point $x(\alpha)$, the solution associated with (8), is minimal. It is easy to see that $(I+\alpha\Gamma)x_J(\alpha) = \eta x_J$ for some positive $\eta$, and from this claim we have that $\text{sign}(x) = \text{sign}(x(\alpha))$ for every small $\alpha$. We have shown that for each $x \in \mathcal{N}(A)$ minimal, there exists $x(\alpha) \in \mathcal{N}(A(\alpha))$ minimal such that $\text{sign}(x) = \text{sign}(x(\alpha))$ for all $\alpha$ sufficiently small. Therefore, $\underline{\text{sign}}(\mathcal{N}(A)) \subseteq \underline{\text{sign}}(\mathcal{N}(A(\alpha)))$.

Suppose now that there exist a minimal $\hat{x}(\alpha) \in \mathcal{N}(A+\alpha(\Delta A))$ such that for every sufficiently small $\alpha$, $\text{sign}(\hat{x}(\alpha)) \notin \underline{\text{sign}}(\mathcal{N}(A))$. Then the system of equations (3) for $J = J_+(\hat{x}(\alpha)) \cup J_-(\hat{x}(\alpha))$ must have multiple solutions. Choose a minimal $\bar{x}_J$ in $\mathcal{N}(A_J)$. Complete with zeros to define $\bar{x} \in \mathbb{R}^n$. Clearly $\bar{x}$ is a minimal solution for $\mathcal{N}(A)$. For $J = J_+(\bar{x}) \cup J_-(\bar{x})$, (8) has a unique solution if and only if (11) has a unique solution. Therefore, by the first part of the proof, the unique solution of (11) has the same sign vector as $\bar{x}$. We arrived at a contradiction. Thus, $\underline{\text{sign}}(\mathcal{N}(A)) \supseteq \underline{\text{sign}}(\mathcal{N}(A(\alpha)))$. $\square$

The condition in the hypothesis implies that condition $C1$ holds. The converse is not true, as shown in the next example. Consider the matrices $A := \begin{pmatrix} 0 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$ and $\Delta A := \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, and the minimal vector $x := (\frac{1}{2}, \frac{1}{2}, 0)^T \in \mathcal{N}(A(\alpha))$, for every $\alpha \in \mathbb{R}$. It is clear that $1 \in J_+(x) \cup J_-(x)$ and $(\Delta A)_1 \notin \mathcal{R}(A_{J_+(x)\cup J_-(x)})$, but for every $\alpha \in \mathbb{R}$, $\underline{\text{sign}}(\mathcal{N}(A)) = \underline{\text{sign}}(\mathcal{N}(A(\alpha)))$; that is, $\Delta A$ satisfies condition $C1$ for $A$.

## 5.2 Sensitivity of $\sigma(A)$

As mentioned in the previous section, we shall study the sensitivity of $\sigma(A)$ for a perturbation matrix satisfying $C4$. For each $i \in B$ we define

$$\sigma_{P_i}(A) := \max\left\{x_i : A_B x_B = 0, x_B \geq 0, e^T x_B = 1\right\}.$$

By the duality theorem of linear programming we have

$$\sigma_{P_i}(A) = \min\left\{t : A_B^T w + te \geq e_i\right\} = \min_{s\in\mathcal{R}(A_B^T)}\left\{\max\left\{\min_{j\in B\setminus\{i\}}\{s_j\}, 1+s_i\right\}\right\},$$

also see Theorem 3.1 of [37] and Lara-Gonzaga [21] for another characterization. Associated to each $i \in B$ there exists a minimal $x^{(i)}$ in $\mathcal{N}(A)$, an optimal solution of the primal problem defining $\sigma_{P_i}(A)$ ($\sigma_{P_i}(A) = x_i^{(i)}$).

**Lemma 5.3** *Consider the condition number $\sigma_P(A)$, $\sigma_{P_i}(A)$, and the perturbation matrix $\Delta A \in \mathbb{R}^{m\times n}$ satisfying $C4$ for $A$. Then there exists $\hat{\alpha} > 0$ such that for all $i \in B$ and $\alpha \in [0, \hat{\alpha}]$*

$$\sigma_{P_i}(A + \alpha\Delta A) = \sigma_{P_i}(A) - \frac{\alpha}{1 - e^T[A(\alpha)]_{\bar{B}}^{-1}[A(\alpha)]_i} \left[w^{(i)}\right]^T \left\{(\Delta A)_i - (\Delta A)_{\bar{B}}[A(\alpha)]_{\bar{B}}^{-1}[A(\alpha)]_i\right\},$$

*where $\bar{B} \in \mathcal{B}(A)$, $\bar{B} \cup \{i\} \supseteq J_+(\bar{x})$ for some $\bar{x}$ minimal in $\mathcal{N}(A)$, $\bar{x} \geq 0$, $\bar{x}_i = \sigma_{P_i}(A)$ and $(w^{(i)}, t^{(i)})$ is the unique solution of system $A_J^T w + et = e_i$ with $J = \bar{B} \cup \{i\}$.*

**Proof.** By condition $C4$, the strict complementarity partition $[B, N]$ does not change after the perturbation. Consider minimal $x^{(i)}$ such that $\text{sign}(x^{(i)}) \in \underline{\text{sign}}_+(\mathcal{N}(A))$ and $\sigma_{P_i}(A) = x_i^{(i)}$. Consider also $x^{(i)}(\alpha)$ satisfying the conditions given by Lemma 5.1. Since $i \in J_+(x^{(i)})$, by Corollary 5.1 we have

$$x_i^{(i)}(\alpha) = x_i^{(i)} - \frac{\alpha}{1 - e^T[A(\alpha)]_{\bar{B}}^{-1}[A(\alpha)]_i} \left[w^{(i)}\right]^T \left\{(\Delta A)_i - (\Delta A)_{\bar{B}}\,[A(\alpha)]_{\bar{B}}^{-1}\,[A(\alpha)]_i\right\}.$$

Since $x_i^{(i)}$ is the maximum objective value of the LP problem maximizing $x_i$ over all feasible $x$ in the original problem $(FP)$, and $x^{(i)}(\alpha) \to x^{(i)}$ as $\alpha \to 0$, we have that there exists $\hat{\alpha} > 0$ such that $x_i^{(i)}(\alpha)$ is the maximum value of the LP problem of maximizing $x_i$ over all feasible $x$ in the perturbed problem defined by $[A + \alpha(\Delta A)]$. □

**Theorem 5.3** *Consider the conditions of the lemma above. Suppose that there is a unique index $\ell$ in $B$ such that $\sigma_{P_\ell}(A) = \sigma_P(A)$. Then we have the following facts.*

*(a) There exists $\bar{\alpha} > 0$ such that for all $\alpha \in [0, \bar{\alpha}]$*

$$\sigma_P(A(\alpha)) = \sigma_{P_\ell}(A(\alpha)).$$

*(b) The directional derivative $\sigma_P'(A, \Delta A)$ exists and it is calculated as*

$$\sigma_P'(A, \Delta A) = -\frac{1}{1 - e^T A_{\bar{B}}^{-1} A_\ell} \left[w^{(\ell)}\right]^T \left\{(\Delta A)_\ell - (\Delta A)_{\bar{B}}\,[A(\alpha)]_{\bar{B}}^{-1}\,[A(\alpha)]_\ell\right\},$$

*where $\bar{B}$ and $w^{(\ell)}$ are as described in Lemma 5.3.*

**Proof.** (a) Since $\ell$ is the unique index such that $\sigma_{P_\ell}(A) = \sigma_P(A)$, by continuity arguments, there exists $\bar{\alpha} > 0$ such that $\sigma_{P_\ell}(A(\alpha)) = \sigma_P(A(\alpha))$ for every $\alpha \in [0, \bar{\alpha}]$.
(b) This follows from part (a) and Lemma 5.3, by driving $\alpha$ to zero. □

### 5.3 Sensitivity of $\xi(A)$

Now we study $\xi_P(A)$. For each $i \in B$ we define

$$\xi_{P_i}(A) := \min\{x_i \neq 0 \ : \ x \in \mathcal{N}(A), \|x\|_1 = 1, \operatorname{sign}(x) \in \underline{\operatorname{sign}}_+(\mathcal{N}(A))\}.$$

$\xi_{P_i}(A) = 1$ if for the set over which the min is computed is empty. Clearly, $\xi_P(A) = \min_{i \in B}\{\xi_{P_i}(A)\}$. For each $i$ such that $\xi_{P_i}(A) < 1$ we have a minimal $x^{(i)}$ in $\mathcal{N}(A)$ such that $\xi_{P_i}(A) = x_i^{(i)}$. Let us define $\xi_{P_i}(A(\alpha))$ accordingly, for the perturbed matrix.

**Lemma 5.4** *Consider the condition numbers $\xi_P(A)$, $\xi_{P_i}(A)$, and the perturbation matrix $\Delta A$ satisfying $C3$ for $A$. Then there exists positive $\hat{\alpha}$ such that for all $\alpha \in [0, \hat{\alpha}]$ we have*

$$\xi_{P_i}(A(\alpha)) = \xi_{P_i}(A) - \frac{\alpha}{1 - e^T [A(\alpha)]_{\bar{B}}^{-1} [A(\alpha)]_i} \left[w^{(i)}\right]^T \left\{(\Delta A)_i - (\Delta A)_{\bar{B}} [A(\alpha)]_{\bar{B}}^{-1} [A(\alpha)]_i\right\},$$

*where $\bar{B} \in \mathcal{B}(A)$, $\bar{B} \cup \{i\} \supseteq J_+(\bar{x})$ for some $\bar{x}$ minimal in $\mathcal{N}(A)$, $\bar{x} \geq 0$ and $\bar{x}_i = \xi_{P_i}(A)$ and $\left(w^{(i)}, t^{(i)}\right)$ is the unique solution of the system $A_J^T w + et = e_i$ with $J = \bar{B} \cup \{i\}$.*

**Proof.** Note that condition $C3$ implies condition $C4$ by Theorem 5.2. The rest of the proof can be completed as in the proof of Lemma 5.3. □

**Theorem 5.4** *Consider the condition numbers $\xi_P(A)$, $\xi_{P_i}(A)$, and the perturbation matrix $\Delta A$ satisfying $C3$ for $A$. Suppose that there is a unique index $k \in B$ such that $\xi_P(A) = \xi_{P_k}(A)$. Then we have the following facts.*

*(a) There exists $\tilde{\alpha} > 0$ such that for all $\alpha \in [0, \tilde{\alpha}]$ we have $\xi_P(A(\alpha)) = \xi_{P_i}(A(\alpha))$.*

*(b) The directional derivative $\xi_P'(A, \Delta A)$ exists and it is calculated as*

$$\xi_P'(A, \Delta A) = -\frac{1}{1 - e^T A_{\bar{B}}^{-1} A_k} \left[w^{(k)}\right]^T \left\{(\Delta A)_k - (\Delta A)_{\bar{B}} [A(\alpha)]_{\bar{B}}^{-1} [A(\alpha)]_k\right\},$$

*where $\bar{B}$ and $w$ are as described in Lemma 5.3.*

**Proof.** Similar to the proof of Theorem 5.3. □

Note that the above set up and the results can also be used to work with the directional derivatives of $\bar{\chi}(\cdot)$ under condition $C1$. in particular, we can utilize the above theorem together with Theorem 4.2 to get the directional derivatives of $\bar{\chi}_\infty(\cdot)$.

# 6 Generalizations to Convex Cones and Relationships with Renegar's Condition Measure

Consider a convex cone $K$ in a finite dimensional normed space $E$. The dual cone $K^*$ is defined as

$$K^* := \{s \in E^* \; : \; \langle s, x\rangle \geq 0, \text{ for every } x \in K\}.$$

We will assume that $K$ is a closed, convex cone, has a nonempty interior and is pointed (i.e., contains no line). It follows that $K^*$ has all of these properties too. We also have the fact that

$$\text{int}(K) = \{x \in E : \langle s, x\rangle > 0, \; \forall s \in K^* \backslash \{0\}\}. \tag{12}$$

*The width of* $K$ is given by:

$$\tau_K := \max_{x \in K, r \in \mathbb{R}_+} \left\{ \frac{r}{\|x\|} : B(x,r) \subset K \right\}.$$

Note that $\tau_K \in (0, 1]$. Since $K$ is pointed and has nonempty interior, $\tau_K$ is attained for some $(\bar{x}, \bar{r})$ as well as along the ray $(\alpha\bar{x}, \alpha\bar{r})$ for all $\alpha > 0$. By choosing the value of $\alpha$ appropriately, we can find $u \in K$ such that

$$\|u\| = 1 \text{ and } \tau_K \text{ is attained for } (x, r) = (u, \tau_K). \tag{13}$$

*The norm approximation coefficient of* $K$ is defined by

$$\delta_K := \text{dist}\,(0, \partial\text{conv}\,[K(1) \cup -K(1)])\,, \tag{14}$$

where $K(1) := \{x \in K \; : \; \|x\| \leq 1\}$, and $\partial\text{conv}\,[K(1) \cup -K(1)]$ is the boundary of the convex hull of the set $K(1) \cup (-K(1))$.

Epelman and Freund [8] showed that $\delta_K \geq \frac{\tau_K}{1+\tau_K} \geq \frac{\tau_K}{2}$. Also see Freund-Vera [13] for an earlier approach to similar geometric measures. We demonstrate here that when all the norms in the above definition are taken as the Euclidean 2-norm, the denominator in the previous relation is not necessary.

**Proposition 6.1** *Let $K$ be as above and suppose all the norms in the definitions of $\delta_K$ and $\tau_K$ are the Euclidean 2-norm. Then*

$$\tau_K \leq \delta_K.$$

**Proof.** Let $\bar{x}$ be an arbitrary point in $B(0, \tau_K)$. To show our claim, it suffices to prove that $\bar{x} \in \text{conv}\,[K(1) \cup -K(1)]$. If $\bar{x} \in K$ we are done. Assume $\bar{x} \notin K$. Take $u$ as in (13), and consider the two-dimensional subspace $S$ generated by $u$ and $\bar{x}$. $S$ is isomorphic to $\mathbb{R}^2$ because $\bar{x}$ and $u$ are not collinear, and each $y \in S$ can be written as $y = \alpha u + \beta \bar{x}$ where $(\alpha, \beta) \in \mathbb{R}^2$. In the following, the action happens in $S$: Choose $v \in S$ such that $u^T v = 0$, $\|v\|_2 = 1$ and $\bar{x}^T v > 0$. Define $\gamma(t) := \bar{x} + tu$ for $t > 0$. It is clear that $\gamma(1) \in B(u, \tau_K) \subset K$. Since $\gamma(0) = \bar{x} \notin K$ and $\gamma(1) \in K$ there exist a unique $\bar{t} \in (0,1)$ such that $\gamma(\bar{t}) \in \partial K$. Denote it by $\bar{\gamma}$. Define by $\theta$ the angle between $u$ and $\bar{\gamma}$. We claim that $\sin\theta \geq \tau_K$. Suppose for a contradiction that $\sin\theta < \tau_K$. Then there exists $\gamma(\hat{t}) := \hat{\gamma}$ near $\bar{\gamma}$ such that $\sin\hat{\theta} < \tau_K$ and $\hat{\gamma} \notin K$ (here $\hat{\theta}$ stands for the angle between $u$ and $\hat{\gamma}$). Now, we choose $\beta > 0$ such that $\beta\hat{\gamma}$ is the orthogonal projection of $u$ onto the ray $\alpha\hat{\gamma}$. It is clear that $\|\beta\hat{\gamma} - u\|_2 = \sin\hat{\theta} < \tau_K$, so $\beta\hat{\gamma} \in K$ which is a contradiction. So necessarily $\sin\theta \geq \tau_K$. Now, we have

$$\tau_K \leq \sin\theta = \cos(\pi/2 - \theta) = \frac{v^T\bar{\gamma}}{\|\bar{\gamma}\|_2} = \frac{v^T(\bar{x} + \bar{t}u)}{\|\bar{\gamma}\|_2} = \frac{v^T\bar{x}}{\|\bar{\gamma}\|_2}$$

So,

$$\|\bar{\gamma}\|_2 \leq \frac{v^T\bar{x}}{\tau_K} \leq 1.$$

We have shown that $\bar{\gamma} \in K(1)$. We can take in a similar way $\tilde{\gamma} = \bar{x} - \tilde{t}u$ such that $\hat{\gamma} \in (-K(1))$. Then we can write $\bar{x} = \frac{\tilde{t}}{2(\bar{t}+\tilde{t})}\bar{\gamma} + \left[1 - \frac{\tilde{t}}{2(\bar{t}+\tilde{t})}\right]\tilde{\gamma}$. □

The following example shows that the relation above is not necessarily an equality: Consider the cone $K = \mathbb{R}^n_+$. For this cone, $\tau_K = \frac{1}{\sqrt{n}}$ and $u = (\frac{1}{\sqrt{n}})e$. We claim that $\delta_K \geq \frac{1}{\sqrt{2}}$. To prove this, it is enough to show that every point $x$ in $B(0, \frac{1}{\sqrt{2}})$ also belongs to $\operatorname{conv}[K(1) \cup -K(1)]$. In fact, consider arbitrary $\bar{x} \in B\left(0, \frac{1}{\sqrt{2}}\right)$. If $\bar{x} \in K(1)$ or $\bar{x} \in -K(1)$ then $\bar{x} \in \operatorname{conv}[K(1) \cup -K(1)]$ as desired. Suppose that $\bar{x} \notin [K(1) \cup -K(1)]$. So the signing partition is proper; that is, there exist positive and negative components in $\bar{x}$. Denote by $B$ the index set of the nonnegative components, and by $N$ the index set of the negative components. Denote by $x^{(1)}$ the vector in $\mathbb{R}^n$ such that $x^{(1)}_B = \frac{1}{\|\bar{x}_B\|_2}\bar{x}_B$ and $x^{(1)}_N = 0$ and by $x^{(2)} \in \mathbb{R}^n$ such that $x^{(2)}_B = 0$ and $x^{(2)}_N = \frac{1}{(1-\|\bar{x}_B\|_2)}\bar{x}_N$. The vector $\bar{x}$ can be written as $\bar{x} = \alpha x^{(1)} + (1-\alpha)x^{(2)}$, where $\alpha = \|\bar{x}_B\|_2 \in (0, \frac{1}{\sqrt{2}}) \subset (0,1)$. It remains to show that $x^{(1)} \in K(1)$ and $x^{(2)} \in -K(1)$. The first claim is trivial. To prove the second one, note that $x^{(2)} \in -K$ and its norm is the square root of:

$$\frac{\|\bar{x}_N\|^2}{(1-\alpha)^2} \leq \frac{1}{(1-\alpha)^2}\left(\frac{1}{2} - \alpha^2\right) = \frac{1 - 2\alpha^2}{2(1-\alpha)^2} \leq 1.$$

The first inequality is true because $\bar{x} \in B(0, 1/\sqrt{2})$. The maximum of the last function in the interval $(0, \frac{1}{\sqrt{2}})$ is attained at $\alpha = \frac{1}{2}$, with value 1. So we have the result.

**Proposition 6.2** *For $K = \mathbb{R}^n_+$, if we choose all the norms in the definitions as the Euclidean 2-norm, then $\delta_K = \frac{1}{\sqrt{2}}$.*

**Proof.** We already proved that $\delta_K \geq 1/\sqrt{2}$. So, it suffices to prove the reverse inequality. Clearly, we can assume $n \geq 2$. Let $k$ be an integer such that $1 \leq k \leq n/2$. Consider the nonnegative vector $x^{(+)}$ with $1/\sqrt{k}$ in the first $k$ positions and zeros everywhere else. Similarly, let $x^{(-)}$ denote the vector with $-1/\sqrt{k}$ in the last $k$ positions and zeros everywhere else. Clearly, $x^{(+)} \in B(0,1) \cap \mathbb{R}^n_+$ and $x^{(-)} \in B(0,1) \cap -\mathbb{R}^n_+$; moreover, $\left\|\frac{1}{2}x^{(-)} + \frac{1}{2}x^{(+)}\right\|_2 = \frac{1}{\sqrt{2}}$. It remains to show that this midpoint is on the boundary of $\operatorname{conv}[K(1) \cup -K(1)]$. To see the latter, let $\epsilon > 0$ be arbitrarily small, add $\epsilon$ to the positive entries of the midpoint and subtract $\epsilon$ from the negative entries of the midpoint. The resulting vector cannot be written as $\left[\lambda x^{(-)} + (1-\lambda)x^{(+)}\right]$ for $x^{(-)} \in B(0,1) \cap -\mathbb{R}^n_+$, $x^{(+)} \in B(0,1) \cap \mathbb{R}^n_+$ and $\lambda \in [0,1]$. Therefore, the midpoint described above is on the boundary of $\operatorname{conv}[K(1) \cup -K(1)]$ as desired. $\square$

We can generalize the above proposition to all self-dual cones.

**Theorem 6.1** *Let $K$ be a convex cone in $\mathbb{R}^n$ as described at the beginning of the section. If $K = K^*$ under the inner product $\langle \cdot, \cdot \rangle$, then*

$$\delta_K = \delta_{K^*} = \frac{1}{\sqrt{2}},$$

*where the distances (and the balls) are defined with respect to the norm $\|x\| := \langle x, x\rangle^{1/2}$.*

**Proof.** The claim that $\delta_K = \delta_{K^*}$ is obvious.

First, we show that there exists $\gamma \in \partial\operatorname{conv}[K(1) \cup -K(1)]$ such that $\|\gamma\| = \frac{1}{\sqrt{2}}$. Let $v \in \operatorname{ext}(K)$. Denote by $F(v)$ the minimal face of $K$ containing $v$. Consider $[F(v)]^\Delta$, the dual face ($F^\Delta := \{y \in K^* : \langle x, y\rangle = 0,\ \forall x \in F\}$). Let $w \in \operatorname{ext}(K) \cap [F(v)]^\Delta$. Note that $\langle v, w\rangle = 0$. We claim that $\gamma := \frac{1}{2}v + \frac{1}{2}(-w) \in \partial\operatorname{conv}[K(1) \cup -K(1)]$. It is clear that $\gamma \in \operatorname{conv}[K(1) \cup -K(1)]$ and that $\|\gamma\| = \frac{1}{\sqrt{2}}$. Define $\gamma(\epsilon) := (1+\epsilon)\gamma$. We will prove that $\gamma(\epsilon) \notin \operatorname{conv}[K(1) \cup -K(1)]$. To do so, let us denote by $S$ the two-dimensional subspace generated by $v$ and $w$. Since $\gamma(\epsilon) \in S$ it suffices to show that $\gamma(\epsilon) \notin S \cap \operatorname{conv}[K(1) \cup -K(1)]$. This last set can be characterized in terms of $v$ and $w$:

$$S \cap \operatorname{conv}[K(1) \cup -K(1)] = [\operatorname{cone}\{v,w\}](1) \cup [\operatorname{cone}\{-v,-w\}](1) \cup \operatorname{conv}\{0, v, -w\} \cup \operatorname{conv}\{0, -v, w\}.$$

Clearly, for no $\epsilon > 0$, $\gamma(\epsilon)$ is in the above set. Therefore, $\gamma$ is on the boundary of $\operatorname{conv}[K(1) \cup -K(1)]$ as claimed. We proved that $\delta_K \leq \frac{1}{\sqrt{2}}$.

Secondly, to demonstrate the reverse inequality, we prove that

$$B\left(0, \frac{1}{\sqrt{2}}\right) \subseteq \operatorname{conv}[K(1) \cup -K(1)].$$

Take an arbitrary $\bar{x} \in B\left(0, \frac{1}{\sqrt{2}}\right)$. If $\bar{x} \in K(1)$ or $\bar{x} \in -K(1)$, then we are done. So, we can assume that $\bar{x} \notin [K(1) \cup -K(1)]$. Let $u$ be the closest point to $\bar{x}$ in $K$ and define $v := \bar{x} - u$.

We claim that $v$ is in $-K$. Since $u$ is the closest point to $\bar{x}$ in $K$, we have (by the Kolmogorov Criterion)

$$\langle \bar{x} - u, x - u \rangle \leq 0, \quad \forall x \in K$$

which is equivalent to

$$\langle \bar{x} - u, x \rangle \quad \leq \quad \langle u, \bar{x} \rangle - \|u\|^2, \quad \forall x \in K. \tag{15}$$

Since $K$ is a cone, and the right-hand-side is a constant, the above implies

$$\langle \bar{x} - u, x \rangle \leq 0, \quad \forall x \in K.$$

That is, $v \in -K$.

Also note that

$$\langle u, v \rangle = \langle u, \bar{x} - u \rangle = \langle u, \bar{x} \rangle - \|u\|^2 \geq 0.$$

The last inequality above follows from (15) and the fact that $0 \in K$. But $u \in K$ and $v \in -K^*$. Thus, $\langle u, v \rangle = 0$. (In fact, at this point, we can easily verify the Kolmogorov Criterion for $v$ to be the closest point to $\bar{x}$ in $-K$.) Since $\|\bar{x}\| \leq \frac{1}{\sqrt{2}}$, $\bar{x} = u + v$, and $\langle u, v \rangle = 0$, we conclude that $\|u\| \in (0, 1)$. Define

$$x^{(1)} := \frac{u}{\|u\|} \text{ and } x^{(2)} := \frac{v}{1 - \|u\|}.$$

Then $\bar{x} = \alpha x^{(1)} + (1 - \alpha) x^{(2)}$, where $\alpha := \|u\| \in (0, 1)$. Clearly, $x^{(1)} \in K(1)$. We claim that $x^{(2)} \in -K(1)$. $x^{(2)} \in -K$ was already established. Moreover,

$$\|x^{(2)}\|^2 = \frac{\|v\|^2}{(1-\alpha)^2} \leq \frac{1}{(1-\alpha)^2}\left(\frac{1}{2} - \alpha^2\right) = \frac{1 - 2\alpha^2}{2(1-\alpha)^2} \leq 1.$$

The first inequality is true because $x \in B(0, \frac{1}{\sqrt{2}})$ and $\|v\|^2 = \|\bar{x}\|^2 - \|u^2\|$. The maximum of the last function in $(0, 1/\sqrt{2})$ is attained at $\alpha = 1/2$, with value 1. $\square$

A quantity related to $\tau_K$ was mentioned by Sturm in the context of symmetric cones [34]. Let $\kappa(K)$ denote the Carathéodory number of $K$ (for a definition see [17]). Sturm remarks that the radius of the smallest circular cone (scaled second-order cone) which contains a *symmetric cone* $K$ (see [9] for a definition) is $[\kappa(K) - 1]$ times the radius of the largest inscribed circular cone. In particular, Sturm's remark about the largest inscribed circular cone is equivalent to the following fact (we omit the proof).

**Theorem 6.2** *Let $K$ be a symmetric cone and define the width of $K$ by using the norm induced by the inner product under which $K = K^*$. Then*

$$\tau_K = \frac{1}{\sqrt{\kappa(K)}}.$$

Cheung and Cucker [5] propose a condition measure for the linear feasibility problem. They establish the theory in terms of the cone of solutions $\text{sol}(A^T) := \{y \in \mathbb{R}^n \ : \ A^T y < 0\}$ and its dual cone $\text{sol}(A^T)^* = \{w \in \mathbb{R}^m \ : \ \langle w, y\rangle \geq 0, \ \forall y \in \text{sol}(A^T)\}$. We shall define this condition measure for a (general) pointed convex cone $K$. For arbitrary vectors $x, y$ in $E$, define the angle $\theta(y, x)$ between these vectors as:

$$\theta(y, x) := \arccos \frac{\langle y, x\rangle}{\|y\|\|x\|}.$$

Also define the angle $\theta(K, x)$ between $x$ and a pointed convex cone $K$ as

$$\theta(K, x) := \min_{u \in \text{ext}(K)} \{\theta(u, x)\},$$

where $\text{ext}(K)$ denotes the set of normalized extreme rays of $K$. We denote by $\bar{x} \in K$ any vector satisfying

$$\theta(-K^*, \bar{x}) = \max_{x \in E} \{\theta(-K^*, x)\}.$$

We are now ready define the generalization of Cheung and Cucker's measure:

$$\mathcal{C}^{CC}(K) := \frac{1}{|\cos \theta(-K^*, \bar{x})|}.$$

In the sequence, we establish some properties of $\mathcal{C}^{CC}$. The next lemma generalizes Lemma 1 of [5].

**Lemma 6.1** *Let $x \in E$, $\bar{x}$ defined as above, and $y \in (-K^*)$. Then*

*(a)* $\langle y, x\rangle < 0 \Leftrightarrow \cos\theta(y, x) < 0 \Leftrightarrow \theta(y, x) > \frac{\pi}{2}$.

*(b)* $x \in \text{int}(K) \Leftrightarrow \theta(-K^*, x) > \frac{\pi}{2} \Leftrightarrow \cos\theta(-K^*, x) < 0$.

*(c)* $\text{int}(K) \neq \emptyset \Leftrightarrow \bar{x} \in \text{int}(K)$.

**Proof.**

(a) This follows directly from the definition of $\theta(y, x)$.

(b) Note that

$$\begin{aligned}
\theta(-K^*, x) > \tfrac{\pi}{2} \quad &\Leftrightarrow \theta(y, x) > \tfrac{\pi}{2}, \ \forall y \in \text{ext}(-K^*) \\
&\Leftrightarrow \langle y, x\rangle < 0, \ \forall y \in \text{ext}(-K^*) && \text{(by part (a))} \\
&\Leftrightarrow \langle y, x\rangle < 0, \ \forall y \in -K^* && \text{(by convexity of } K^*) \\
&\Leftrightarrow x \in \text{int}(K) && \text{(by the equation (12)).}
\end{aligned}$$

Furthermore, since $\theta(y, x) \in [0, \pi]$ for all $y \in \text{ext}(-K^*)$, we have $\theta(-K^*, x) \in [0, \pi]$. Therefore $\theta(-K^*, x) > \frac{\pi}{2} \Leftrightarrow \cos\theta(-K^*, x) < 0$.

(c) Let $\mathrm{int}(K) \neq \emptyset$. Then there exists $\tilde{x} \in \mathrm{int}(K)$. By part (b), $\theta(-K^*, \tilde{x}) > \frac{\pi}{2}$. By the definition of $\bar{x}$, $\theta(-K^*, \bar{x}) \geq \theta(-K^*, \tilde{x}) > \frac{\pi}{2}$. By part (b) again $\bar{x} \in \mathrm{int}(K)$. The converse is obvious.

□

Suppose that $\mathrm{int}(K) \neq \emptyset$. For every $x, y \in E$, let $\beta(y,x)$ be the acute angle, i.e. $0 < \beta(y,x) \leq \frac{\pi}{2}$ between $x$ and the subspace $y^\perp$; that is,

$$\beta(y,x) := \arccos\left(\frac{\langle x, P_{y^\perp} x\rangle}{\|x\|\|P_{y^\perp} x\|}\right).$$

Also define

$$\beta(K,x) := \min_{y \in \mathrm{ext}(K)} \beta(y,x).$$

We shall prove that

$$\text{if } \mathrm{int}(K) \neq \emptyset \text{ then } \max\nolimits_{x \in K} \sin\beta(-K^*, x) = |\cos\theta(-K^*, \bar{x})|,$$

that is

$$\mathcal{C}^{CC}(K) = \min_{x \in K} \frac{1}{\sin\beta(-K^*, x)}.$$

The next lemma generalizes Lemma 7 of Cheung and Cucker [5].

**Lemma 6.2** *Consider a pointed closed convex cone $K \subset E$, and $\bar{x}$ as defined above. Suppose that* $\mathrm{int}(K) \neq \emptyset$. *Then*

*(a)* $\theta(-K^*, \bar{x}) = \max_{x \in K} \theta(-K^*, x)$.

*(b) For all $x \in \mathrm{int}(K)$ and for all $y \in \mathrm{ext}(-K^*)$, $\theta(y,x) = \beta(y,x) + \frac{\pi}{2}$.*

*(c)* $\beta(-K^*, \bar{x}) = \max_{x \in K} \beta(-K^*, x)$.

*(d)* $|\cos\theta(-K^*, \bar{x})| = \sin\beta(-K^*, \bar{x})$.

**Proof.**

(a) This follows from part (c) of the previous lemma.

(b) Let $x \in \mathrm{int}(K)$ and $y \in \mathrm{ext}(-K^*)$. Then $\langle x, y\rangle < 0$. So $y$ and $x$ are not in the same half-space with respect to $y^\perp$. It is clear that $P_{y^\perp} x$ is in the same two-dimensional subspace generated by $x$ and $y$, so we can think about $\theta$ and $\beta$ as the angles between some vectors in $\mathbb{R}^2$. From this, it follows that $\theta(y,x) = \beta(y,x) + \frac{\pi}{2}$.

(c) Since $\theta(-K^*, x) = \min_{y \in \mathrm{ext}(-K^*)} \theta(y, x)$ we have for all $x \in K$, $\theta(-K^*, x) = \beta(-K^*, x) + \frac{\pi}{2}$. This together with part (a) imply that $\beta(-K^*, \bar{x}) = \max_{x \in K} \beta(-K^*, x)$.

(d) This follows from parts (a), (b) and (c).

□

The next proposition generalizes Proposition 3 of [5].

**Proposition 6.3** *For every pointed closed convex cone $K$ such that* $\mathrm{int}(K) \neq \emptyset$,*we have*

$$\max_{x \in K} \sin \beta(-K^*, x) = |\cos \theta(-K^*, \bar{x})|.$$

**Proof.** Utilizing part (c) of the previous lemma we have

$$\begin{aligned} \max_{x \in K} \sin \beta(-K^*, x) &= \sin \max_{x \in K} \beta(-K^*, x) \\ &= \sin \beta(-K^*, \bar{x}) \\ &= |\cos \theta(-K^*, \bar{x})|. \end{aligned}$$

□

The definitions and properties described above coincide with those given by Cheung and Cucker [5] in the case of $K = \{w \in \mathbb{R}^n \ : \ A^T w \leq 0\}$. In such a special case, $-K^*$ is the cone generated by the columns of $A$. The following result establishes a link between our generalization of this condition measure and the width of $K$.

**Theorem 6.3** *For every pointed, closed convex cone $K$ with nonempty interior, we have*

$$\mathcal{C}^{CC}(K) = \frac{1}{\tau_K}.$$

**Proof.** First note that $\frac{1}{\mathcal{C}^{CC}(K)} = \sin \beta(-K^*, \bar{x})$. It suffices to show that $\tau_K = \sin \beta(-K^*, \bar{x})$: For arbitrary $x \in \mathrm{int}(K)$

$$\begin{aligned} \sin \beta(-K^*, x) &= \min_{y \in \mathrm{ext}(-K^*)} \sin \beta(y, x) \\ &= \min_{y \in \mathrm{ext}(-K^*)} \sin \left( \arccos \frac{\langle x, P_{y^\perp} x \rangle}{\|x\| \|P_{y^\perp} x\|} \right) \\ &= \min_{y \in \mathrm{ext}(-K^*)} \frac{\|x - P_{y^\perp} x\|}{\|x\|}. \end{aligned}$$

Since $P_{y^\perp}x \in y^\perp$ and $y \in \text{ext}(-K^*)$, we have that $\theta(-K^*, P_{y^\perp}x) = \frac{\pi}{2}$ and by Lemma 6.1 part (b), we conclude that $P_{y^\perp}x \notin \text{int}(K)$. This means that $\sin\beta(-K^*, x) \geq \max_{r\in\mathbb{R}_+}\left\{\frac{r}{\|x\|} : B(x,r) \subset K\right\}$. Taking the maximum in each side over all $x \in K$, we obtain that $\sin\beta(-K^*, \bar{x}) \geq \tau_K$.

Now, we take $w$ on the boundary of $K$ such that $\tau_K \geq \frac{\|\bar{x}-w\|}{\|\bar{x}\|} =: \tau_w$ and $B(\bar{x}, \tau_w) \subset K$. We denote by $F$ the minimal face of $K$ containing $w$. Consider the dual face $F^\Delta$ in $K^*$ (that is, $F^\Delta := \{y \in K^* : \langle x, y\rangle = 0,\ \forall x \in F\}$). Each element $y \in F^\Delta$ defines a supporting hyperplane $y^\perp$ of $K$ in $w$ such that $F \subset y^\perp$. Since each of these hyperplanes should also be a supporting hyperplane of the ball $B(\bar{x}, \tau_w)$ at $w$, we conclude that $F^\Delta$ is a singleton (so, $y \in \text{ext}(-K^*)$), because a full dimensional ball admits only one supporting hyperplane at a boundary point. Therefore, $y$ satisfies $y \in \text{ext}(-K^*)$, $w \perp y$ and $w = P_{y^\perp}x$. Then we have

$$\begin{aligned}\tau_K &\geq \frac{\|\bar{x}-w\|}{\|\bar{x}\|} = \frac{\|\bar{x}-P_{y^\perp}\bar{x}\|}{\|\bar{x}\|}\\ &= \sin\arccos\frac{\langle\bar{x}, P_{y^\perp}\bar{x}\rangle}{\|\bar{x}\|\|P_{y^\perp}\bar{x}\|}\\ &= \sin\beta(y, \bar{x})\\ &\geq \sin\beta(-K^*, \bar{x})\end{aligned}$$

as desired. □

The above fact was independently observed by Hauser-Cucker-Cheung [18], in the case of polyhedral convex cones.

### 6.1 Strong infeasibility certificates

In this paper, our focus is on the infeasibility certificates and the related complexity and condition measures. So, in this subsection, we focus on the convex feasibility problems described as a convex cone intersected with a linear subspace. Therefore, for a given linear subspace $S$ of $\mathbb{R}^n$ and a given pointed closed convex cone $K$ in $\mathbb{R}^n$ with nonempty interior, we define the following pair of problems:

$$(CFP)\ \left\{\begin{array}{rcl} x & \in & S,\\ \|x\| & = & 1,\\ x & \in & K,\end{array}\right.$$

$$(CFD)\ \left\{\begin{array}{rcl} s & \in & S^\perp,\\ \|s\| & = & 1,\\ s & \in & K^*,\end{array}\right.$$

where we allow different norms if so desired. We must note that we no longer have a good analogue of the uniquely determined partition $[B, N]$. Of course, even the strong duality type statements require additional assumptions in this general setting. With these warnings in place, we must also mention that the infeasibility certificates still are interesting since they characterize *strong infeasibility.* Moreover, we can study the complexity measures for any closed, convex pointed cone by simply redefining the space so that in the smaller space the cone has a nonempty interior.

It follows from the separating hyperplane theorem that

$$S \cap K = \{0\} \iff S^{\perp} \cap \mathrm{int}(K^*) \neq \emptyset.$$

The implication $S^{\perp} \cap \mathrm{int}(K^*) \neq \emptyset \Rightarrow S \cap K = \{0\}$ is trivial. For the converse, suppose $S^{\perp} \cap \mathrm{int}(K^*) = \emptyset$. Then there exists $\bar{x} \in \mathbb{R}^n$ such that $\langle \bar{x}, s\rangle \leq 0$, for every $s \in S^{\perp}$ and $\langle \bar{x}, s\rangle > 0$, for every $s \in \mathrm{int}(K^*)$. Thus, $\langle \bar{x}, s\rangle = 0$ for every $s \in S^{\perp}$ (since $S^{\perp}$ is a subspace) and $\bar{x} \in K \backslash \{0\}$. Therefore, $S \cap K \supset \{0\}$ as desired.

From now on, we will talk about the complexity and condition measures of subspace-cone pairs: $(S, K)$, $(S^{\perp}, K^*)$ etc. In the previous sections we only indicated the subspace $S$ in the notation since the underlying cone was always $\mathbb{R}^n_+$ and under the Euclidean inner product, we have $(\mathbb{R}^n_+)^* = \mathbb{R}^n_+$.

Let us think of $S$ as $\mathcal{N}(A)$. Then

$$\mathrm{sym}(A, K) := \mathrm{sym}\left\{Ax : \|x\| \leq 1,\ x \in K\right\}.$$

Renegar's condition number for $(CFP)$ can then be defined as the reciprocal of the relative distance to infeasibility, that is, the reciprocal of the smallest $\frac{\|\epsilon\|}{\|A\|}$ such that $\epsilon \in \mathbb{R}^{m \times n}$ and $\mathcal{N}(A+\epsilon) \cap K = \emptyset$. We denote Renegar's condition number for $(CFP)$ by $\mathcal{C}_P^R(A, K)$.

**Theorem 6.4** *(Epelman and Freund [8]) Suppose $\mathcal{N}(A) \cap K \neq \{0\}$. Then there exists an $m \times m$ positive definite matrix $U$ such that*

$$\mathrm{sym}(A, K) \leq \frac{1}{\mathcal{C}_P^R(UA, K)} \leq \frac{\sqrt{m}}{\delta_K}\mathrm{sym}(A, K).$$

Suppose for a moment that we use the Euclidean 2-norm in all the definitions. Then using the above theorem, Proposition 6.1 and Theorem 6.3 we have the following fact.

**Corollary 6.1** *Suppose $\mathcal{N}(A) \cap K \neq \{0\}$. Then, there exists an $m \times m$ positive definite matrix $U$ such that*

$$\mathrm{sym}(A, K) \leq \frac{1}{\mathcal{C}_P^R(UA, K)} \leq \sqrt{m}\, \mathcal{C}^{CC}(K)\, \mathrm{sym}(A, K).$$

Note that $(S \cap K)$ is always a pointed closed convex cone. When it is at least one-dimensional (note that $\{0\}$ is always in the intersection—which is not interesting for us), we can identify the linear span $\bar{E}$ of the cone $(S \cap K)$. Then, $(S \cap K)$ restricted to $\bar{E}$ is a pointed closed convex cone with nonempty interior. Therefore, our proposed generalization of Cheung-Cucker measure is applicable to it. We denote by $(S \cap K)|_{\bar{E}}$ the pointed closed convex cone in $\bar{E}$ (note that this is geometrically the same object, we simply redefined our space). This cone has a nonempty interior in $\bar{E}$. Therefore, we define

$$\mathcal{C}_P^{CC}(S, K) := \mathcal{C}^{CC}\left((S \cap K)|_{\bar{E}}\right),$$

and

$$\mathcal{C}_D^{CC}(S, K) := \mathcal{C}_P^{CC}(S^\perp, K^*),$$

where we define the underlying value to be 1 if the intersection of $S$ and $K$ is $\{0\}$, similarly for the dual measure; and finally

$$\mathcal{C}^{CC}(S, K) := \max\left\{\mathcal{C}_P^{CC}(S, K), \mathcal{C}_D^{CC}(S, K)\right\}.$$

Note that our geometric generalization of the Cheung-Cucker measure to cones is only geared towards pointed, closed convex cones with non-empty interior. For our purposes, when $(CFP)$ or $(CFD)$ is infeasible, we assign the corresponding infeasibility measure problem, the complexity value 1. (Recall the definitions describing $\sigma(A)$ for $(FP)$ and $(FD)$.) In contrast, Cheung-Cucker measure for systems of linear inequalities (even though less data dependent than the Renegar condition measure) still measures the distance from a scaled version of the data to ill-posed instances (and hence it is well-defined and very meaningful for infeasible systems too).

Our generalization is much less data dependent and more geometric; hence fitting the approach that we have taken in this paper. Indeed, in this paper we focused on problems $(FP)$, $(FD)$, $(CFP)$ and $(CFD)$ whose solutions are potential infeasibility certificates for an actual optimization problem with additional data.

**Acknowledgment** We thank Robert M. Freund for very useful comments on an earlier draft of this paper.

# Appendix

**Proof of Theorem 5.2:** Consider the linear programming problem whose optimal solution determines $\sigma(A)$:

$$\begin{array}{ll} \text{maximize} & u_i \\ \text{subject to} & u \in S \\ & \|u\|_1 = 1 \\ & u \geq 0, \end{array}$$

where $S$ is either $\mathcal{N}(A)$ or $\mathcal{R}(A^T)$ (and $i \in B$ or $i \in N$). The optimal value of the above linear programming problem is $\sigma(A)$. So, there exists $\bar{u} \in S$ which is an extreme point of the feasible region such that $\bar{u}_i = \sigma(A)$. But $\bar{u}$ is a minimal element of $S$. Therefore, $\xi(A) \leq \sigma(A)$. This also establishes $\underline{\xi}(A) \leq \underline{\sigma}(A)$.

Next, we prove $\underline{\sigma}(A) \leq \xi(A)$. We can assume that $\xi(A) = \xi_P(A)$. Let $\bar{x}$ and $J$ determine $\xi(A)$ so that $\xi(A)$ is the optimal value of

$$\begin{array}{ll} \text{maximize} & x_i \\ \text{subject to} & A_J x_J = 0 \\ & e^T x_J = 1 \\ & x_J \geq 0 \end{array}$$

determined by the unique optimal solution $\bar{x}_J$. Let $(\bar{w}, \bar{t}) \in \mathbb{R}^m \oplus \mathbb{R}$ be an optimal solution of the dual of this LP problem. Then, $\bar{t} = \bar{x}_i > 0$. For every $j \notin J$ such that $\bar{w}^T A_j < 0$, multiply $A_j$ by $(-1)$. Now, for this signing of $A$, consider $\bar{x}$ (extended by zeros to $\mathbb{R}^n$) and $(\bar{w}, \bar{t})$. Then $\bar{x}$ is a feasible solution of

$$\begin{array}{ll} \text{maximize} & x_i \\ \text{subject to} & (A_{-G})\, x = 0 \\ & e^T x = 1 \\ & x \geq 0 \end{array}$$

and $(\bar{w}, \bar{t})$ is a feasible solution of its dual. Moreover, $\bar{x}_i = \bar{t}$. Therefore, $\bar{x}$ is an optimal solution of the last LP problem and we conclude

$$\underline{\sigma}(A) \leq \sigma(A_{-G}) \leq \bar{x}_i = \xi(A).$$

Thus, $\underline{\sigma}(A) \leq \xi(A)$. The last inequality we proved is valid for any signing of $A$ and the left-hand-side is invariant under signings of $A$. Therefore, we have $\underline{\sigma}(A) \leq \underline{\xi}(A)$. Reverse inequality was already established; thus, we have equality as desired. The rest of the relations follow from Theorem 2 of [36] and the fact that $\underline{\sigma}(A) = \underline{\xi}(A)$. □

## References

[1] I. Adler and R. D. C. Monteiro, A geometric view of parametric linear programming, *Algorithmica* 8 (1992) 161–176.

[2] D. Azé and J.-N. Corvellec, On the sensitivity analysis of Hoffman constants for systems of linear inequalities, *SIAM J. Opt.* 12 (2002) 913–927.

[3] A. Björner, M. Las Vergnas, B. Sturmfels, N. White and G. Ziegler, *Oriented matroids*, Cambridge University Press, 1993.

[4] R. G. Bland and M. Las Vergnas, Orientability of matroids, *J. Combinatorial Theory B* 23 (1978) 94–123.

[5] D. Cheung and F. Cucker, A new condition number for linear programming, *Math. Prog. A* 91 (2001) 163–174.

[6] L. Danzer, B. Grünbaum and V. Klee, Helly's theorem and its relatives, 1963 Proc. Sympos. Pure Math., Vol. VII, Amer. Math. Soc., Providence, RI, USA, pp. 101–180.

[7] S. Deng, Perturbation analysis of a condition number for convex inequality systems and global error bounds for anayltic systems, *Math. Prog. A* 83 (1998) 263–276.

[8] M. Epelman and R. M. Freund, A new condition measure, pre-conditioners, and relations between different measures of conditioning for conic linear systems, *SIAM J. Opt.* 12 (2002) 627–655.

[9] J. Faraut and A. Korányi, *Analysis on Symmetric Cones*, Oxford University Press, NY, USA, 1994.

[10] A. Forsgren, On linear least-squares problems with diagonally dominant weight matrices, *SIAM J. Matrix Anal. Appl.* 17 (1996) 763–788.

[11] R. M. Freund, Complexity of convex optimization using geometry-based measures and a reference point, Sept. 2001.

[12] R. M. Freund, On the primal-dual geometry of level sets in linear and conic optimization, July 2001.

[13] R. M. Freund and J. R. Vera, Some characterizations and properties of the "distance to ill-posedness" and the condition measure of a conic linear system, *Math. Prog. A* 86 (1999) 225-260.

[14] C. C. Gonzaga and H. Lara, A note on properties of condition numbers, *Linear Algebra And Its Applications* 261 (1997) 269–273.

[15] H. J. Greenberg, Matrix sensitivity analysis from an interior solution of a linear program, *INFORMS J. Comput.* 11 (1999) 316–327.

[16] H. J. Greenberg, Simultaneous primal-dual right-hand-side sensitivity analysis from a strictly complementary solution of a linear program, *SIAM J. Optim.* 10 (2000) 427–442.

[17] O. Güler and L. Tunçel, Characterization of the barrier parameter of homogeneous convex cones, *Math. Prog. A* 81 (1998) 55–76.

[18] R. A. Hauser, F. Cucker and D. Cheung, Distribution tails of condition numbers for the polyhedral conic feasibility problem, *Talk presented at the SIAM Conf. Optim.*, May 2002, Toronto, Canada.

[19] R. Hirabayashi, H. Th. Jongen and M. Shida, Stability for linearly constrained optimization problems, *Math. Prog. A* 66 (1994) 351–360.

[20] J. C. K. Ho and L. Tunçel, Reconciliation of various complexity and condition measures for linear programming problems and a generalization of Tardos' theorem, In: *Foundations of Computational Mathematics, Proceedings of the Smalefest 2000*, F. Cucker and J. M. Rojas (Eds.) World Scientific, 2002, pp. 93–147.

[21] H. Lara and C. C. Gonzaga, A characterization of the smallest large variable condition number, *Investigación Operativa*, Volume 9, N 1,2 and 3 (2000).

[22] Z.-Q. Luo and P. Tseng, Perturbation analysis of a condition number for linear systems, *SIAM J. Matrix Anal. Appl.* 5 (1994) 636–660.

[23] N. Megiddo and M. Shub, Boundary behavior of interior point algorithms in linear programming, *Math. Oper. Res.* 14 (1989) 97–146.

[24] R. D. C. Monteiro and S. Mehrotra, A general parametric analysis approach and its implication to sensitivity analysis in interior point methods, *Math. Prog. A* 72 (1996) 65–82.

[25] D. P. O'Leary, On bounds for scaled projections and pseudoinverses, *Linear Algebra And Its Applications* 132 (1990) 115–117.

[26] J. Renegar, Linear programming, complexity theory and elementary functional analysis, *Math. Prog. A* 70 (1995) 279–351.

[27] J. Renegar, Some perturbation theory for linear programming, *Math. Prog. A* 65 (1994) 73–91.

[28] S. M. Robinson, Local structure of feasible sets in nonlinear programming. III. Stability and sensitivity, *Math. Prog. Stud.* 30 (1987) 45–66.

[29] S. M. Robinson, Local structure of feasible sets in nonlinear programming. II. Nondegeneracy, *Math. Prog. Stud.* 22 (1984) 217–230.

[30] S. M. Robinson, A characterization of stability in linear programming, *Oper. Res.* 25 (1977) 435–447.

[31] S. M. Robinson, Stability theory for systems of inequalities, Part II: Differentiable nonlinear systems, *SIAM J. Numer. Anal.* 13 (1976) 497–513.

[32] S. M. Robinson, Stability theory for systems of inequalities, Part I: Linear systems, *SIAM J. Numer. Anal.* 12 (1975) 754–769.

[33] G. W. Stewart, On scaled projections and pseudoinverses, *Linear Algebra And Its Applications* 112 (1989) 189–193.

[34] J. F. Sturm, Similarity and other spectral relations for symmetric cones, *Linear Algebra And Its Applications* 312 (2000) 135–154.

[35] M. J. Todd, A Dantzig-Wolfe-like variant of Karmarkar's interior-point linear programming algorithm, *Oper. Res.* 38 (1990) 1006–1018.

[36] M. J. Todd, L. Tunçel and Y. Ye, Probabilistic analysis of two complexity measures for linear programming problems, *Math. Prog. A* 90 (2001) 59–69.

[37] L. Tunçel, On the condition numbers for polyhedra in Karmarkar's form, *Operations Research Letters* 24 (1999) 149–155.

[38] S. A. Vavasis, Stable numerical algorithms for equilibrium systems, *SIAM J. Matrix Anal. Appl.* 15 (1994) 1108–1131.

[39] S. A. Vavasis and Y. Ye, A primal-dual interior point method whose running time depends only on the constraint matrix, *Math. Prog. A* 74 (1996) 79–120.

[40] S. A. Vavasis and Y. Ye, Condition numbers for polyhedra with real number data, *Operations Research Letters* 17 (1995) 209–214.

[41] Y. Ye, Toward probabilistic analysis of interior-point algorithms for linear programming, *Mathematics of Operations Research* 19 (1994) 38–52.

[42] Y. Ye, On the finite convergence of interior-point algorithms for linear programming, *Math. Prog. B* 57 (1992) 325-335.

[43] E. A. Yıldırım and M. J. Todd, An interior-point approach to sensitivity analysis in degenerate linear programs, *SIAM J. Optim.* 12 (2002) 692–714.