Reconciliation of Various Complexity and Condition Measures for Linear Programming Problems and a Generalization of Tardos' Theorem*

Jackie C. K. Ho[†] Levent Tunçel[‡]

CORR 2000–33
June 12, 2000 (revised: April 2001)

Abstract

First, we review and clarify the relationships amongst various complexity and condition measures for linear programming problems. Then, we generalize Tardos' Theorem for linear programming problems with integer data to linear programming problems with real number data. Our generalization, in contrast to the only previous such generalization due to Vavasis and Ye, shows that many conventional, polynomial-time (in the sense of the Turing Machine Model, with integer data) primal-dual interior-point algorithms can be adapted in a Tardos' like scheme, to solve linear programming problems with real number data in time polynomial in the dimensions of the coefficient matrix and the logarithms of certain measures of the coefficient matrix (independent of the objective function and the right-hand-side vectors).

Keywords: linear programming, computational complexity, complexity measures, interior-point methods

AMS Subject Classification: 90C05, 90C60, 49K40, 90C31, 90C51

*Part of this work was done while both authors were visiting The Fields Institute, Toronto, Ontario, Canada (Fall, 1999).

[†]Ph.D. student, 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 an OGSST, an NSERC PGSB, and by an NSERC research grant and a PREA of the second author.

[‡]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 a Faculty Research Grant from University of Waterloo, research grants from NSERC, and a PREA from Ontario, Canada.

1 Introduction

Let $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$. In this paper, one of our main concerns is the computational complexity of solving linear programming (LP) problems with data (A, b, c) in a way that the number of arithmetic operations is bounded by polynomial functions determined only by A.

For $t \in \mathbb{R}_+$, poly(t) denotes a polynomial function of t. For $\alpha \in \mathbb{Z}$, we define

$$\text{size}(\alpha) := \lceil \log(|\alpha| + 1) \rceil + 1;$$

for $A \in \mathbb{Z}^{m\times n}$,

$$\text{size}(A) := \sum_{i,j} \text{size}(a_{ij}).$$

When $A \in \mathbb{Z}^{m\times n}$, $b \in \mathbb{Z}^m$, $c \in \mathbb{Z}^n$, Tardos [22] proved that the existence of an algorithm for LP which performs only polynomially many elementary arithmetic operations in size(A, b, c) implies the existence of an algorithm for LP which performs only poly(size(A)) elementary arithmetic operations. (Her results also apply in the more general case $A \in \mathbb{Q}^{m\times n}$, $b \in \mathbb{Q}^m$, $c \in \mathbb{Q}^n$, also see [23] for network flow problems.)

Tardos' proof is constructive in the sense that it shows how to use *any* polynomial time algorithm for LP as a subroutine to achieve the goal of solving LP problems in poly(size(A)) time complexity. However, the proof requires calling the subroutine (the LP solver with poly(size(A, b, c)) time complexity), polynomially many times using modified data so that the sizes of the modified LP instances can be bounded by poly(size(A)).

Later Vavasis and Ye [29], in another seminal paper (with many new insights), proposed a new kind of interior-point algorithm and proved that their algorithm can solve LP problems with data $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, in $O\left(n^{3.5}\left(\log \bar{\chi}(A) + \log(n)\right) \log\log \bar{\chi}(A)\right)$ interior-point iterations. Also, see Adler and Beling's [1] paper which is more specialized than the Vavasis-Ye paper since it is concerned with the polynomial-time LP algorithms over the algebraic numbers. When specialized to integer (or rational) data, Vavasis-Ye result gives another proof of Tardos' theorem (using $\bar{\chi}(A) = 2^{O(\text{size}(A))}$—see Section 2). So, in this sense, Vavasis-Ye result generalizes Tardos' theorem to LP problems with data $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$. Vavasis-Ye proof is even "more constructive" in the sense that their algorithm is a specialized algorithm designed for such a purpose, and need not be called many times (except to guess an upper bound for $\bar{\chi}(A)$—accounted for in the above quoted iteration bound by the $\log\log \bar{\chi}(A)$ term; also see [15]).

One advantage of Vavasis-Ye algorithm is that it has the potential of becoming a practical algorithm. However, theoretically speaking, Vavasis and Ye left open the question of whether conventional polynomial time interior-point algorithms (or perhaps some others) can be adapted in a scheme more directly related to Tardos' to solve the LP problems with data $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ in polynomially many elementary arithmetic operations where the polynomial bound depends only on the (properly defined) "size" of $A \in \mathbb{R}^{m\times n}$. In fact, Vavasis and Ye [29] state that

"Tardos uses the assumption of integer data in a fairly central way: an important tool in [22] is the operation of rounding down to the nearest integer. It is not clear how to generalize the rounding operation to noninteger data."

For example, let $A \in \mathbb{Z}^{m\times n}$, $c \in \mathbb{Z}^n$. Then if d is an extreme ray of $\{x \in \mathbb{R}^n : Ax = 0, x \geq 0\}$ such that $c^T d < 0$, then we know that there exists an integral extreme ray d in the above cone such that

$c^T d \leq (-1)$. Of course, such arguments do not directly apply in general when the entries of A and c are real numbers. When A and c have only rational entries, the data can be multiplied by a large enough (but not too large) integer such that the new scaled data contain only integers. This again ensures a notion of a "unit" to round to, even after a normalization of the integral d such that $\sum_{j=1}^{n} d_j = 1$, so that the arguments similar to the above still work (e.g., after such a normalization, $c^T d \leq -1/\Delta(A)$, where $\Delta(A)$ denotes the largest absolute value of a subdeterminant of A). In addition to this, a few other obstacles arise in an attempt to obtain such a generalization of Tardos' theorem and proof to the real number model.

In this paper, we overcome these obstacles, and generalize Tardos' theorem *and* a significant part of her proof to the case when $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$. Our results also generalize Vavasis and Ye's result in the sense that in our scheme almost any polynomial time LP algorithm can be adapted, whereas their result uses a new, specialized algorithm.

Before we describe the generalization of Tardos' theorem, we review and clarify (with many new results) relationships amongst various complexity and condition measures such as $\chi(A)$, $\bar{\chi}(A)$, the condition number of (AA^T) denoted by $\kappa(AA^T)$, Hoffman's bound (or the Lipschitz bound) for systems of linear inequalities, Ye's complexity measure for LP (also known as the smallest large variable bound), $\Delta(A)$ and the smallest nonzero absolute value of a subdeterminant of A, denoted $\delta(A)$. Special emphasis is put on establishing various fundamental properties of $\bar{\chi}$, which becomes one of the central tools in the last section when we deal with generalization of Tardos' result. While our proof of the generalization of her theorem is very similar to hers, a key part of the proof which makes it work in the real number case, is the generalization of the rounding operation to noninteger data (in the sense of choosing an appropriate "unit" for the data at hand). For this, we rely heavily on those fundamental properties of $\bar{\chi}$ mentioned above. We first perform our analysis on deciding the feasibility of a system of inequalities, and then use the resulting algorithm as a subroutine to solve the whole primal-dual LP problem. In both cases, we solve the original problem by solving a sequence of polynomially many "nicer" or smaller LP problems, each of which has integral right hand side vector (and cost vector, in the latter case) whose size is bounded by a polynomial function of our complexity measures. This is one of the fundamental tools for eliminating the dependence on b and c in the overall complexity bound of the algorithms. Solving these "nicer" LP problems gives us important information about the structure of the optimal solutions of the main LP problem in terms of the linear algebraic structures of the input data. For example, "there exists an optimal solution at which the jth inequality is tight" or "at all optimal solutions, the jth inequality is strictly satisfied." Such information helps us reduce the dimensions of the problem at hand; but, it also requires us to analyze the complexity measures for the subproblems.

The sizes of all the integers making up the right hand side and objective vectors of these "nicer" LP problems are bounded above by a polynomial function of n and the logarithm of $\left(\frac{\Delta(A)}{\delta(A)}\right)$. Many are also bounded by a polynomial function of n and $\log \bar{\chi}(A)$.

As mentioned, we need to use an LP solver as a subroutine in our proof of Tardos' theorem. While any polynomial time LP solver can be used, we describe a very useful formulation – the homogeneous self-dual form – in Section 5. The complexity of running an interior-point algorithm (with a certain termination rule) on such a form can be expressed in terms of Ye's complexity measure, which becomes convenient in our complexity analysis.

This paper is organized as follows. In Section 2, we review definitions and characterizations of some complexity measures which are relevant to our stated interest in this paper. We also present some new results in this section. Section 3 includes the Cauchy-Binet formula and an application of it to obtain a bound on the condition number of (AA^T). In Section 4, we discuss Hoffman's Theorem and relate the

Hoffman constant to $\chi(A)$. In Section 5, we discuss Ye's complexity measure for LP problems and relate it to the Hoffman constant. Also in Section 5, we show that the number of iterations of many primal-dual interior-point algorithms to solve LP problems with data (A, b, c), with arbitrary A and special b and c, can be bounded by a polynomial function of n and logarithms of certain complexity measures. We review a sensitivity bound result of Cook, Gerards, Schrijver, Tardos [3] in Section 6 and establish various variants of it based on the complexity measures $\chi(A)$ and $\bar{\chi}(A)$. Section 7 contains our main result – a generalization of Tardos' Theorem – based on the results obtained in the preceding sections. We conclude with a very brief discussion of the special cases when A is integral and totally unimodular.

2 Complexity and Condition Measures: χ and $\bar{\chi}$

We denote by $\mathcal{N}(A)$, the null-space of A; $\mathcal{R}(A)$ denotes the range (or column-space) of A. We assume $A \neq 0, n > m \geq 3$. Recall the definitions:

$$\begin{aligned} \|A\|_p &:= \max_{\|x\|_p=1} \|Ax\|_p, \text{ for } 1 \leq p \leq \infty, \\ \|A\|_F &:= \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |A_{ij}|^2}. \end{aligned}$$

It is not hard to show that

$$\|A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{m} |A_{ij}|, \tag{1}$$

$$\|A\|_\infty = \max_{1\leq i\leq m} \sum_{j=1}^{n} |A_{ij}|. \tag{2}$$

We have the following well-known matrix norm inequalities:

$$\|A\|_2 \leq \|A\|_F \leq \sqrt{n}\|A\|_2, \tag{3}$$

$$\frac{1}{\sqrt{n}}\|A\|_\infty \leq \|A\|_2 \leq \sqrt{m}\|A\|_\infty, \tag{4}$$

$$\frac{1}{\sqrt{m}}\|A\|_1 \leq \|A\|_2 \leq \sqrt{n}\|A\|_1, \tag{5}$$

We also have the submultiplicative property for p-norms, $1 \leq p \leq \infty$. For all $A \in \mathbb{R}^{m\times n}, C \in \mathbb{R}^{n\times q}$, we have

$$\|AC\|_p \leq \|A\|_p\|C\|_p. \tag{6}$$

For the rest of the paper, the 2-norm is assumed when norms are mentioned, unless stated otherwise.

We assume throughout this section that A has full row rank. Define

$$\bar{\chi}(A) := \sup\{\|A^T(ADA^T)^{-1}AD\| : D \in \mathcal{D}\},$$

where $\mathcal{D}$ is the set of all positive definite $n \times n$ diagonal matrices. Note that $\bar{\chi}(RA) = \bar{\chi}(A)$ for all nonsingular $R \in \mathbb{R}^{m\times m}$. In fact, $\bar{\chi}(A)$ depends only on the pair of orthogonal subspaces, $\mathcal{N}(A)$ and $\mathcal{R}(A^T)$. So, it can be defined on subspaces instead. Note that for all $D \in \mathcal{D}$,

$$\|A^T\| = \|A^T(ADA^T)^{-1}ADA^T\| \leq \|A^T(ADA^T)^{-1}AD\| \cdot \|A^T\|.$$

Hence $\|A^T(ADA^T)^{-1}AD\| \geq 1$, and thus $\bar{\chi}(A) \geq 1$.

Similarly we define

$$\chi(A) := \sup\{\|(ADA^T)^{-1}AD\| : D \in \mathcal{D}\}.$$

Note that both $\bar{\chi}(A)$ and $\chi(A)$ are finite. Also,

$$\begin{aligned} \|A^T(ADA^T)^{-1}AD\| &\leq \|A^T\| \cdot \|(ADA^T)^{-1}AD\|, \\ \|(ADA^T)^{-1}AD\| &\leq \|(AA^T)^{-1}A\| \cdot \|A^T(ADA^T)^{-1}AD\|. \end{aligned}$$

Therefore, we have

$$\frac{1}{\|A\|}\bar{\chi}(A) \leq \chi(A) \leq \|(AA^T)^{-1}A\|\bar{\chi}(A) = \frac{\sqrt{\kappa(AA^T)}\bar{\chi}(A)}{\|A\|}, \tag{7}$$

where $\kappa(R) := \|R\| \cdot \|R^{-1}\|$, the condition number of R, for any nonsingular matrix R. Note that if $m = n$, then $\kappa(AA^T) = \|A\|^2 \cdot \|A^{-1}\|^2 = (\|A\|\chi(A))^2$.

An equivalent way to define these parameters is in terms of weighted least squares:

$$\begin{aligned} \bar{\chi}(A) &= \sup\left\{\frac{\| A^T y \|}{\| c \|} : y \text{ minimizes } \| D^{1/2}(A^T y - c) \| \text{ for some } c \in \mathbb{R}^n, D \in \mathcal{D}\right\}, \\ \chi(A) &= \sup\left\{\frac{\| y \|}{\| c \|} : y \text{ minimizes } \| D^{1/2}(A^T y - c) \| \text{ for some } c \in \mathbb{R}^n, D \in \mathcal{D}\right\}. \end{aligned}$$

Let us define, for $1 \leq \alpha, \beta \leq \infty$,

$$\rho_{\alpha,\beta}(A) := \inf\{\|x - y\|_\beta : x \in X, y \in Y_\alpha\},$$

where $X := \{D\xi : \xi \in \mathcal{N}(A), D \in \text{cl}(\mathcal{D})\}, Y_\alpha := \{\gamma : \gamma \in \mathcal{R}(A^T), \|\gamma\|_\alpha = 1\}$, and $\text{cl}(\mathcal{D})$ denotes the closure of the set $\mathcal{D}$, that is, the set of nonnegative diagonal matrices. Note that $\rho_{\alpha,\beta}(\cdot) > 0$. If we have $\|\cdot\|_\beta \leq c\|\cdot\|_\alpha$, then $\rho_{\alpha,\beta}(\cdot) \leq c$, as $0 \in X$. In particular, $\rho_{\alpha,\alpha}(\cdot) \leq 1$. Also note that the definition of $\rho_{\alpha,\beta}(A)$ depends only on $\mathcal{R}(A^T)$ and its orthogonal complement $\mathcal{N}(A)$. Gonzaga and Lara [8] prove that when $\alpha = \beta = 2$, the subspaces $\mathcal{N}(A)$ and $\mathcal{R}(A^T)$ can be interchanged in the definition of $\rho_{\alpha,\beta}(A)$. In the following, we denote $\rho_{\alpha,\alpha}$ simply by ρ_α. We are mostly interested in ρ_2, which we denote simply by ρ.

All vector p-norms are equivalent, that is, given α, β such that $1 \leq \alpha, \beta \leq \infty$, there exist positive c_1, c_2 such that $c_1\|\cdot\|_\alpha \leq \|\cdot\|_\beta \leq c_2\|\cdot\|_\alpha$. This property also applies to $\rho_{\alpha,\beta}$:

Proposition 2.1 *Suppose* $1 \leq \alpha, \beta, \gamma, \delta \leq \infty$, *and* $c_1, c_2, d_1, d_2 > 0$ *such that* $d_1\|\cdot\|_\alpha \leq \|\cdot\|_\delta \leq d_2\|\cdot\|_\alpha$ *and* $c_1\|\cdot\|_\gamma \leq \|\cdot\|_\beta \leq c_2\|\cdot\|_\gamma$. *Then*

$$c_1 d_1 \rho_{\delta,\gamma}(\cdot) \leq \rho_{\alpha,\beta}(\cdot) \leq c_2 d_2 \rho_{\delta,\gamma}(\cdot).$$

Proof
Note that $\rho_{\delta,\beta}(A) = \inf\{\|x - y\|_\beta : x \in X, y \in Y_\delta\}$ is attained, by say, $\bar{x}$ and $\bar{y}$. Let $y^* := \bar{y}/\|\bar{y}\|_\alpha$. Then $y^* \in Y_\alpha$, and we have

$$\rho_{\delta,\beta}(A) = \|\bar{x} - \bar{y}\|_\beta = \|\bar{y}\|_\alpha \left\| \frac{\bar{x}}{\|\bar{y}\|_\alpha} - y^* \right\|_\beta \geq \frac{\|\bar{y}\|_\delta}{d_2}\rho_{\alpha,\beta}(A) = \frac{1}{d_2}\rho_{\alpha,\beta}(A).$$

By considering the infimum in $\rho_{\alpha,\beta}(A)$, similarly we have $\rho_{\alpha,\beta}(A) \geq d_1\rho_{\delta,\beta}(A)$. Combining the above, we have $d_1\rho_{\delta,\beta}(A) \leq \rho_{\alpha,\beta}(A) \leq d_2\rho_{\delta,\beta}(A)$. By using $c_1\|\cdot\|_\gamma \leq \|\cdot\|_\beta \leq c_2\|\cdot\|_\gamma$, we have $c_1d_1\rho_{\delta,\gamma}(\cdot) \leq \rho_{\alpha,\beta}(\cdot) \leq c_2d_2\rho_{\delta,\gamma}(\cdot)$. □

In particular, we have

$$\frac{1}{\sqrt{n}}\rho(\cdot) \leq \rho_\infty(\cdot) \leq \sqrt{n}\rho(\cdot). \tag{8}$$

The following is a well-known fact.

Proposition 2.2 *(Stewart [21])*

$$\bar{\chi}(A) = 1/\rho(A).$$

A *basis* of A is a set of indices $B \subseteq \{1,\ldots,n\}$ such that $|B| = m$ and the columns of A_B are linearly independent. We denote the set of all bases of A by $\mathcal{B}(A)$.

Proposition 2.3 *(Vavasis and Ye [29], Todd, Tunçel and Ye [24])*

$$\bar{\chi}(A) = \max\{\|A_B^{-1}A\| : B \in \mathcal{B}(A)\}.$$

Here, "$\geq$" is proven in [29] and "$\leq$" is proven in [24]. It is known and not hard to show that an analogous characterization for $\chi(A)$ also exists:

$$\chi(A) = \max\{\|A_B^{-1}\| : B \in \mathcal{B}(A)\}. \tag{9}$$

Using the above proposition, we prove that $\bar{\chi}$ cannot increase if any column is removed.

Proposition 2.4 *Suppose $\tilde{A}$ is obtained by removing a column $a \in \mathbb{R}^m$ from $A \in \mathbb{R}^{m\times n}$. We have the following:*

- *If* $\operatorname{rank}(\tilde{A}) = m$, *then* $\bar{\chi}(\tilde{A}) \leq \bar{\chi}(A)$.
- *If* $\operatorname{rank}(\tilde{A}) \leq m-1$, *then let* $\bar{A}$ *be obtained by removing any dependent row from* $\tilde{A}$. *We have* $\operatorname{rank}(\bar{A}) = m-1$ *and* $\bar{\chi}(\bar{A}) = \bar{\chi}(A)$.

Proof
If $\operatorname{rank}(\tilde{A}) = m$, we have

$$\begin{aligned}\bar{\chi}(\tilde{A}) &= \|\tilde{A}_B^{-1}\tilde{A}\|, \text{ for some basis } B \text{ of } \tilde{A}\\ &\leq \|[\tilde{A}_B^{-1}\tilde{A}|\tilde{A}_B^{-1}a]\| = \|\tilde{A}_B^{-1}A\| \leq \bar{\chi}(A).\end{aligned}$$

We used the fact that B is also a basis of A. Now, consider the case where $\operatorname{rank}(\tilde{A}) \leq m-1$. Without loss of generality, assume a is the last column of A. Then by row reduction, there exists a nonsingular $G \in \mathbb{R}^{m\times m}$ such that

$$GA = G[\tilde{A}|a] = \begin{pmatrix} A' & 0 \\ 0^T & 1 \end{pmatrix},$$

for some $A' \in \mathbb{R}^{(m-1)\times(n-1)}$ having full row rank (hence, $\text{rank}(\bar{A}) = m-1$). Then

$$\mathcal{R}(\bar{A}^T) = \mathcal{R}(\tilde{A}^T) = \mathcal{R}((G\tilde{A})^T) = \mathcal{R}(A'^T),$$

and hence $\mathcal{N}(\bar{A}) = \mathcal{N}(A')$. So, $\bar{\chi}(\bar{A}) = \bar{\chi}(A')$. Now, since every basis of GA must include the last column,

$$\begin{aligned} \bar{\chi}(GA) &= \left\| \begin{pmatrix} (A'_B)^{-1} & 0 \\ 0^T & 1 \end{pmatrix} \begin{pmatrix} A' & 0 \\ 0^T & 1 \end{pmatrix} \right\|, \text{ for some basis } B \text{ of } A' \\ &= \left\| \begin{pmatrix} (A'_B)^{-1} A' & 0 \\ 0^T & 1 \end{pmatrix} \right\| = \max\{\| (A'_B)^{-1} A' \|, 1\} \le \bar{\chi}(A'). \end{aligned}$$

The proof of $\bar{\chi}(A') \le \bar{\chi}(GA)$ is similar. Therefore, we have

$$\bar{\chi}(\bar{A}) = \bar{\chi}(A') = \bar{\chi}(GA) = \bar{\chi}(A).$$

□

Consider $A \in \mathbb{Q}^{m\times n}$. Let L denote the total number of bits required to store A. We have the following.

Proposition 2.5 *(Vavasis and Ye [29])*
If $A \in \mathbb{Q}^{m\times n}$, $\bar{\chi}(A)$ and $\chi(A)$ are both bounded by $2^{O(L)}$.

Khachiyan [14] proved that approximating $\chi(A)$ within a factor of $2^{\text{poly}(n)}$ is NP-hard. Similarly, approximating $\bar{\chi}(A)$ within a factor of $2^{\text{poly}(n)}$ is also NP-hard [25].

The following observation is due to O'Leary [17]. Naturally, for $\alpha \in \mathbb{R}$, sign(α) is either $+, 0$, or $-$ depending on the sign of α.

Proposition 2.6 *(O'Leary [17])*
Considering J, γ, ξ as the variables, we have

$$\begin{aligned} \rho_{\alpha,\beta}(A) = \quad \min \quad & \|\gamma_J\|_\beta \\ \text{subject to} \quad \text{sign}(\gamma_j) &= \text{sign}(\xi_j), j \notin J \\ \|\gamma\|_\alpha &= 1, \\ \gamma &\in \mathcal{R}(A^T), \\ \xi &\in \mathcal{N}(A), \\ J &\subseteq \{1, 2, \ldots, n\}, J \neq \emptyset. \end{aligned}$$

Consider the matrix:

$$A_C := \begin{pmatrix} A & 0 \\ C & C \end{pmatrix},$$

where C is an $n \times n$ invertible matrix. Obviously A_C also has full row rank. We have the following result.

Proposition 2.7 *(Ho [11])*

$$\bar{\chi}(A_C) = \sqrt{2}\bar{\chi}(A).$$

Proof
It is easy to see that

$$\begin{aligned} \mathcal{N}(A_C) &= \left\{ \begin{pmatrix} \xi \\ -\xi \end{pmatrix} : \xi \in \mathcal{N}(A) \right\} \text{ and} \\ \mathcal{R}(A_C^T) &= \left\{ \begin{pmatrix} \gamma + y \\ y \end{pmatrix} : \gamma \in \mathcal{R}(A^T), y \in \Re^n \right\}. \end{aligned}$$

We will prove that $\rho(A) = \sqrt{2}\rho(A_C)$ using the characterization of ρ in Proposition 2.6 with $\mathcal{R}(A^T)$ and $\mathcal{N}(A)$ interchanged (which we can do since we are working with the 2-norms). Let us denote this minimization problem as $Q(A)$.

1. $\rho(A) \geq \sqrt{2}\rho(A_C)$
   Let (ξ^*, γ^*, J^*) be an optimal solution of $Q(A)$. We now define a y^* that satisfies certain sign conditions. If $j \in J^*$, let y_j^* be such that $\text{sign}(-\xi_j^*) = \text{sign}(y_j^*)$. Therefore $\text{sign}(\xi_j^*) \neq \text{sign}(\gamma_j^* + y_j^*)$. Now if $j \notin J^*$, we can let y_j^* be such that $\text{sign}(-\xi_j^*) = \text{sign}(y_j^*)$, and $\text{sign}(\xi_j^*) = \text{sign}(\gamma_j^* + y_j^*)$, by ensuring $|y_j^*|$ is small enough. Thus, the 3-tuple

   $$\left(\frac{1}{\sqrt{2}} \begin{pmatrix} \xi^* \\ -\xi^* \end{pmatrix}, \begin{pmatrix} \gamma^* + y^* \\ y^* \end{pmatrix}, J^* \right)$$

   is feasible for $Q(A_C)$. Therefore, $\sqrt{2}\rho(A_C) \leq \|\xi_{J^*}^*\| = \rho(A)$.

2. $\rho(A) \leq \sqrt{2}\rho(A_C)$
   Let

   $$\left(\begin{pmatrix} \xi^* \\ -\xi^* \end{pmatrix}, \begin{pmatrix} \gamma^* + y^* \\ y^* \end{pmatrix}, J^* \right)$$

   be an optimal solution of $Q(A_C)$. Let

   $$\hat{J} := \{ j \in \{1, \ldots, n\} : \text{sign}(\xi_j^*) \neq \text{sign}(\gamma_j^*) \}.$$

   Since ξ^* is orthogonal to γ^*, and $\xi^* \neq 0$, there must exist j such that $\text{sign}(\xi_j^*) \neq \text{sign}(\gamma_j^*)$. Hence $\hat{J} \neq \emptyset$. The 3-tuple $(\sqrt{2}\xi^*, \gamma^*, \hat{J})$ is feasible for $Q(A)$, and therefore $\rho(A) \leq \sqrt{2}\|\xi_{\hat{J}}^*\|$. Now take any $j \in \hat{J}$. Since $\text{sign}(\xi_j^*) \neq \text{sign}(\gamma_j^*)$, there does not exist a y_j which satisfies both $\text{sign}(\xi_j^*) = \text{sign}(\gamma_j^* + y_j)$ and $\text{sign}(-\xi_j^*) = \text{sign}(y_j)$ at the same time. Hence, at least one of j or $n + j$ is in J^*. Therefore we have

   $$\frac{1}{\sqrt{2}}\rho(A) \leq \|\xi_{\hat{J}}^*\| \leq \left\| \begin{pmatrix} \xi^* \\ -\xi^* \end{pmatrix}_{J^*} \right\| = \rho(A_C).$$

□

Using a proof similar to the above or using Proposition 2.3, we easily prove the following fact.

Proposition 2.8 *(Ho [11])*

$$\bar{\chi}([A| - A]) = \sqrt{2}\bar{\chi}(A).$$

Recall that the *singular values* of A are the square roots of the eigenvalues of the matrix $A^T A$. The largest singular value of A is simply $\|A\|_2$. Let $\sigma_{\min}(A)$ denote the smallest nonzero singular value of A. We have the following connection to $\rho(A)$.

Proposition 2.9 *(Stewart [21] and O'Leary [17])*
Let the columns of $U \in \mathbb{R}^{n\times m}$ form an orthonormal basis for $\mathcal{R}(A^T)$. Then

$$\rho(A) = \min_{\emptyset \neq I \subseteq \{1,\ldots,n\}} \sigma_{\min}(U_I),$$

where U_I denotes the submatrix formed from a set I of rows of U.

First, Stewart proved "$\leq$", next O'Leary proved "$\geq$". A nonzero $x \in \mathcal{N}(A)$ (with nonzero entries in positions $\{i_1, \ldots, i_p\} \subseteq \{1, \ldots, n\}$) is said to define a *minimal linear dependence* amongst the columns of A if for every subset I of size at most $(p-1)$ of $\{i_1, \ldots, i_p\}$, the columns of A indexed by I are linearly independent. We have the following proposition due to Vavasis.

Proposition 2.10 *(Vavasis [28])*
Let $x \in \mathcal{N}(A)$ be a nonzero vector defining a minimal linear dependence amongst the columns of A. We have

$$\frac{\min\{|x_j| : x_j \neq 0\}}{\max\{|x_j| : x_j \neq 0\}} \geq \rho(A).$$

We now give a new proof that is different from Vavasis'.

Proof
Let k and l be such that $\min\{|x_j| : x_j \neq 0\} = |x_k| = l$. Let us denote the jth column of A as A_j, for all $j \in \{1, \ldots, n\}$. Then there exists $J \subseteq \{1, \ldots, n\} \setminus \{k\}$ such that $A_J x_J = \pm l A_k$, where x_J contains precisely the nonzero entries of x other than x_k. Since x defines a minimal linear dependence, the columns of A_J must be linearly independent. So we can extend J to a basis B of A to get $A_B x_B = \pm l A_k$. Now,

$$\|x\|_\infty = \|x_B\|_\infty \leq \|x_B\| = l\|A_B^{-1} A_k\| \leq l\|A_B^{-1} A\| \leq l\bar{\chi}(A).$$

In other words,

$$\bar{\chi}(A) \geq \frac{\|x\|_\infty}{l},$$

or equivalently,

$$\rho(A) \leq \frac{l}{\|x\|_\infty} = \frac{\min\{|x_j| : x_j \neq 0\}}{\max\{|x_j| : x_j \neq 0\}}.$$

□

Using these arguments, it is not hard to show that the same result holds for any extreme ray x of the cone $\{x : Ax = 0, x \geq 0\}$.

Corollary 2.11 *Suppose $\{d \in \mathbb{R}^n : Ad = 0, e^T d = 1, d \geq 0\}$ is not empty. Then, it is compact and every extreme point $\bar{d}$ of it has the property*

$$\min\{\bar{d}_j : \bar{d}_j \neq 0\} \geq \frac{\rho(A)}{n}.$$

Proof
Compactness of the set is clear. Every extreme point corresponds to an extreme ray (and hence a basic feasible direction) of $\{x : Ax = 0, x \geq 0\}$. For every basic feasible direction $\bar{x}$, we identify the smallest nonzero component $\bar{x}_k$ first, and then $B \in \mathcal{B}(A)$ such that all other nonzero components of $\bar{x}$ are determined by the system of equations

$$A_B x_B = -\bar{x}_k A_k.$$

Then, as in the proof of Proposition 2.10, we get $||\bar{x}||_\infty \leq \bar{x}_k \bar{\chi}(A)$. Letting $\bar{d} := \bar{x}/(e^T \bar{x})$, we see that

$$\min\{\bar{d}_j : \bar{d}_j \neq 0\} \geq \frac{\bar{x}_k}{\bar{x}_k\,[m\bar{\chi}(A) + 1]} \geq \frac{1}{n\bar{\chi}(A)} = \frac{\rho(A)}{n}.$$

We used the facts that $n \geq (m+1)$ and $\bar{\chi}(A) \geq 1$. □

Recall that $\Delta(A)$ and $\delta(A)$ denote the maximum and minimum (respectively) of the absolute values of the determinants of all the nonsingular square submatrices of A. We have the following relationship among ρ_∞, $\Delta(A)$ and $\delta(A)$, proven via exploitation of the sign pattern characterization and Cramer's Rule.

Proposition 2.12 *(Tunçel [27])*

$$\rho_\infty(A) \geq \frac{\delta(A)}{m\Delta(A)}.$$

Proof
Recall the definition

$$\rho_\infty(A) := \inf\{||Dx - y||_\infty : D \in \text{cl}(\mathcal{D}), x \in \mathcal{N}(A), y \in \mathcal{R}(A^T), ||y||_\infty = 1\}.$$

Clearly here we can restrict x to be in $\{x \in \mathcal{N}(A) : ||x|| \leq 1\}$. Let $\{(D^k, x^k, y^k)\}$ be a sequence of feasible solutions such that $||D^k x^k - y^k||_\infty$ converges to $\rho_\infty(A)$. Since $\{x^k\}$ and $\{y^k\}$ are in compact feasible sets, we may assume $\{(x^k, y^k)\}$ converges to, say, (x^*, y^*). Let J^* be the set of indices such that the signs of x^* and y^* disagree. Note that $J^* \neq \emptyset$ because otherwise we can choose $D \in \text{cl}(\mathcal{D})$ such that $Dx^* - y^* = 0$, contradicting the fact that $\rho_\infty(A) > 0$. Note that for the pair (x^*, y^*), a best D^* is such that

$$D^*_{ii} = \begin{cases} 0, & i \in J^*, \\ 1, & i \notin J^*, x^*_i = 0, \\ \frac{y^*_i}{x^*_i}, & i \notin J^*, x^*_i \neq 0. \end{cases}$$

So $\rho_\infty(A) = ||y^*_{J^*}||_\infty$, and therefore

$$\rho_\infty(A) = \min\{||y_{J^*}||_\infty : y \in \mathcal{R}(A^T), ||y||_\infty = 1, \text{sign}(y) = \text{sign}(y^*)\}.$$

Then it is easy to see that

$$\begin{aligned}\frac{1}{\rho_\infty(A)} &= \max\{\|y\|_\infty : y \in \mathcal{R}(A^T), \mathrm{sign}(y) = \mathrm{sign}(y^*), \|y_{J^*}\|_\infty \le 1\} \\ &= \max\{\|A^T w\|_\infty : \mathrm{sign}(A^T w) = \mathrm{sign}(y^*), \|(A^T w)_{J^*}\|_\infty \le 1\}.\end{aligned}$$

Let w^* be a maximizer of this expression,

$$\begin{aligned}\epsilon &:= \min\{|(A^T w^*)_j| : (A^T w^*)_j \neq 0\}, \\ F(\mathrm{sign}(y^*), J^*) &:= \left\{ w : \begin{array}{ll} (A^T w)_j \ge \epsilon, & \text{if } \mathrm{sign}(y_j^*) = 1, \\ (A^T w)_j = 0, & \text{if } \mathrm{sign}(y_j^*) = 0, \\ (A^T w)_j \le -\epsilon, & \text{if } \mathrm{sign}(y_j^*) = -1, \\ (A^T w)_j \le 1, & \text{if } j \in J^* \end{array} \right\}.\end{aligned}$$

Then

$$\frac{1}{\rho_\infty(A)} = \max\{\|A^T w\|_\infty : w \in F(\mathrm{sign}(y^*), J^*)\} = \max\{a^T w : w \in F(\mathrm{sign}(y^*), J^*)\},$$

where a is a column of A (or its negation) such that $a^T w^* = \|A^T w\|_\infty$. Equivalently this is the optimal value of the LP:

$$\begin{array}{llcl} (P) \quad \max & & \eta & \\ & w & & \in \ F(\mathrm{sign}(y^*), J^*), \\ & a^T w & - \ \eta & \ge \ 0. \end{array}$$

Suppose the feasible region of (P) contains a line. So there exist (w, η) and $(d, t) \neq 0$ such that $w + kd \in F(\mathrm{sign}(y^*), J^*)$ and $a^T(w + kd) \ge \eta + kt$, for all $k \in \mathbb{R}$. So $A^T d = 0$. If $d \neq 0$, then it contradicts that fact that A has full row rank. So $d = 0$ and $t \neq 0$. But then $a^T w = a^T(w + kd) \ge \eta + kt$ for all $k \in \mathbb{R}$ also gives a contradiction. So the feasible region of (P) is pointed, and hence contains an optimal basic feasible solution. Let $f(\epsilon)$ be the vector representing the right-hand-side values in the definition of $F(\mathrm{sign}(y^*), J^*)$ (entries of $f(\epsilon)$ are $0,1,\epsilon,-\epsilon$). Then using Cramer's Rule, we have

$$\frac{1}{\rho_\infty(A)} = \left| \frac{\mathrm{subdet}\begin{bmatrix} \begin{pmatrix} A^T \\ A^T_{J^*} \end{pmatrix} & f(\epsilon) \\ a^T & 0 \end{bmatrix}}{\mathrm{subdet}\begin{bmatrix} \begin{pmatrix} A^T \\ A^T_{J^*} \end{pmatrix} & 0 \\ a^T & -1 \end{bmatrix}} \right| \le \frac{m\Delta(A)}{\delta(A)}.$$

Here, we used that fact that $\epsilon \le 1$; because $0 \neq \|(A^T w^*)_{J^*}\|_\infty \le 1$ (as otherwise, it would contradict $\rho_\infty(A) > 0$). □

In fact, the above was originally stated for $A \in \mathbb{Z}^{m \times n}$ in [27], in which case we have $\rho_\infty(A) \ge 1/(m\Delta(A))$.

Proposition 2.5 is a consequence of Proposition 2.12 and (8). Indeed,

$$\bar{\chi}(A) = \frac{1}{\rho(A)} \le \frac{\sqrt{n}}{\rho_\infty(A)} \le \sqrt{n} m \frac{\Delta(A)}{\delta(A)}. \tag{10}$$

Therefore, for $A \in \mathbb{Q}^{m \times n}$,

$$\log(\bar{\chi}(A)) \leq \log\left(\frac{\Delta(A)}{\delta(A)}\right) + \log(m) + \frac{1}{2}\log(n) = O(L).$$

Directly utilizing equation (9) and Proposition 2.3, we also bound χ and $\bar{\chi}$ in terms of Δ/δ in the following two propositions.

Proposition 2.13

$$\chi(A) \leq m\frac{\Delta(A)}{\delta(A)}.$$

Proof
Suppose $B \in \mathcal{B}(A)$ maximizes (9). Let y be such that $||y|| = 1$ and $||A_B^{-1}|| = ||A_B^{-1}y||$. Let $x \in \mathbb{R}^m$ such that $A_B x = y$. Then by Cramer's rule, for each $i \in \{1, \ldots, m\}$,

$$|x_i| = \frac{|\text{subdet}([A_B|y])|}{|\det(A_B)|} \leq ||y||_1 \frac{\Delta(A)}{\delta(A)} \leq \sqrt{m}\frac{\Delta(A)}{\delta(A)}.$$

So,

$$||A_B^{-1}||^2 = ||x||^2 \leq m^2\frac{\Delta(A)^2}{\delta(A)^2}.$$

Therefore,

$$\chi(A) = ||A_B^{-1}|| \leq m\frac{\Delta(A)}{\delta(A)}.$$

□

Proposition 2.14

$$\bar{\chi}(A) \leq \sqrt{m(n-m)+1}\ \frac{\Delta(A)}{\delta(A)}.$$

Proof
Suppose $B \in \mathcal{B}(A)$ maximizes the expression in Proposition 2.3. Let $\{y^1, \ldots, y^{n-m}\}$ be the columns of A that are not in A_B. Let $x^l \in \mathbb{R}^m$ such that $A_B x^l = y^l$, for all $l \in \{1, \ldots, n-m\}$. Then by Cramer's rule, for each $i \in \{1, \ldots, m\}$,

$$|x_i^l| = \frac{|\det(C)|}{|\det(A_B)|} \leq \frac{\Delta(A)}{\delta(A)},$$

for some $m \times m$ submatrix C of A. So, denoting the maximum eigenvalue of a matrix by $\lambda_{\max}(\cdot)$, we have

$$\begin{aligned} ||A_B^{-1}A||^2 &= ||[I|x^1|\cdots|x^{n-m}]||^2 = \lambda_{\max}[I + x^1(x^1)^T + \cdots + x^{n-m}(x^{n-m})^T] \\ &\leq 1 + (x^1)^T x^1 + \cdots + (x^{n-m})^T x^{n-m} \\ &\leq 1 + m(n-m)\frac{\Delta(A)^2}{\delta(A)^2} \leq [m(n-m)+1]\frac{\Delta(A)^2}{\delta(A)^2}. \end{aligned}$$

Therefore,

$$\bar{\chi}(A) = \|A_B^{-1}A\| \leq \sqrt{m(n-m)+1}\frac{\Delta(A)}{\delta(A)}.$$

□

Facts similar to those given in last three propositions can also be obtained by employing the Cauchy-Binet Formula. This goes back at least to Dikin [4]. (For a historical account and related results, see Forsgren [7] and the references therein.)

3 Cauchy-Binet Formula and the Condition Number of AA^T

Recall that $\mathcal{B}(A)$ denotes the set of all bases of A. We represent each basis B of A as a m-subset of the set of numbers from the natural numbering of the columns of A.

Proposition 3.1 *(Cauchy-Binet Formula)*
Let $A, \tilde{A} \in \mathbb{R}^{m\times n}$ with full row rank. Then

$$\det(A\tilde{A}^T) = \sum_{B\in\mathcal{B}(A)\cap\mathcal{B}(\tilde{A})} \det(A_B)\det(\tilde{A}_B).$$

Using this, we can prove the following relationship among κ, Δ and δ.

Proposition 3.2 *Suppose $A \in \mathbb{R}^{m\times n}$ has full row rank. Then*

$$\kappa(AA^T) \leq m^{3/2}n^{m+1}\frac{\Delta(A)^4}{\delta(A)^2}.$$

Proof
We have $|A_{ij}| \leq \Delta(A)$ for all i, j, and hence by (3),

$$\|AA^T\| = \|A\|^2 \leq \|A\|_F^2 \leq mn\Delta(A)^2.$$

On the other hand,

$$\|(AA^T)^{-1}\| \leq \sqrt{m}\|(AA^T)^{-1}\|_\infty \leq \frac{m^{3/2}\Delta(AA^T)}{\det(AA^T)}.$$

Now, by Proposition 3.1,

$$\begin{aligned}
\Delta(AA^T) &= \det(A_{I,*}A_{J,*}^T) \text{ (for some sets } I, J \subseteq \{1,\ldots,m\}, |I| = |J|) \\
&= \sum_{B\in\mathcal{B}(A_{I,*})\cap\mathcal{B}(A_{J,*})} \det(A_{I,*B})\det(A_{J,*B}) \\
&\leq \binom{n}{m}\Delta(A)^2 \leq \frac{n^m\Delta(A)^2}{m},
\end{aligned}$$

and $\det(AA^T) = \sum_{B\in\mathcal{B}(A)} \det(A_B)^2 \geq \delta(A)^2$. Therefore,

$$\kappa(AA^T) = \|AA^T\| \cdot \|(AA^T)^{-1}\| \leq \frac{m^{3/2}n^{m+1}\Delta(A)^4}{\delta(A)^2}.$$

□

4 Hoffman's Bound and χ

For a vector $u \in \mathbb{R}^n$, let $\mathrm{pos}(u) \in \mathbb{R}^n$ be such that $(\mathrm{pos}(u))_j := \max\{u_j, 0\}$ for each $j \in \{1, \ldots, n\}$. The following result gives an upper bound on the distance of a point to a polyhedron, in terms of its violation of the constraints defining the polyhedron.

Theorem 4.1 *(Hoffman [12])*
Let $A \in \mathbb{R}^{m\times n}$ (not necessarily full row rank) and let $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ be norms on $\mathbb{R}^m$ and on $\mathbb{R}^n$, respectively. Then there exists a scalar $K_{\alpha,\beta}(A)$, such that for every $c \in \mathbb{R}^n$ for which the set $\{y \in \mathbb{R}^m : A^Ty \leq c\} \neq \emptyset$, and for every $y' \in \mathbb{R}^m$,

$$\min_{y:A^Ty\leq c} \| y - y' \|_\alpha \leq K_{\alpha,\beta}(A) \| pos(A^Ty' - c) \|_\beta .$$

The coefficient $K_{\alpha,\beta}(A)$ is sometimes called a *Lipschitz bound* of A. For a norm $\|\cdot\|$ on $\mathbb{R}^n$, let $\|\cdot\|^*$ be the *dual norm* defined by

$$\| v \|^* := \max\{v^Tx : x \in \mathbb{R}^n, \| x \| \leq 1\},$$

for each $v \in \mathbb{R}^n$. Note that for p-norms $(1 \leq p \leq \infty)$, we have $\|\cdot\|_p^* = \|\cdot\|_q$, where q is such that $p^{-1} + q^{-1} = 1$. In particular, $\|\cdot\|_2^* = \|\cdot\|_2$. Let $\mathrm{ext}(S)$ denote the set of extreme points of a set S. We have the following geometric representation of the Lipschitz bound.

Proposition 4.2 *(Güler, Hoffman and Rothblum [9])*
Theorem 4.1 holds with $K_{\alpha,\beta}(A) := \max\{\| v \|_\beta^* : v \in \mathrm{ext}(V_\alpha(A))\}$, where $V_\alpha(A) := \{v \in \mathbb{R}^n : v \geq 0, \| Av \|_\alpha^* \leq 1\}$.

We write $K_2(A) := K_{2,2}(A)$ for all A. There is also a representation of the Lipschitz bound via singular values. For any $E \in R^{n\times m}$. Let $U(E)$ be the set of subsets of $\{1, \ldots, n\}$ for which the corresponding rows of E are linearly independent. Let $U^*(E)$ be the maximal elements in $U(E)$.

Proposition 4.3 *(Güler, Hoffman and Rothblum [9])*

$$K_2(A) \leq \max_{J\in U^*(A^T)} \frac{1}{\sigma_{\min}(A_J^T)}.$$

Note that $\min_{J\in U^*(A^T)} \sigma_{\min}(A_J^T) = \min_{\emptyset\neq J\subseteq\{1,\ldots,n\}} \sigma_{\min}(A_J^T)$. To prove this, first note that "$\geq$" is clear. Take $A\in\mathbb{R}^{m\times n}$ with rank, say, r. Take any nonempty $J\subseteq\{1,\ldots,n\}$. Let $\sigma_i(E)$ denote the ith largest singular value of any matrix E, and $k := \operatorname{rank}(A_J^T)$. Then $\sigma_{\min}(A_J^T) = \sigma_k(A_J^T)$. Let $I\subseteq J$ be such that $\operatorname{rank}(A_I^T) = k = |I|$. Then by the interlacing property of singular values, $\sigma_k(A_I^T)\leq\sigma_k(A_J^T)$. Let $M\in U^*(A^T)$ be such that $I\subseteq M$. Then

$$\sigma_{\min}(A_M^T) = \sigma_r(A_M^T) \leq \sigma_k(A_I^T) \leq \sigma_{\min}(A_J^T),$$

where we used the interlacing property again in the first inequality above. Therefore,

$$\max_{J\in U^*(A^T)} \frac{1}{\sigma_{\min}(A_J^T)} = \max_{\emptyset\neq J\subseteq\{1,\ldots,n\}} \frac{1}{\sigma_{\min}(A_J^T)}.$$

The next proposition gives a connection between K_2 and $\bar{\chi}$ via singular values.

Proposition 4.4 *Suppose $A\in\mathbb{R}^{m\times n}$ has full row rank. Then*

$$\|A\|K_2(A) \leq \bar{\chi}(A).$$

Proof
Consider the singular value decomposition of A. Let $A = UDV^T$, where $U\in\mathbb{R}^{m\times m}$ is orthogonal, $D\in\mathbb{R}^{m\times n}$ is diagonal (with singular values $\sigma_1,\ldots,\sigma_m$ of A on the diagonal, in that order), and $V\in\mathbb{R}^{n\times n}$ is orthogonal as well. Suppose $V = [v_1|\cdots|v_n]$, i.e., $\{v_1,\ldots,v_n\}$ are the columns of V. Let $\bar{V} := [v_1|\cdots|v_m]$ and $\Sigma := \operatorname{Diag}(\sigma_1,\ldots,\sigma_m)$. Then $A = U\Sigma\bar{V}^T$. Since A has full row rank, $\sigma_1,\ldots,\sigma_m > 0$, and hence Σ is invertible. We have $A^T = \bar{V}\Sigma U^T$, and $\bar{V} = A^T U\Sigma^{-1}$. So $\mathcal{R}(A^T) = \mathcal{R}(\bar{V})$, and $\bar{V}$ has orthonormal columns. By Propositions 2.9 and 4.3,

$$K_2(\bar{V}^T) \leq \max_{I\in U^*(\bar{V})} \frac{1}{\sigma_{\min}(\bar{V}_I)} = \max_{\emptyset\neq I\subseteq\{1,\ldots,n\}} \frac{1}{\sigma_{\min}(\bar{V}_I)} = \bar{\chi}(A).$$

Now it remains to show $\|A\|K_2(A)\leq K_2(\bar{V}^T)$. Note that

$$\|A^T y\|^2 = y^T U\Sigma\bar{V}^T\bar{V}\Sigma U^T y = \|\Sigma U^T y\|^2.$$

Therefore,

$$\|A\| = \|A^T\| = \max_{\|y\|=1}\|A^T y\| = \|\Sigma U^T\| = \|U\Sigma\|.$$

Now we consider the relationship between $K_2(A)$ and $K_2(\bar{V}^T)$. Suppose $K_2(A) = \|\hat{v}\|$, where $\hat{v}$ is an extreme point of $V_2(A)$. Let $\bar{v} := \|A\|\hat{v}$. We will prove that $\bar{v}$ is an extreme point of $V_2(\bar{V}^T)$. Suppose $\bar{v} = \lambda w + (1-\lambda)z$, where $\lambda\in(0,1)$, and $w,z\in V_2(\bar{V}^T)$. Then

$$\hat{v} = \lambda\frac{w}{\|A\|} + (1-\lambda)\frac{z}{\|A\|}.$$

Since $w\in V_2(\bar{V}^T)$, $w\geq 0$ and therefore $w/\|A\|\geq 0$. Also,

$$\left\|A\left(\frac{w}{\|A\|}\right)\right\| = \frac{1}{\|A\|}\|U\Sigma\Sigma^{-1}U^T Aw\| \leq \frac{1}{\|A\|}\|U\Sigma\|\|\bar{V}^T w\| \leq 1.$$

So $w/\|A\| \in V_2(A)$, and similarly so does $z/\|A\|$. Therefore, $w = z$, implying that $\bar{v}$ is an extreme point of $V_2(\bar{V}^T)$. Now,

$$\|A\|K_2(A) = \|A\|\|\hat{v}\| = \|\bar{v}\| \leq K_2(\bar{V}^T) \leq \bar{\chi}(A).$$

□

As a corollary, since $\bar{\chi}(A) \leq \|A\|\chi(A)$, we have $K_2(A) \leq \chi(A)$. During the review of our paper, we became aware of [33]. Note that the relation $K_2(A) \leq \chi(A)$ implies Theorem 3.6 from [33] which states that Theorem 4.1 holds with $K_{\alpha,\beta}(A)$ replaced by $\chi(A)$, when $\alpha = \beta = 2$ and A has full row rank. Also Lemmas 3.3, 3.4 and 3.5 of [33] follow from equation (9) and the fact that whenever $\{x : Ax = b, x \geq 0\}$ is nonempty, it contains a basic feasible solution.

We also note that, by Proposition 2.12, we have

$$K_2(A) \leq \frac{m\Delta(A)}{\|A\|\delta(A)}.$$

Let $\mathcal{G}$ be the set of diagonal matrices in $\mathbb{R}^{n\times n}$ with diagonal entries from $\{1, -1\}$. Take $G \in \mathcal{G}$. Then $\|AG\| = \|A\|$. Also for any diagonal matrix $D \in \mathbb{R}^{n\times n}, \|(AG)^T(AGD(AG)^T)^{-1}AGD\| = \|A^T(ADA^T)^{-1}AD\|$, and hence $\bar{\chi}(AG) = \bar{\chi}(A)$. (Similarly, $\chi(AG) = \chi(A)$.) Therefore, we have

$$\max_{G\in\mathcal{G}} K_2(AG) \leq \frac{\bar{\chi}(A)}{\|A\|} \leq \chi(A). \tag{11}$$

Also, $\Delta(AG) = \Delta(A)$ and $\delta(AG) = \delta(A)$. So we also have

$$\max_{G\in\mathcal{G}} K_2(AG) \leq \frac{m\Delta(A)}{\|A\|\delta(A)}.$$

We now characterize the extreme points of $V_1(A)$. Recall that

$$V_1(A) = \left\{ v \in \mathbb{R}^n : \begin{pmatrix} A \\ -A \\ -I \end{pmatrix} v \leq \begin{pmatrix} e \\ e \\ 0 \end{pmatrix} \right\},$$

which is a polyhedron, and the constraint matrix in the above description has full column rank. Let $J \subseteq \{1, \ldots, n\}$ such that $|J| \leq m$. Then we pick $I_1, I_2 \subseteq \{1, \ldots, m\}$ such that $I_1 \cap I_2 = \emptyset$ and $|I_1| + |I_2| = |J|$. Assume that the matrix

$$\begin{pmatrix} A_{I_1,J} \\ -A_{I_2,J} \end{pmatrix}$$

is nonsingular. Here $A_{I_1,J}$ denotes the submatrix of A with rows indexed by I_1 and columns indexed by J. Let $x \in \mathbb{R}^n$ be such that $x_j = 0$ if $j \notin J$ and

$$\begin{pmatrix} A_{I_1,J} \\ -A_{I_2,J} \end{pmatrix} x_J = \begin{pmatrix} e \\ e \end{pmatrix}.$$

If $x \in V_1(A)$, then x is an extreme point of $V_1(A)$. Vice versa, any given $x \in \text{ext}(V_1(A))$ must satisfy the above for some J, I_1 and I_2. So using Cramer's rule, for each $j \in J$,

$$x_j = \frac{\left|\text{subdet}\begin{pmatrix} A_{I_1,J} & e \\ -A_{I_2,J} & e \end{pmatrix}\right|}{\left|\det\begin{pmatrix} A_{I_1,J} \\ -A_{I_2,J} \end{pmatrix}\right|} \leq \frac{|J|\Delta(A)}{\delta(A)} \leq \frac{m\Delta(A)}{\delta(A)}.$$

So,

$$K_1(A) = \|x\|_\infty \leq \frac{m\Delta(A)}{\delta(A)}, \text{ and } K_{1,\infty}(A) = \|x\|_1 \leq \frac{m^2\Delta(A)}{\delta(A)}.$$

Therefore, we have

$$\max_{G\in\mathcal{G}} K_1(AG) \leq \frac{m\Delta(A)}{\delta(A)}, \text{ and } \max_{G\in\mathcal{G}} K_{1,\infty}(AG) \leq \frac{m^2\Delta(A)}{\delta(A)}.$$

In fact, the extreme points of $V_1(AG)$ can be characterized in a similar way. The only difference is that we require x to satisfy the sign pattern given by G, instead of $x \geq 0$. Now, we give another proof of the implication $K_2(AG) \leq \chi(A)$, $\forall G \in \mathcal{G}$ of (11). We use the following characterization of $K_2(A)$ for this purpose.

Lemma 4.5

$$K_2(A) = \max\left\{\|A_B^{-1}A\gamma\| : \ \gamma \in \mathcal{R}(A^T), B \in \mathcal{B}(A), \|A\gamma\| = 1, \gamma_B \leq -A_B^{-1}A_N\gamma_N\right\}.$$

Proof
Note that

$$K_2(A) = \max\left\{\|v\| : \ v \in \text{ext}(V_2(A)) \cap \left\{\xi - \gamma : \ \xi \geq \gamma, \xi \in \mathcal{N}(A), \gamma \in \mathcal{R}(A^T), \|A\gamma\| = 1\right\}\right\}.$$

Let $\gamma \in \mathcal{R}(A^T)$ such that $\|A\gamma\| = 1$; also let $B \in \mathcal{B}(A)$ such that $\gamma_B \leq -A_B^{-1}A_N\gamma_N$ and $\|A_B^{-1}A\gamma\|$ is equal to the maximum value in the statement of the lemma. Define $\xi \in \mathbb{R}^n$ as follows. $\xi_N := \gamma_N$, $\xi_B := -A_B^{-1}A_N\gamma_N$. Thus, we have $\xi \in \mathcal{N}(A)$, $\xi \geq \gamma$. Next, we claim $v := (\xi - \gamma) \in \text{ext}(V_2(A))$. Suppose not. Then, there exist $v^{(1)}, v^{(2)} \in V_2(A)$ such that $\frac{1}{2}\left(v^{(1)} + v^{(2)}\right) = \xi - \gamma$, $v^{(1)} \neq v^{(2)}$. We immediately have $v_N^{(1)} = v_N^{(2)} = 0$. Thus,

$$1 = \|A_Bv_B\| \leq \frac{1}{2}\|A_Bv_B^{(1)}\| + \frac{1}{2}\|A_Bv_B^{(2)}\| \leq 1$$

which implies

$$\|A_Bv_B\| = \|A_Bv_B^{(1)}\| = \|A_Bv_B^{(2)}\| = 1.$$

Therefore (since $v_B = \frac{1}{2}v_B^{(1)} + \frac{1}{2}v_B^{(2)}$), by the characterization of the equality case in the Cauchy-Schwarz inequality, we must have $A_Bv_B = A_Bv_B^{(1)} = A_Bv_B^{(2)}$. Since $v_B^{(1)} \neq v_B^{(2)}$, A_B must be singular, we arrived at a contradiction. In addition to $(\xi - \gamma)$ being an extreme point of $V_2(A)$, we have

$$\|\xi - \gamma\| = \|A_B^{-1}A_N\gamma_N + \gamma_B\| = \|A_B^{-1}A\gamma\|.$$

Therefore,

$$K_2(A) \geq \max\left\{\|A_B^{-1}A\gamma\|:\ \gamma \in \mathcal{R}(A^T), B \in \mathcal{B}(A), \|A\gamma\| = 1, \gamma_B \leq -A_B^{-1}A_N\gamma_N\right\},$$

as desired.

To prove the reversed inequality, we let $\xi \in \mathcal{N}(A)$, $\gamma \in \mathcal{R}(A^T)$ such that $\|A\gamma\| = 1$, $\xi \geq \gamma$, $(\xi - \gamma) \in \text{ext}\,(V_2(A))$ and $\|\xi - \gamma\| = K_2(A)$. Let $J \subset \{1, 2, \ldots, n\}$ be such that $\xi_J = \gamma_J$ and $\xi_{\bar{J}} > \gamma_{\bar{J}}$. Then since $(\xi - \gamma)$ is in ext $(V_2(A))$, we must have $\text{rank}(A_{\bar{J}}) = |\bar{J}| \leq m$ (otherwise, we can find $\bar{\xi} \in \mathcal{N}(A_{\bar{J}})\backslash\{0\}$ such that

$$\tilde{\xi}_j := \begin{cases} 0 & \text{if } j \in J \\ \bar{\xi}_j & \text{if } j \notin J; \end{cases}$$

now $\tilde{\xi} \in \mathcal{N}(A)$ and for small enough $\epsilon > 0$, $(\xi + \epsilon\tilde{\xi} - \gamma)$ and $(\xi - \epsilon\tilde{\xi} - \gamma) \in V_2(A)$, a contradiction). Complete $\bar{J}$ to a basis B of A. Then $\xi_N = \gamma_N$ and $A_B\xi_B = -A_N\gamma_N$. The latter implies $\xi_B = -A_B^{-1}A_N\gamma_N$. Thus,

$$K_2(A) = \|\xi - \gamma\| = \|\xi_B - \gamma_B\| = \|A_B^{-1}A_N\gamma_N + \gamma_B\| = \|A_B^{-1}A\gamma\|.$$

Hence yielding the desired inequality

$$K_2(A) \leq \max\left\{\|A_B^{-1}A\gamma\|:\ \gamma \in \mathcal{R}(A^T), B \in \mathcal{B}(A), \|A\gamma\| = 1, \gamma_B \leq -A_B^{-1}A_N\gamma_N\right\}.$$

□

Theorem 4.6

$$\max_{G \in \mathcal{G}} K_2(AG) \leq \chi(A).$$

Proof
By Lemma 4.5,

$$K_2(A) = \max\{\|A_B^{-1}A\gamma\| : \gamma \in \mathcal{R}(A^T), B \in \mathcal{B}(A), \|A\gamma\| = 1, \gamma_B \leq -A_B^{-1}A_N\gamma_N\}.$$

So,

$$\begin{aligned}
\max_{G \in \mathcal{G}} K(AG) &\leq \max\{\|A_B^{-1}AG\gamma\| : \gamma \in \mathcal{R}(GA^T), B \in \mathcal{B}(AG), \|AG\gamma\| = 1, G \in \mathcal{G}\} \\
&= \max\{\|A_B^{-1}AG\gamma\| : G\gamma \in \mathcal{R}(A^T), B \in \mathcal{B}(A), \|AG\gamma\| = 1, G \in \mathcal{G}\} \\
&= \max\{\|A_B^{-1}A\gamma\| : \gamma \in \mathcal{R}(A^T), B \in \mathcal{B}(A), \|A\gamma\| = 1\} \\
&= \max\{\|A_B^{-1}Ax\| : x \in \mathbb{R}^n, B \in \mathcal{B}(A), \|Ax\| = 1\} \\
&= \max\{\|A_B^{-1}y\| : B \in \mathcal{B}(A), \|y\| = 1\} \\
&= \max\{\|A_B^{-1}\| : B \in \mathcal{B}(A)\} = \chi(A).
\end{aligned}$$

□

We note that the inequality above may be strict. Otherwise, using (11) we would have had $\bar{\chi}(A) = \|A\|\chi(A)$ which is clearly false in general —take for instance $A := \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$.

5 Ye's Complexity Measure for LP and Hoffman's Bound

We are going to look at two more complexity measures, η and symm. These complexity measures relate closely to the symmetry of certain geometric objects of the LP problem. Let us consider the LP problem in the following primal form:

$$\begin{array}{llrcl} (P) & \min & c^T x & & \\ & \text{subject to} & Ax & = & b \\ & & x & \in & \mathbb{R}^n_+, \end{array}$$

and the corresponding dual form:

$$\begin{array}{llrcrcl} (D) & \max & b^T y & & & & \\ & \text{subject to} & A^T y & + & s & = & c \\ & & y & & & \in & \mathbb{R}^m \\ & & & & s & \in & \mathbb{R}^n_+ \end{array}$$

where $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$.

Under the assumption that both (P) and (D) have feasible solutions, Ye [31] first defines a complexity measure for each of the problems (P) and (D):

$$\begin{array}{rcl} \eta_P & := & \min\limits_{j \in B} \max\limits_{x \in \text{opt}(P)} x_j, \\ \eta_D & := & \min\limits_{j \in N} \max\limits_{s \in \text{opt}(D)} s_j. \end{array}$$

Then, Ye [31] defines the complexity measure of the primal dual pair as the minimum of the two:

$$\eta(P, D) \quad := \quad \min\{\eta_P, \eta_D\},$$

where $\text{opt}(P)$ and $\text{opt}(D)$ denote the sets of optimal solutions of (P) and (D) respectively, and (B, N) denotes the strict complementarity partition.

Let us study these measures for feasibility problems over polyhedra expressed in Karmarkar's ([13]) standard form:

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

(This form is relevant in Subsection 5.1 as well.) We assume A to have full row rank and no zero columns because, without loss of generality, we can always eliminate the variables that correspond to zero columns in A. Let $S := \mathcal{N}(A)$ and (hence) $S^\perp = \mathcal{R}(A^T)$. (P) and its dual can now be written as a primal-dual pair of feasibility problems. See Vavasis and Ye [30] and [26].

$$\begin{array}{lrcl} (FP) & x & \in & S, \\ & \|x\|_1 & = & 1, \\ & x & \geq & 0. \end{array} \qquad \begin{array}{lrcl} (FD) & s & \in & S^\perp, \\ & \|s\|_1 & = & 1, \\ & s & \geq & 0. \end{array}$$

(FD) is the dual of (FP) in the sense that every feasible solution to the dual problem of maximizing 0 over the constraints defined by (FP), corresponds to a feasible solution of (FD), except for $s = 0$ which does not correspond to a feasible solution in (FD). In this setting, even though (FP) is always bounded, (FD) can still be infeasible (for example, $A := [1, -1]$).

When (FP) is feasible, there exists a pair (x, s) such that $x \in S, x_N = 0, x_B > 0, s \in S^{\perp}, s_N > 0, s_B = 0$, where $[B, N]$ is the corresponding strict complementarity partition with B nonempty. Furthermore, all feasible solutions of (FP) and (FD) must satisfy $x_N = 0$ and $s_B = 0$. We allow B or N to be empty. The condition $B \neq \emptyset$ is equivalent to (FP) being feasible. Similarly, $N \neq \emptyset$ is equivalent to (FD) being feasible.

Since the problems (FP) and (FD) are written in terms of the subspaces S and $S^{\perp}$, let us redefine Ye's measures accordingly. For any subspace C, define $C(1) := \{x \in C : \|x\|_1 = 1\}$. Let

$$\begin{aligned} \eta(S) &:= \min_{j \in B} \max_{x \in S(1), x \geq 0} x_j, \\ \eta(S^{\perp}) &:= \min_{j \in N} \max_{s \in S^{\perp}(1), s \geq 0} s_j, \\ \eta(A) &:= \min\{\eta(S), \eta(S^{\perp})\}. \end{aligned}$$

We define $\eta(S)$ to be 1, when (FP) is infeasible (similarly, $\eta(S^{\perp})$ is 1 if (FD) is infeasible). Notice that all of $\eta(S), \eta(S^{\perp})$ and $\eta(A)$ are positive for all A. $\eta(S)$ measures some kind of symmetry of the columns vectors of A_B about the origin. The set $\{A_B x_B : \|x_B\|_1 = 1, x_B \geq 0\}$ is the set of all convex combinations of the columns of A_B. Therefore $\{x \in S(1) : x \geq 0\}$ corresponds to the coefficients when 0 is written as convex combinations of the columns of A_B, and hence η measures their sizes. If the columns of A_B are perfectly symmetric about the origin, $\eta(S)$ would be $1/2$. And if the columns are highly asymmetric, $\eta(S)$ would be much smaller than $1/2$.

The following results give dual descriptions for $\eta(S)$ and $\eta(S^{\perp})$. For $v \in \mathbb{R}^n, J \in \{1, 2, \ldots, n\}$, let

$$\begin{aligned} v_J^+ &:= \begin{cases} -\infty & \text{if } v_J \leq 0, \\ \max_{j \in J} v_j & \text{otherwise,} \end{cases} \\ v_J^- &:= \begin{cases} +\infty & \text{if } v_J \geq 0, \\ \min_{j \in J} v_j & \text{otherwise.} \end{cases} \end{aligned}$$

Proposition 5.1 *(Tunçel [26])*
Suppose $\{e_j : j \in \{1, 2, \ldots, n\}\} \cap S = \emptyset$ *and* $B \neq \emptyset$. *Then*

$$\eta(S) = \min\{\gamma_B^+ : \gamma \in S^{\perp}, 0 < \gamma_B^+ < 1, \gamma_B^+ - \gamma_B^- = 1\}.$$

Proposition 5.2 *(Tunçel [26])*
Suppose $\{e_j : j \in \{1, 2, \ldots, n\}\} \cap S^{\perp} = \emptyset$ *and* $N \neq \emptyset$. *Then*

$$\eta(S^{\perp}) = \min\{\xi_N^+ : \xi \in S, 0 < \xi_N^+ < 1, \xi_N^+ - \xi_N^- = 1\}.$$

Note that under our assumptions, we always have $\{e_j : j \in \{1, 2, \ldots, n\}\} \cap S = \emptyset$ because A has no zero columns. Also, the condition $B \neq \emptyset$ is equivalent to (FP) being feasible. Similarly, $N \neq \emptyset$ is equivalent to (FD) being feasible.

Recently, Epelman [5], Epelman and Freund [6] presented another complexity measure based on A. Let $\mathcal{H}(A_B) := \{A_B x_B : x_B \geq 0, \|x_B\|_1 = 1\}$. That is, $\mathcal{H}(A_B)$ is the convex hull of the column vectors of A_B. Let

$$\text{symm}(A) := \max\{t : -tv \in \mathcal{H}(A_B) \text{ for all } v \in \mathcal{H}(A_B)\}.$$

Note that a generalized version of this measure has been used before by Renegar [18] to estimate complexity for convex optimization problems.

It is clear that $\mathrm{symm}(A)$ measures precisely the degree of symmetry of $\mathcal{H}(A_B)$ about the origin in $\mathbb{R}^m$ ($A \in \mathbb{R}^{m\times n}$). When $\mathcal{H}(A_B)$ is centrally symmetric (about the origin), $\mathrm{symm}(A) = 1$.

Proposition 5.3 *(Epelman [5], Epelman and Freund [6])*

$$\frac{\mathrm{symm}(A)}{1+\mathrm{symm}(A)} = \eta(S).$$

The above proposition gives an explicit relation between the two complexity measures, $\eta(S)$ and $\mathrm{symm}(A)$. Since the function $x/(1+x)$ is strictly increasing on $(0,1]$, $\eta(S)$ also measures the degree of symmetry of $\mathcal{H}(A_B)$ about the origin. In fact, by combining Proposition 5.1 and Proposition 5.3, we get the following.

Corollary 5.4 *(Ho [11])*

$$\mathrm{symm}(A) = \min_{\gamma\in S^\perp, \|\gamma_B\|=1} -\frac{\gamma_B^+}{\gamma_B^-}.$$

We can state similar results for $\eta(S^\perp)$. Let us define $H \in \mathbb{R}^{(n-m)\times n}$ to be a full row rank matrix obtained by deleting linearly dependent rows from $P_A := I - A^T(AA^T)^{-1}A$.

Corollary 5.5 *(Ho [11]) Suppose (FD) is feasible and $\{e_j : j \in \{1,2,\ldots,n\}\} \cap S^\perp = \emptyset$. Then*

$$\frac{\mathrm{symm}(H)}{1+\mathrm{symm}(H)} = \eta(S^\perp).$$

Similarly, we can combine Proposition 5.2 and Corollary 5.5.

Corollary 5.6 *(Ho [11]) Suppose (FD) is feasible and $\{e_j : j \in \{1,2,\ldots,n\}\} \cap S^\perp = \emptyset$. Then*

$$\mathrm{symm}(H) = \min_{\xi\in S, \|\xi_N\|=1} -\frac{\xi_N^+}{\xi_N^-}.$$

We now look at a relationship between the complexity measures $\eta(A)$ and $\rho(A)$. We call AG a *signing* of A, where $G \in \mathcal{G}$ and $\mathcal{G}$ is the set of diagonal matrices in $\mathbb{R}^{n\times n}$ with diagonal entries from $\{1,-1\}$. Note that $\bar{\chi}(AG) = \bar{\chi}(A)$.

Define $\underline{\eta}(A) := \min_{G\in\mathcal{G}} \eta(AG)$. We have the following fact.

Proposition 5.7 *(Todd, Tunçel and Ye [24])*

$$\frac{1}{\sqrt{n}}\underline{\eta}(A) \le \rho(A) \le \underline{\eta}(A).$$

The second inequality above can be obtained easily from the results of Vavasis and Ye [30] and Gonzaga and Lara [8]. The first inequality can be proved using Propositions 2.6 and 5.1. The second author [26] showed that in general, η may carry no information about ρ. Indeed, suppose the columns of A define an almost centrally symmetric polytope. Then there is a signing of A such that the new polytope is highly asymmetric and therefore has a very small η value, which in turn implies a very small ρ value. This suggests that $\bar{\chi}$ may not be a good complexity measure as it tends to grossly overestimate the complexity of interior-point algorithms. Even though $\bar{\chi}(A)$ grossly overestimates the amount of computational work to solve LP problems with data (A, b, c), it has been useful in estimating the work for LP problems having A as the coefficient matrix, with arbitrary b and c and arbitrary orientation of inequalities. Also, $\Delta(A)/\delta(A)$ has a similar role.

Proposition 5.7 shows that $\frac{1}{\bar{\chi}(A)}$ behaves like $\underline{\eta}(A)$ or like $\eta(AG)$, where G is "the worst signing of A" in this context. Notice that Theorem 4.6 relates Hoffman's bound to $\chi(A)$ in a similar way. It shows that $\chi(A)$ is at least $K_2(AG)$, where G is "the worst signing of A" in this latter context. Since $\eta(S)$ is essentially symm(A) and we have noticed the above parallel, we give below a brief geometric interpretation of $K_{1,\infty}$, in a special but illustrative case. Note that the essential difference between η and K is that of formulation. They both measure similar quantities; considering the problem (D), K works in the y-space and η in the s-space. See the next section for similar situations between χ and $\bar{\chi}$.

Let us now look at $K_{\alpha,\beta}(A)$ more closely. For this brief discussion, we assume that $V_\alpha(A)$ is bounded. This is true if and only if $\{v : Av = 0, v \geq 0, v \neq 0\} = \emptyset$, if and only if there exists $y \in \mathbb{R}^m$ such that $A^T y > 0$, by LP duality theory. Under this assumption,

$$\begin{aligned} K_{\alpha,\beta}(A) &= \max\{\|v\|_\beta^* : v \geq 0, \|Av\|_\alpha^* \leq 1\} \\ &= \max\{\|v\|_\beta^* : v \geq 0, \|Av\|_\alpha^* = 1\} \\ &= \max\left\{\frac{\|v\|_\beta^*}{\|Av\|_\alpha^*} : v \geq 0, Av \neq 0\right\}. \end{aligned}$$

Also $v = 0$ if and only if $Av = 0$. Hence,

$$\begin{aligned} \frac{1}{K_{\alpha,\beta}(A)} &= \min\left\{\frac{\|Av\|_\alpha^*}{\|v\|_\beta^*} : v \geq 0, v \neq 0\right\} \\ &= \min\{\|Av\|_\alpha^* : v \geq 0, \|v\|_\beta^* = 1\}. \end{aligned}$$

For the case $\alpha = 1, \beta = \infty$, we have

$$\frac{1}{K_{1,\infty}(A)} = \min\{\|Av\|_\infty : v \geq 0, e^T v = 1\}.$$

This is precisely the ∞-norm distance of the origin of $\mathbb{R}^m$ to the convex hull of the column vectors of A. Since we assume that V_1 is bounded, 0 is not in this convex hull. On the other hand,

$$\begin{aligned} \frac{1}{K_{1,\infty}(A)} &= \min\{t : \|Av\|_\infty \leq t, v \geq 0, e^T v = 1\} \\ &= \min\left\{t : \begin{pmatrix} A \\ -A \end{pmatrix} v + te \geq 0, e^T v = 1, v \geq 0\right\}. \end{aligned}$$

This is an LP problem. So by LP duality theory,

$$\frac{1}{K_{1,\infty}(A)} = \max\{\eta : [A^T | -A^T]y + \eta e \leq 0, e^T y = 1, y \geq 0\}$$

$$= \max\{\text{smallest entry of } [-A^T | A^T] y : e^T y = 1, y \geq 0\}.$$

In other words, it is the maximum of the smallest entry of any vector in the convex hull of the rows of A and their negations.

5.1 Linear Programming Solver Subroutine

In Section 7, we generalize Tardos' scheme. To do so, we need to solve LP problems with the following data. Define

$$\bar{q} := \max\left\{2\left\lceil \log\left(\frac{\Delta(A)}{\delta(A)}\right)\right\rceil, n\right\}, \tilde{q} := \bar{q}^2,$$

$$\bar{p} := 2^{\lceil \log\left(2(2m+n)^{3/2}(2mn+1)\right)\rceil} 2^{\bar{q}} \text{ and } \tilde{p} := 2^{\lceil \log\left(2(2m+n)^{3/2}(2mn+1)\right)\rceil} 2^{\tilde{q}}. \tag{12}$$

Let p be a positive integer power of two and $p \leq \tilde{p}$. We will not have any restriction on the entries of A, except that we want A to have full row rank (easily ensured). The rest of the data, b and c, for the LP solver subroutine will be restricted to the following two cases.

(i) We set $l := \left[(p+1), (p+1)^2, \ldots, (p+1)^n\right]^T$, and $b := Al$. We have $c \in \mathbb{Z}^n$ such that $\|c\|_\infty \leq \tilde{p}$.

(ii) We have $b \in \mathbb{Z}^m$, $c \in \mathbb{Z}^n$ such that $\|b\|_\infty \leq \tilde{p}$ and $\|c\|_\infty \leq \tilde{p}$.

In this subsection, we assume that b and c satisfy at least one of (i) and (ii). We also need the following function of A in our estimations.

Definition 5.8 *Let $\bar{\mathcal{A}} := [A|I]$. For every $B \in \mathcal{B}(\bar{\mathcal{A}})$ (N is the complement of B) consider the smallest absolute value of nonzero entries of*

$$\bar{\mathcal{A}}_B^{-1} u, \ \forall\, u \in \mathbb{Z}^m \text{ such that } \|u\|_\infty \leq \tilde{p},$$

$$\bar{\mathcal{A}}_B^{-1} \bar{\mathcal{A}} w, \ \forall\, w \in \mathbb{Z}^n, \text{ with entries from } (p+1), (p+1)^2, \ldots, (p+1)^n,$$

where p is a positive integer power of two and $p \leq \tilde{p}$,

$$\left[-\bar{\mathcal{A}}_N^T \bar{\mathcal{A}}_B^{-T} | I\right] v, \ \forall\, v \in \mathbb{Z}^{n+m}, \text{ such that } \|v\|_\infty \leq \tilde{p}.$$

Also consider the entries of the vectors for the same construction in which $\bar{\mathcal{A}}$ is replaced by

$$\tilde{\mathcal{A}} := \left[A^T | -A^T | -I\right].$$

These generate a finite collection of positive real numbers depending only on A. We call the minimum of all these numbers $\delta_\delta(A)$.

Note that $0 < \delta_\delta(A) \leq 1$ for all $A \in \mathbb{R}^{m \times n}$. If $A \in \mathbb{Z}^{m \times n}$ then $\delta_\delta(A) \geq 1/\Delta(A)$.

The LP problems with b and c described as above (in (i) and (ii)) depend only on A. As we show in this subsection, many algorithms can be adapted to solve such LP problems in poly $\left(n, |\log(\delta_\delta(A))|, \log\left(\frac{\Delta(A)}{\delta(A)}\right)\right)$

elementary arithmetic operations. In particular, we show that such polynomial bounds can be satisfied by employing almost any primal-dual interior-point algorithm with (mild centrality properties and) polynomial-time complexity in the Turing Machine Model. Consider the homogeneous self-dual linear programming problem $(HSDLP)$:

$$
\begin{array}{ll}
\min & (n+1)\theta \\
\text{subject to} & \begin{pmatrix} 0 & A & -b & b-Ae \\ -A^T & 0 & c & e-c \\ b^T & -c^T & 0 & e^Tc+1 \\ (Ae-b)^T & (c-e)^T & -(e^Tc+1) & 0 \end{pmatrix} \begin{pmatrix} y \\ x \\ \tau \\ \theta \end{pmatrix} \begin{array}{c} = \\ \geq \\ \geq \\ = \end{array} \begin{pmatrix} 0 \\ 0 \\ 0 \\ -(n+1) \end{pmatrix}, \\
& \begin{array}{lll} y & \text{free}, & \\ x & \geq & 0, \\ \tau & \geq & 0, \\ \theta & \text{free}. & \end{array}
\end{array}
$$

Note that $(HSDLP)$ is self-dual, and that $\theta = 0$ at every optimal solution of $(HSDLP)$. Let us define the surplus variables for the inequalities above:

$$
\begin{array}{rcl}
s & := & -A^T y + \tau c + \theta(e-c), \\
\psi & := & b^T y - c^T x + \theta(e^T c + 1).
\end{array}
$$

Then $\bar{y} := 0, \bar{x} := e, \bar{s} := e, \bar{\tau} := 1, \bar{\psi} := 1, \bar{\theta} := 1$ is feasible in $(HSDLP)$. For various facts on such formulations, see the book by Roos, Terlaky and Vial [19].

Theorem 5.9 *(Ye, Todd and Mizuno[32])*
Let $(y^*, x^*, \tau^*, \theta^* = 0, s^*, \psi^*)$ be a strictly self-complementary solution for $(HSDLP)$. Then,

1. *(P) has a solution if and only $\tau^* > 0$. In this case, x^*/τ^* is an optimal solution for (P) and $(y^*/\tau^*, s^*/\tau^*)$ is an optimal solution for (D),*
2. *if $\tau^* = 0$, then $\psi^* > 0$, which implies that $c^T x^* - b^T y^* < 0$, that is, at least one of $c^T x^*$ and $-b^T y^*$ is strictly less than zero. If $c^T x^* < 0$, then (D) is infeasible; if $-b^T y^* < 0$, then (P) is infeasible; if both $c^T x^* < 0$ and $-b^T y^* < 0$, then both (P) and (D) are infeasible.*

Consider the setting at the very beginning of Section 5. Assume both (P) and (D) have feasible solutions. Let $\left\{\left(x^{(k)}, s^{(k)}\right)\right\}$, $k \in \mathbb{Z}_+$ denote the iterates of a primal-dual interior-point algorithm (with feasible iterates). Güler and Ye [10] proved that the mild, wide neighborhood condition (or centrality condition)

$$
\frac{\min_j \left\{x_j^{(k)} s_j^{(k)}\right\}}{\left(x^{(k)}\right)^T s^{(k)}} \geq \Omega\left(\frac{1}{n}\right) \tag{13}
$$

guarantees that every limit point of $\left\{\left(x^{(k)}, s^{(k)}\right)\right\}$ is a strictly complementary pair. Mehrotra and Ye [16] and Ye [31] showed how to make such polynomial-time primal-dual interior-point algorithms terminate in $O\left(\sqrt{n}\,|\log(\eta(P,D))|\right)$ iterations.

Results of Ye-Todd-Mizuno [32] and Ye [31] also show how to terminate primal-dual interior-point algorithms (those converging to a strictly complementary pair) after $O(\sqrt{n}|\log(\eta(HSDLP))|)$ iterations. We denoted by $\eta(HSDLP)$, Ye's complexity measure applied to the problem $(HSDLP)$. Since the problem is self-dual, the notation is consistent.

Next, we will estimate $\eta(HSDLP)$. The optimal value of $(HSDLP)$ is 0. Therefore, we can represent the set of optimal solutions of $(HSDLP)$ as $(FHSDLP)$:

$$\begin{aligned} Ax &= \tau b, \\ A^T y + s &= \tau c, \\ b^T y - c^T x &= \psi, \\ e^T x + e^T s + \tau + \psi &= n+1, \\ x, s, \tau, \psi &\geq 0. \end{aligned}$$

By the last equation and the nonnegativity constraints, we have

$$0 < \eta(HSDLP) \leq n+1.$$

It remains to bound $\eta(HSDLP)$ from below and away from zero. We want to maximize each restricted variable (say x_j for some j) subject to $(FHSDLP)$. We will split the analysis into three exhaustive cases:

1. (P) and (D) both have feasible solutions,

2.(a) (D) is infeasible,

2.(b) (P) is infeasible.

As mentioned before, we will assume that b and c satisfy (i) or (ii), and we will differentiate the analysis of these two cases, whenever necessary.

Case 1.: (P) and (D) both have feasible solutions

Every solution of $(FHSDLP)$ satisfies $\psi = b^T y - c^T x = 0$, by LP weak duality and the constraint $\psi \geq 0$. Also, there exists a solution of $(FHSDLP)$ with $\tau > 0$. Let $(\bar{x}, \bar{y}, \bar{s})$ be a basic primal-dual pair of optimal solutions for (P) and (D). So for some $B \in \mathcal{B}(A)$, we have

$$\bar{x}_B = A_B^{-1} b, \bar{s}_N = c_N - A_N^T A_B^{-T} c_B,$$

where $N := \{1, \ldots, n\} \setminus B$. For case (i), we have

$$e^T \bar{x} = \|\bar{x}_B\|_1 \leq \sqrt{m}\|\bar{x}_B\| = \sqrt{m}\|A_B^{-1} A l\| \leq \sqrt{m}\|A_B^{-1} A\| \cdot \|l\| \leq \sqrt{nm}\bar{\chi}(A)(\tilde{p}+1)^n.$$

For case (ii), we have

$$e^T \bar{x} \leq \sqrt{m}\|A_B^{-1} b\| \leq \sqrt{m}\|A_B^{-1}\| \cdot \|b\| \leq m\chi(A)\tilde{p} \leq m^2 \tilde{p}\frac{\Delta(A)}{\delta(A)} \leq \tilde{p}^3 \leq \sqrt{nm}\bar{\chi}(A)(\tilde{p}+1)^n,$$

where the fourth inequality uses Proposition 2.13. Similarly,

$$e^T \bar{s} \leq \|c_N\|_1 + \|A_N^T A_B^{-T} c_B\|_1 \leq \sqrt{n}\bar{\chi}(A)\|c\|_1 \leq n^{3/2}\bar{\chi}(A)\tilde{p}.$$

Let

$$\bar{\tau} := \frac{n+1}{e^T \bar{x} + e^T \bar{s} + 1}.$$

Then $(\bar{\tau}\bar{y}, \bar{\tau}\bar{x}, \bar{\tau}\bar{s}, \bar{\tau}, \bar{\psi} := 0)$ is a solution of $(FHSDLP)$. Hence,

$$\max_{(y,x,s,\tau,\psi)\in(FHSDLP)} \tau \geq \bar{\tau} \geq \frac{n+1}{\sqrt{nm}\bar{\chi}(A)(\tilde{p}+1)^n + n^{3/2}\bar{\chi}(A)\tilde{p}+1} \geq \frac{n+1}{3n^{3/2}\bar{\chi}(A)(\tilde{p}+1)^n}.$$

Let $[B', N']$ be the (unique) strict complementarity partition (restricted to just the indices x, or s) for $(HSDLP)$. Let $j \in B'$. Then there exists a basic primal-dual pair of optimal solutions for (P) and (D), $(\bar{x}, \bar{y}, \bar{s})$, corresponding to some new basis B, such that $\bar{x}_j > 0$. Then all the above arguments apply with this new B. Since $\bar{x}_j = (A_B^{-1}b)_j \geq \delta_\delta(A)$, we have

$$\max_{(y,x,s,\tau,\psi)\in(FHSDLP)} x_j \geq \bar{\tau}\bar{x}_j \geq \frac{(n+1)\delta_\delta(A)}{3n^{3/2}\bar{\chi}(A)(\tilde{p}+1)^n}.$$

Similarly, for each $j \in N'$, there exists $\bar{s}$ corresponding to some basis B such that $\bar{s}_j > 0$. Then $j \in N$ and

$$\bar{s}_j = \left([-A_N^T A_B^{-T} | I] \begin{bmatrix} c_B \\ c_N \end{bmatrix} \right)_j \geq \delta_\delta(A).$$

Hence, we have

$$\max_{(y,x,s,\tau,\psi)\in(FHSDLP)} s_j \geq \bar{\tau}\bar{s}_j \geq \frac{(n+1)\delta_\delta(A)}{3n^{3/2}\bar{\chi}(A)(\tilde{p}+1)^n}.$$

Therefore, since $\delta_\delta(A) \leq 1$,

$$\eta(HSDLP) \geq \frac{(n+1)\delta_\delta(A)}{3n^{3/2}\bar{\chi}(A)(\tilde{p}+1)^n}$$

in this case.

Case 2.(a): (D) is infeasible

Every solution of $(FHSDLP)$ satisfies $\tau = 0$, because if (y, x, s, τ, ψ) is a solution such that $\tau > 0$, then $(y/\tau, s/\tau)$ is a feasible solution of (D). On the other hand, by Farkas' lemma,

$$\min\{c^T x : Ax = 0, e^T x = 1, x \geq 0\} < 0.$$

Let $\bar{x}$ be a basic optimal solution of this problem. So for some $B \in \mathcal{B}(A)$ and $k \in \{1, \ldots, n\} \setminus B$ such that $\bar{x}_k \neq 0$, we have $A_B \bar{x}_B = -A_k \bar{x}_k$ and $\bar{x}_j = 0$ for all $j \notin B \cup \{k\}$. It is easy to see that

$$\left(0, \frac{(n+1)\bar{x}}{1 - c^T\bar{x}}, 0, 0, \frac{-(n+1)c^T\bar{x}}{1 - c^T\bar{x}}\right) \in (FHSDLP).$$

Note that $1 = e^T\bar{x} = \bar{x}_k - \bar{x}_k e^T A_B^{-1} A_k$, which implies

$$\bar{x}_k = \frac{1}{1 - e^T A_B^{-1} A_k} > 0.$$

Now, $1 - e^T A_B^{-1} A_k \leq 1 + \|A_B^{-1} A_k\|_1 \leq 1 + \sqrt{m}\|A_B^{-1} A_k\| \leq 1 + \sqrt{m}\bar{\chi}(A)$. So,

$$0 < -c^T\bar{x} = |c^T\bar{x}| = |c_k\bar{x}_k + c_B^T(-A_B^{-1}A_k\bar{x}_k)| = \frac{|c_k - c_B^T A_B^{-1} A_k|}{1 - e^T A_B^{-1} A_k} \geq \frac{\delta_\delta(A)}{1 + \sqrt{m}\bar{\chi}(A)}. \qquad (14)$$

Also, $-c^T\bar{x} \le \tilde{p}$ (since $\bar{x} \ge 0$ and $e^T\bar{x} = 1$). So,

$$\max_{(y,x,s,\tau,\psi)\in(FHSDLP)} \psi \ge \frac{-(n+1)c^T\bar{x}}{1-c^T\bar{x}} \ge \frac{(n+1)\delta_\delta(A)}{(1+\sqrt{m}\bar{\chi}(A))(\tilde{p}+1)}.$$

Let $j \in B'$ where $[B', N']$ is, as before, the (unique) strict complementarity partition (restricted to just the subvectors x and s) for $(HSDLP)$. Let $\tilde{x}$ be a maximizer of

$$\max\{x_j : Ax = 0, e^T x = 1, x \ge 0\}.$$

Note that $\tilde{x}_j \ge \eta(\mathcal{N}(A)) \ge \eta(A)$. Also $|c^T\tilde{x}| \le \tilde{p}$. Let

$$\hat{x} := (1+\sqrt{m}\bar{\chi}(A))\bar{x} + \frac{\delta_\delta(A)}{\tilde{p}}\tilde{x}.$$

Now, by (14), $c^T\hat{x} \le 0$. So,

$$\left(0, \frac{(n+1)\hat{x}}{e^T\hat{x}-c^T\hat{x}}, 0, 0, \frac{-(n+1)c^T\hat{x}}{e^T\hat{x}-c^T\hat{x}}\right) \in (FHSDLP).$$

Note that $-c^T\hat{x} \le (1+\sqrt{m}\bar{\chi}(A))\tilde{p} + \delta_\delta(A) \le (1+\sqrt{m}\bar{\chi}(A))\tilde{p} + 1$ and $e^T\hat{x} \le 2 + \sqrt{m}\bar{\chi}(A)$. Therefore,

$$\max_{(y,x,s,\tau,\psi)\in(FHSDLP)} x_j \ge \frac{(n+1)\hat{x}_j}{e^T\hat{x}-c^T\hat{x}} \ge \frac{(n+1)\delta_\delta(A)\eta(A)}{\tilde{p}\left[(1+\sqrt{m}\bar{\chi}(A))(\tilde{p}+1)+2\right]}.$$

Now let $j \in N'$. Consider the problem

$$\max\{s_j : s \in \mathcal{R}(A^T), e^T s = 1, s \ge 0\}.$$

First note that if this problem is infeasible, then every solution of $(FHSDLP)$ satisfies $s = 0$ and hence N' is empty; and we are done. So we assume the problem has a feasible solution and because the feasible set is compact, the maximum is attained by some basic solution, say $\tilde{s}$, corresponding to some basis $B \in \mathcal{B}(A)$. Note that $\tilde{s}_j \ge \eta(\mathcal{R}(A^T)) \ge \eta(A)$. We (again) let $N := \{1, \ldots, n\} \setminus B$. Let $\tilde{y}$ be the unique vector in $\mathbb{R}^m$ such that $A^T\tilde{y} = -\tilde{s}$. For case (i), we let $l \in \mathbb{R}^n$ be as in the assumption given before. For case (ii), we let $l \in \mathbb{R}^n$ be such that $l_B := A_B^{-1}b$ and $l_N := 0$. In both cases, we have $Al = b$ and

$$b^T\tilde{y} = l^T A^T\tilde{y} = -l^T\tilde{s}.$$

For case (i), it is clear that $|l^T\tilde{s}| \le (\tilde{p}+1)^n$. This is also true for case (ii) because

$$\begin{aligned} |l^T\tilde{s}| &= |l_B^T\tilde{s}_B| = |(A_B^{-1}b)^T\tilde{s}_B| \le \|A_B^{-1}b\| \cdot \|\tilde{s}_B\| \le \chi(A)\sqrt{m} \cdot \|b\|_\infty \|\tilde{s}_B\|_1 \\ &\le \sqrt{m}\tilde{p}\chi(A) \le m^{3/2}\tilde{p}\left(\frac{\Delta(A)}{\delta(A)}\right) \le \tilde{p}^3 \le (\tilde{p}+1)^n, \end{aligned}$$

where we use Proposition 2.13 and the fact that $n \ge 3$. If $l^T\tilde{s} \le 0$, then

$$\left(\frac{(n+1)\tilde{y}}{1-l^T\tilde{s}}, 0, \frac{(n+1)\tilde{s}}{1-l^T\tilde{s}}, 0, \frac{-(n+1)l^T\tilde{s}}{1-l^T\tilde{s}}\right) \in (FHSDLP).$$

We then have

$$\max_{(y,x,s,\tau,\psi)\in(FHSDLP)} s_j \ge \frac{(n+1)\tilde{s}_j}{1-l^T\tilde{s}} \ge \frac{(n+1)\eta(A)}{(\tilde{p}+1)^n+1}.$$

If $l^T\tilde{s} > 0$, then we can easily show that

$$\left(\frac{-(n+1)(c^T\bar{x})\tilde{y}}{l^T\tilde{s}-c^T\bar{x}}, \frac{(n+1)(l^T\tilde{s})\bar{x}}{l^T\tilde{s}-c^T\bar{x}}, \frac{-(n+1)(c^T\bar{x})\tilde{s}}{l^T\tilde{s}-c^T\bar{x}}, 0, 0\right) \in (FHSDLP).$$

Now, using the fact that $-c^T\bar{x} \le \tilde{p}$, we have

$$\max_{(y,x,s,\tau,\psi)\in(FHSDLP)} s_j \ge \frac{-(n+1)(c^T\bar{x})\tilde{s}_j}{l^T\tilde{s}-c^T\bar{x}} \ge \frac{(n+1)\delta_\delta(A)\eta(A)}{(1+\sqrt{m}\bar{\chi}(A))\left[(\tilde{p}+1)^n+\tilde{p}\right]}.$$

Case 2.(b): (P) is infeasible

Note that this case does not apply to case (i), since by construction, $Al = b, l \ge 0$, and therefore (P) must have a feasible solution. So we only need to consider case (ii).

Every solution of $(FHSDLP)$ satisfies $\tau = 0$, because if (y, x, s, τ, ψ) is a solution such that $\tau > 0$, then x/τ is a feasible solution of (P). On the other hand, by Farkas' lemma,

$$\max\{b^Ty : A^Ty \le 0, e^TA^Ty = 1\} > 0.$$

Let $s = -A^Ty$. Then as before, we have $b^Ty = -l^Ts$. So the above problem can be rewritten as

$$\max\{-l^Ts : s \in \mathcal{R}(A^T), e^Ts = 1, s \ge 0\}.$$

Now let $D \in \mathbb{R}^{(n-m)\times n}$ be such that the rows are precisely a basis of $\mathcal{N}(A)$. We know $\mathcal{R}(D^T) = \mathcal{N}(A)$ and $\mathcal{N}(D) = \mathcal{R}(A^T)$. In particular, $\bar{\chi}(D) = \bar{\chi}(A)$, which we will use later on. Therefore, the above problem can be further rewritten as

$$\max\{-l^Ts : Ds = 0, e^Ts = 1, s \ge 0\}.$$

Let $\bar{s}$ be a basic optimal solution of this problem. So for some $N \in \mathcal{B}(D)$ and $k \in \{1, \ldots, n\} \setminus N$ such that $\bar{s}_k \ne 0$, we have $D_N\bar{s}_N = -D_k\bar{s}_k$ and $\bar{s}_j = 0$ for all $j \notin N \cup \{k\}$. Let $\bar{y}$ be the unique vector in $\mathbb{R}^m$ such that $A^T\bar{y} = -\bar{s}$. It is easy to see that

$$\left(\frac{(n+1)\bar{y}}{1+b^T\bar{y}}, 0, \frac{(n+1)\bar{s}}{1+b^T\bar{y}}, 0, \frac{(n+1)b^T\bar{y}}{1+b^T\bar{y}},\right) \in (FHSDLP).$$

Note that $1 = e^T\bar{s} = \bar{s}_k - \bar{s}_ke^TD_N^{-1}D_k$, which implies

$$\bar{s}_k = \frac{1}{1-e^TD_N^{-1}D_k} > 0.$$

Now,

$$1 - e^TD_N^{-1}D_k \le 1 + \|D_N^{-1}D_k\|_1 \le 1 + \sqrt{n-m}\|D_N^{-1}D_k\| \le 1 + \sqrt{n-m}\bar{\chi}(D) = 1 + \sqrt{n-m}\bar{\chi}(A).$$

Since the choice of B in the definition of l (for case (ii) in case 2(a)) does not affect the previous arguments, we can redefine l using $B := \{1, \ldots, n\} \setminus N$. It is not hard to see that $B \in \mathcal{B}(A)$. So we have

$$0 < b^T\bar{y} = |l^T\bar{s}| = |l_B^T\bar{s}_B| = |(A_B^{-1}b)^T\bar{s}_B| = |(A_B^{-1}b)_k|\bar{s}_k \ge \frac{\delta_\delta(A)}{1+\sqrt{n-m}\bar{\chi}(A)}. \qquad (15)$$

Also, $b^T\bar{y} = -l^T\bar{s} \leq (\tilde{p}+1)^n$, as we have shown before. So,

$$\max_{(y,x,s,\tau,\psi)\in(FHSDLP)} \psi \geq \frac{(n+1)b^T\bar{y}}{1+b^T\bar{y}} \geq \frac{(n+1)\delta_\delta(A)}{(1+\sqrt{n-m}\bar{\chi}(A))\left[(\tilde{p}+1)^n+1\right]}.$$

Recall that $[B', N']$ denotes, as in the previous cases, the (unique) strict complementarity partition (restricted to the subvectors x and s) for $(HSDLP)$. Now let $j \in N'$. Let $\tilde{s}$ be a maximizer of

$$\max\{s_j : s \in \mathcal{R}(A^T), e^T s = 1, s \geq 0\}.$$

Note that $\tilde{s}_j \geq \eta(A)$, and $|l^T\tilde{s}| \leq (\tilde{p}+1)^n$. Let

$$\hat{s} := (\tilde{p}+1)^n(1+\sqrt{n-m}\bar{\chi}(A))\bar{s} + \delta_\delta(A)\tilde{s}.$$

Let $\hat{y}$ be the unique vector in $\mathbb{R}^m$ such that $A^T\hat{y} = -\hat{s}$. Now, by (15), $l^T\hat{s} \leq 0$. So we have $b^T\hat{y} = -l^T\hat{s} \geq 0$. Therefore,

$$\left(\frac{(n+1)\hat{y}}{e^T\hat{s}+b^T\hat{y}}, 0, \frac{(n+1)\hat{s}}{e^T\hat{s}+b^T\hat{y}}, 0, \frac{(n+1)b^T\hat{y}}{e^T\hat{s}+b^T\hat{y}}\right) \in (FHSDLP).$$

Now,

$$-l^T\hat{s} \leq (\tilde{p}+1)^{2n}(1+\sqrt{n-m}\bar{\chi}(A)) + (\tilde{p}+1)^n.$$

Therefore,

$$\max_{(y,x,s,\tau,\psi)\in(FHSDLP)} s_j \geq \frac{(n+1)\hat{s}_j}{e^T\hat{s}+b^T\hat{y}} \geq \frac{(n+1)\delta_\delta(A)\eta(A)}{\left[(\tilde{p}+1)^{2n}+(\tilde{p}+1)^n\right](1+\sqrt{n-m}\bar{\chi}(A)) + (\tilde{p}+1)^n+1}.$$

If B' is empty, then we are done. Otherwise, let $j \in B'$. Consider the problem

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

Let $\tilde{x}$ be a basic optimal solution of this problem such that $A_B\tilde{x}_B = -\tilde{x}_j A_j$, where we called the corresponding basis B. First, we have $\tilde{x}_j \geq \eta(S) \geq \eta(A)$ by definitions. Also, $|c^T\tilde{x}| \leq \tilde{p}$. If $c^T\tilde{x} \leq 0$, then

$$\left(0, \frac{(n+1)\tilde{x}}{1-c^T\tilde{x}}, 0, 0, -\frac{(n+1)c^T\tilde{x}}{1-c^T\tilde{x}}\right) \in (FHSDLP).$$

If $c^T\tilde{x} > 0$, then

$$\left(\frac{(n+1)(c^T\tilde{x})\bar{y}}{b^T\bar{y}+c^T\tilde{x}}, \frac{(n+1)(b^T\bar{y})\tilde{x}}{b^T\bar{y}+c^T\tilde{x}}, \frac{(n+1)(c^T\tilde{x})\bar{s}}{b^T\bar{y}+c^T\tilde{x}}, 0, 0\right) \in (FHSDLP).$$

Using $\frac{\delta_\delta(A)}{1+\sqrt{n-m}\bar{\chi}(A)} \leq b^T\bar{y} \leq (\tilde{p}+1)^n$ and $c^T\tilde{x} \leq \tilde{p}$, we conclude

$$\max_{(y,x,s,\tau,\psi)\in(FHSDLP)} x_j \geq \frac{(n+1)(b^T\bar{y})\eta(A)}{b^T\bar{y}+c^T\tilde{x}} \geq \frac{(n+1)\delta_\delta(A)\eta(A)}{\left[(\tilde{p}+1)^n+\tilde{p}\right](1+\sqrt{n-m}\bar{\chi}(A))}.$$

The above lower bound on x_j also applies in the case that $c^T\tilde{x} \leq 0$. We proved the following fact.

Theorem 5.10 *Consider feasible-start primal-dual interior-point algorithms satisfying condition (13) above and have been proven to run in polynomial time, with $O(\sqrt{n}\,|\log(\eta(P,D)|)$ iteration complexity.*

Every such algorithm when applied to (HSDLP) with the staring point $\bar{y} := 0, \bar{x} := e, \bar{s} := e, \bar{\tau} := 1, \bar{\psi} := 1, \bar{\theta} := 1$, *terminates correctly in*

$$O\left(\sqrt{n}\left(|\log(\delta_\delta(A))| + n\log\left(\frac{\Delta(A)}{\delta(A)}\right) + n\log(n)\right)\right)$$

iterations.

Here we used Propositions 5.7 and 2.14 to see that

$$\eta(A) \geq \rho(A) = \frac{1}{\bar{\chi}(A)} \geq \frac{1}{\sqrt{m(n-m)+1}} \cdot \frac{\delta(A)}{\Delta(A)},$$

and so conclude that

$$|\log(\eta(A))| \leq O\left(\log(n) + \log\left(\frac{\Delta(A)}{\delta(A)}\right)\right).$$

The last inequality above can also be obtained directly from the definition of $\eta(A)$ by utilizing the techniques in Section 2. Note that the above theorem stays valid if we replace A by any submatrix of it. This is one of the reasons why in Definition 5.8, we chose $\bar{\mathcal{A}}$ as $[A|I]$, rather than just A. Each iteration can be performed in $O(n^3)$ elementary arithmetic operations.

6 Sensitivity Analysis, Hoffman's Bound, $\chi, \bar{\chi}, \Delta$, and δ.

Given an LP $\max\{b^T y : A^T y \leq c\}$, we are interested in the change in the set of optimal solutions as the vector c is varied. Let $\bar{\Delta}(A)$ denote the maximum of the absolute values of the entries of C^{-1} over all nonsingular submatrices C of A.

Proposition 6.1 *(Cook, Gerards, Schrijver, Tardos [3], [20])*
Suppose $A \in \mathbb{R}^{m\times n}$ *(not necessarily full row rank),* $c, c' \in \mathbb{R}^n$, *and* $b \in \mathbb{R}^m$, *such that both LP problems* $\max\{b^T y : A^T y \leq c\}$ *and* $\max\{b^T y : A^T y \leq c'\}$ *have optimal solutions. Then for every optimal solution* $\bar{y}$ *of* $\max\{b^T y : A^T y \leq c\}$, *there exists an optimal solution* $\bar{y}'$ *of* $\max\{b^T y : A^T y \leq c'\}$ *with*

$$\|\bar{y} - \bar{y}'\|_\infty \leq m\bar{\Delta}(A)\|c - c'\|_\infty.$$

Note that $\bar{\Delta}(A) \leq \Delta(A)/\delta(A)$ for all A, by Cramer's Rule. In particular, if $A \in \mathbb{Z}^{m\times n}$, then $\bar{\Delta}(A) \leq \Delta(A)$. In fact, Cook et al. state the above proposition in [3] for integral A, and $\bar{\Delta}(A)$ above is replaced by $\Delta(A)$.

We define, for A with full row rank,

$$\chi_1(A) := \max\{\|A_B^{-1}\|_1 : B \in \mathcal{B}(A)\},$$

and

$$\bar{\chi}_1(A) := \max\{\|A_B^{-1}A\|_1 : B \in \mathcal{B}(A)\}.$$

Using almost exactly the same arguments as in the above proof, together with Proposition 2.3, we can give an alternative sensitivity bound in terms of $\chi(A)$.

Corollary 6.2 *If the A in Proposition 6.1 has full row rank, then*

$$\|\bar{y} - \bar{y}'\|_\infty \leq \chi_1(A)\|c - c'\|_\infty.$$

Following the proof of Cook et al. we also have the following useful theorem in terms of $\bar{\chi}(A)$.

Theorem 6.3 *Let $A \in \mathbb{R}^{m \times n}$, $\operatorname{rank}(A) = m$, $c, c' \in \mathbb{R}^n$, and $b \in \mathbb{R}^m$, such that both LP problems $\max\{b^T y : A^T y + s = c, s \geq 0\}$ and $\max\{b^T y : A^T y + s = c', s \geq 0\}$ have optimal solutions. Then for every optimal solution $(\bar{y}, \bar{s})$ of the former problem, there exists an optimal solution $(\bar{y}', \bar{s}')$ of the latter problem with*

$$\|\bar{s} - \bar{s}'\|_\infty \leq (\bar{\chi}_1(A) + 1)\|c - c'\|_\infty.$$

Proof
We first show the inequality for the special case $b = 0$. Then we use the special case to establish the theorem. Assume for now that $b = 0$. Suppose for a contradiction that there exists $(\bar{y}, \bar{s})$ feasible for the first problem such that no feasible solution $(\bar{y}', \bar{s}')$ of the latter problem satisfies

$$\|\bar{s} - \bar{s}'\|_\infty \leq (\bar{\chi}_1(A) + 1)\|c - c'\|_\infty.$$

Then the system

$$A^T y + s = c', s \leq \bar{s} + pe, -s \leq -\bar{s} + pe, s \geq 0,$$

where $p := (\bar{\chi}_1(A) + 1)\|c - c'\|_\infty$, has no solution. By Farkas' lemma, there exist $x \in \mathbb{R}^n, u, v \in \mathbb{R}^n_+$ such that

$$Ax = 0, x + u - v \geq 0, (c')^T x + \bar{s}^T(u - v) + p(e^T u + e^T v) < 0.$$

Note that if $u = v = 0$, then the above x proves that the system $\{A^T y + s = c', s \geq 0\}$ is infeasible, a contradiction. Therefore, $u + v \neq 0$. Let

$$\bar{u} := \frac{u}{\|u + v\|_1}, \bar{v} := \frac{v}{\|u + v\|_1},$$

so that $\|\bar{u} + \bar{v}\|_1 = 1$. Let $\bar{x}$ be a basic optimal solution of

$$\min\{(c')^T x : Ax = 0, x \geq -(\bar{u} - \bar{v})\}.$$

Note that this problem has a feasible solution (for example, $x/\|u + v\|_1$). Also it is bounded, because otherwise there exists $d \in \mathbb{R}^n_+$ such that $d \neq 0, Ad = 0, (c')^T d < 0$ which implies that $\{A^T y + s = c', s \geq 0\}$ is infeasible, a contradiction. Note that $\bar{x} = \tilde{x} - (\bar{u} - \bar{v})$, where, for some $B \in \mathcal{B}(A)$,

$$\tilde{x}_B = A_B^{-1} A(\bar{u} - \bar{v}) \geq 0, \tilde{x}_N = 0.$$

Thus,

$$\|\bar{x}\|_1 \leq \|\tilde{x}\|_1 + \|\bar{u} - \bar{v}\|_1 \leq \|A_B^{-1} A(\bar{u} - \bar{v})\|_1 + \|\bar{u} + \bar{v}\|_1 \leq \|A_B^{-1} A\|_1 + 1 \leq \bar{\chi}_1(A) + 1.$$

This gives a contradiction since

$$\begin{aligned}
0 &> (c')^T \left(\frac{x}{\|u + v\|_1}\right) + \bar{s}^T(\bar{u} - \bar{v}) + p \\
&\geq (c')^T \bar{x} + \bar{s}^T(\bar{u} - \bar{v}) + p \\
&\geq (c')^T \bar{x} - (c - A^T \bar{y})^T \bar{x} + p \\
&= (c' - c)^T \bar{x} + p \\
&\geq -\|c - c'\|_\infty \|\bar{x}\|_1 + p \\
&\geq -(\bar{\chi}_1(A) + 1)\|c - c'\|_\infty + p = 0.
\end{aligned}$$

So there exists $(\bar{y}', \bar{s}')$ feasible in the second system of the theorem such that

$$||\bar{s} - \bar{s}'||_\infty \leq (\bar{\chi}_1(A) + 1)||c - c'||_\infty.$$

This completes the proof for the special case $b = 0$.

Now, consider the general case. Let $(\bar{y}, \bar{s}, \bar{x})$ be an optimal solution of

$$\max\{b^T y : A^T y + s = c, s \geq 0\}$$

and its dual. Let $J := \{j : \bar{s}_j = 0\}$. Let (y^*, s^*) be an optimal solution of

$$\max\{b^T y : A^T y + s = c', s \geq 0\}.$$

We have, by complementary slackness, $\bar{x}_j = 0$ for all $j \notin J$, and so

$$A_J \bar{x}_J = b, \bar{x}_J \geq 0.$$

Also,

$$A_J^T \bar{y} = c_J \geq c'_J - ||c_J - c'_J||_\infty e \geq A_J^T y^* - ||c - c'||_\infty e.$$

We proved that

$$A^T \bar{y} \leq c, -A_J^T \bar{y} \leq ||c - c'||_\infty e - A_J^T y^*.$$

Also the system

$$A^T y \leq c', -A_J^T y \leq -A_J^T y^*$$

has a feasible solution (for example, y^*). Therefore, by applying the first part of the proof (with $b = 0$) to these two systems of inequalities, we conclude that there exists $(\bar{y}', \bar{s}')$ such that

$$A^T \bar{y}' + \bar{s}' = c', -A_J^T \bar{y}' \leq -A_J^T y^*, \bar{s}' \geq 0,$$

and

$$||\bar{s} - \bar{s}'||_\infty \leq (\bar{\chi}_1([A| - A_J]) + 1)||c - c'||_\infty.$$

Note that

$$b^T \bar{y}' = \bar{x}_J^T A_J^T \bar{y}' \geq \bar{x}_J^T A_J^T y^* = b^T y^*.$$

Therefore, $(\bar{y}', \bar{s}')$ is an optimal solution of $\max\{b^T y : A^T y + s = c', s \geq 0\}$. We have (trivially, from (1))

$$\bar{\chi}_1([A| - A_J]) = \bar{\chi}_1(A).$$

We conclude

$$||\bar{s} - \bar{s}'||_\infty \leq (\bar{\chi}_1(A) + 1)||c - c'||_\infty$$

and this completes the proof. □

Using (5), we easily have the following facts.

Corollary 6.4 *Under the same assumptions as in Theorem 6.3, we have*

$$||\bar{y} - \bar{y}'||_\infty \leq \sqrt{m}\chi(A)||c - c'||_\infty$$

and

$$||\bar{s} - \bar{s}'||_\infty \leq (\sqrt{m}\bar{\chi}(A) + 1)||c - c'||_\infty.$$

Note that converting norms inside the proof of Theorem 6.3 would also give the same constant for the bound in terms of χ; however, for $\bar{\chi}$, we would have to resort to Proposition 2.8, leading to an unnecessary factor of $\sqrt{2}$ in the upper bound.

For the LP problems in the primal form, we define

$$\bar{\chi}_\infty(A) := \max\{\|A_B^{-1}A\|_\infty : B \in \mathcal{B}(A)\}$$

and prove by the above techniques the following fact.

Theorem 6.5 *Suppose $A \in \mathbb{R}^{m\times n}$ has full row rank, $c \in \mathbb{R}^n$ and $l, l' \in \mathbb{R}^n$ such that both LP problems $\min\{c^T x : Ax = 0, x \geq -l\}$ and $\min\{c^T x : Ax = 0, x \geq -l'\}$ have optimal solution(s). Then for every optimal solution $\bar{x}$ of the former problem, there exists an optimal solution $\bar{x}'$ of the latter problem with*

$$\|\bar{x} - \bar{x}'\|_\infty \leq (\bar{\chi}_\infty(A) + 2) \cdot \|l - l'\|_\infty.$$

7 Tardos' Theorem

Tardos [22] shows that any LP problem $\max\{b^T y : A^T y \leq c\}$ (with integer or rational data) can be solved in at most poly(size(A)) elementary arithmetic operations on numbers of size polynomially bounded by size(A, b, c). Here we extend her ideas to the case of real number data. The following proofs are very similar to Tardos', and Schrijver's presentation in [20].

7.1 Assumptions

Tardos [22] works with integer (can also easily handle rational numbers) data and the Turing Machine Model. So, not only the number of arithmetic operations but also the sizes of the numbers in intermediate steps are to be bounded by polynomial functions of the input size. In this section, we work with real numbers and utilize Blum-Shub-Smale (BSS) Model (see the book by Blum, Cucker, Shub and Smale [2]). Our final complexity bounds involve complexity measures of the input other than the dimension n. Therefore, to unify the approaches of Vavasis-Ye and Tardos, we introduce below some integers to the complexity model. The sizes of the integers are polynomially bounded in terms of the sizes of the integers closest to our complexity measures. We allow comparison of real numbers to such integers in $O(1)$ time. As a result, determining the "ceiling" of a real number arising from the input data in polynomially many steps of BSS model becomes a polynomial operation for our purposes in this paper. For simplicity, we assume that we can compute the ceiling of such real numbers in $O(1)$ time and consider this operation an *elementary operation.*

Here are some other assumptions that we will make:

1. $A \in \mathbb{R}^{m\times n}$ has full row rank.
2. We can solve the LP problems of the form $(D) : \max\{b^T y : A^T y \leq c\}$, where $c \in \{-1, 0, 1\}^n$, $b \in \{-1, 0, 1\}^m$, in at most poly($n, \log(\bar{\chi}(A))$) elementary arithmetic operations.

As we noted before in various settings, Assumption 1 can be made without loss of generality, and is assumed throughout Section 7. Also note that Assumption 2 holds for the Vavasis-Ye algorithm. It is possible that there exists simpler algorithms than Vavasis-Ye's (and with better complexity bounds) for LP problems with the above-mentioned special data.

In this section, we first do our analysis under Assumption 2. This will lay down most of the main ideas and main technical tools needed. Using these, we then show that removing Assumption 2 is possible by utilizing the results of Subsection 5.1.

Proposition 7.1 *Suppose Assumption 2 holds. Then we can solve* (D)*, where* $c \in \mathbb{R}^n \setminus \{0\}, b \in \mathbb{R}^m \setminus \{0\}$*, in at most*

$$\text{poly}\left(n, \log(\bar{\chi}(A)), \log\left(\frac{\|c\|_\infty}{\min_{c_j \neq 0} |c_j|}\right)\right)$$

elementary arithmetic operations.

Proof
The feasible set $\{A^T y \leq c\}$ can be rewritten as $\{CA^T y \leq Cc\}$, where $C \in \mathbb{R}^{n \times n}$, diagonal, such that for all $j \in \{1, \ldots, n\}$,

$$C_{jj} := \begin{cases} 1/|c_j|, & \text{if } c_j \neq 0, \\ 1/\|c\|_\infty, & \text{if } c_j = 0. \end{cases}$$

Now the problem $\max\{b^T y : CA^T y \leq Cc\}$ is equivalent to $\max\{(Bb)^T w : CA^T Bw \leq Cc\}$, where $w := B^{-1} y$ and $B \in \mathbb{R}^{m \times m}$, diagonal, such that for all $i \in \{1, \ldots, m\}$,

$$B_{ii} := \begin{cases} 1/|b_i|, & \text{if } b_i \neq 0, \\ 1/\|b\|_\infty, & \text{if } b_i = 0. \end{cases}$$

Now $Cc \in \{-1, 0, 1\}^n$ and $Bb \in \{-1, 0, 1\}^m$. So by Assumption 2, we can solve $\max\{b^T y : A^T y \leq c\}$ in at most $\text{poly}(n, \log(\bar{\chi}(BAC)))$ elementary arithmetic operations. Now $\bar{\chi}(BAC) = \bar{\chi}(AC)$ since B is nonsingular. Also,

$$\begin{aligned} \|(AC)^T (ACD(AC)^T)^{-1} ACD\| &= \|CA^T (A(CDC)A^T)^{-1} A(CDC)C^{-1}\| \\ &\leq \|C\| \cdot \|C^{-1}\| \cdot \|A^T (A(CDC)A^T)^{-1} A(CDC)\|, \end{aligned}$$

for all positive definite diagonal $n \times n$ matrices D. Therefore,

$$\bar{\chi}(AC) \leq \|C\| \cdot \|C^{-1}\| \cdot \bar{\chi}(A) = \frac{\max_j C_{jj}}{\min_j C_{jj}} \bar{\chi}(A) = \frac{\|c\|_\infty}{\min_{c_j \neq 0} |c_j|} \bar{\chi}(A).$$

So we get the bound

$$\text{poly}\left(n, \log(\bar{\chi}(A)), \log\left(\frac{\|c\|_\infty}{\min_{c_j \neq 0} |c_j|}\right)\right).$$

□

7.2 Deciding the Feasibility of $A^T y \leq c$

In this subsection, we describe an iterative algorithm to determine whether $A^T y \leq c$ has a solution and if not, find a certificate of its infeasibility.

We first use Gaussian elimination to remove any redundant rows of A, to get $\bar{A}$. (Clearly, the given data A has no redundant rows since it has full row rank; but, this procedure is necessary beyond the first iteration as our A changes.) As before, we can replace A by $\bar{A}$ without changing our problem. Now A has full row rank.

Let $c' := (I - A^T(AA^T)^{-1}A)c$. Then for all $d \in \mathcal{N}(A)$,

$$c'^T d = c^T d - d^T A^T (AA^T)^{-1} Ac = c^T d.$$

Since c' is the orthogonal projection of c onto $\mathcal{N}(A)$,

$$\{y : A^T y \leq c\} = \emptyset \quad \Leftrightarrow \quad \{y : A^T y \leq c'\} = \emptyset.$$

Therefore, we can replace c by c' without changing our problem. Now we have $c \in \mathcal{N}(A)$.

If $c = 0$, then $y = 0$ is a feasible solution, and we are done. So, we replace c by $c/||c||_\infty$. This does not change our problem since the feasibility of the system is invariant under positive scalar multiplication of c (or independently A). Now we have $||c||_\infty = 1$.

Suppose we are given an integer p such that $p \geq 2n^{3/2}(\bar{\chi}(A))^2$. We first solve $A^T y \leq \lceil pc \rceil$. If it has no solution, then we have a $d \geq 0$ such that $Ad = 0$ and $\lceil pc \rceil^T d < 0$. This d is also a certificate of the infeasibility of $A^T y \leq c$, since $(pc)^T d \leq \lceil pc \rceil^T d < 0$, which implies $c^T d < 0$. So we stop.

Therefore, we assume we get $(\bar{y}, \bar{s})$ such that

$$A^T \bar{y} + \bar{s} = \lceil pc \rceil, \bar{s} \geq 0. \tag{16}$$

Lemma 7.2 *Let $c \in \mathcal{N}(A)$, $c \neq 0$. Suppose (y, s) is given such that $A^T y + s = c$. Then $||s|| \geq ||c||/\bar{\chi}(A)$.*

Proof
We use Proposition 2.6. Note that since the 2-norms are used here, we can interchange $\mathcal{R}(A^T)$ and $\mathcal{N}(A)$ in Proposition 2.6, as we noted earlier. Let $\gamma := c/||c||, \xi := A^T y$, and

$$J := \{j \in \{1, \ldots, n\} : \text{sign}(\gamma_j) \neq \text{sign}(\xi_j)\}.$$

Note that $J \neq \emptyset$ because otherwise $\text{sign}(c) = \text{sign}(A^T y)$ together with $c \in \mathcal{N}(A)$ would imply $c = 0$, a contradiction. So (γ, ξ, J) is a feasible solution to the minimization problem in Proposition 2.6, and hence $||c_J|| \geq ||c||/\bar{\chi}(A)$. Now, for each $j \in J$,

$$|s_j| = |c_j + (-(A^T y)_j)| = |c_j| + |(A^T y)_j| \geq |c_j|,$$

where the second equality above uses the fact that c_j and $-(A^T y)_j$ either have the same sign or at least one of them is 0. So, $||s|| \geq ||s_J|| \geq ||c_J|| \geq ||c||/\bar{\chi}(A)$. □

From (16), we have

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

and hence by Lemma 7.2,

$$\|\bar{s} + pc - \lceil pc \rceil\| \geq \frac{\|pc\|}{\bar{\chi}(A)} \geq \frac{p\|c\|_\infty}{\bar{\chi}(A)} = \frac{p}{\bar{\chi}(A)}.$$

So,

$$\|\bar{s}\| \geq \frac{p}{\bar{\chi}(A)} - \|pc - \lceil pc \rceil\| > \frac{p}{\bar{\chi}(A)} - \sqrt{n},$$

and hence,

$$\|\bar{s}\|_\infty \geq \frac{\|\bar{s}\|}{\sqrt{n}} > \frac{p}{\sqrt{n}\bar{\chi}(A)} - 1 \geq 2n\bar{\chi}(A) - 1 \geq n\bar{\chi}(A). \tag{17}$$

Let $J := \{j \in \{1, \ldots, n\} : \bar{s}_j < \|\bar{s}\|_\infty\}$. (We could have defined $J := \{j \in \{1, \ldots, n\} : \bar{s}_j < n\bar{\chi}(A)\}$ and the following arguments would work as well. But the difficulty is we cannot compute $n\bar{\chi}(A)$ efficiently.)

Lemma 7.3 *The system $A^T y \leq c$ has a feasible solution if and only if $A_J^T y \leq c_J$ has a feasible solution.*

Proof
Clearly, if $A^T y \leq c$ has a feasible solution, so does $A_J^T y \leq c_J$ since the latter has possibly fewer constraints. If $A^T y \leq c$ has no solution, then by Farkas' lemma, there exists $d \geq 0$ such that $Ad = 0, c^T d < 0$, and (without loss of generality) $e^T d = 1$. We can assume that d is an extreme point of the compact set

$$\{d : Ad = 0, e^T d = 1, d \geq 0\}.$$

So, by Corollary 2.11, we have

$$\min\{|d_j| : d_j \neq 0\} \geq \frac{1}{n\bar{\chi}(A)}.$$

Now,

$$d^T \bar{s} = d^T(\lceil pc \rceil - pc) + pc^T d - d^T A^T \bar{y} < 1.$$

For each $j \notin J, \bar{s}_j \geq n\bar{\chi}(A)$, and so if $d_j > 0$, then $d_j \bar{s}_j \geq 1$, which contradicts $d^T \bar{s} < 1$. Therefore, $d_j = 0$ for all $j \notin J$. So d_J satisfies $d_J \geq 0, A_J d_J = 0$ and $c_J^T d_J < 0$. Hence by Farkas' lemma, $A_J^T y \leq c_J$ has no solution. □

If $A_J^T y \leq c_J$ has no solution, then we have a $d_J \geq 0$ such that $A_J d_J = 0$ and $c_J^T d_J < 0$. By inserting zero(es) to d_J, we have a $d \geq 0$ such that $Ad = 0$ and $c^T d < 0$. This is a certificate of the infeasibility of $A^T y \leq c$.

Therefore, we can repeat this algorithm with the data (A_J, c_J). Since we remove at least one column from A to get A_J in each iteration, the algorithm will terminate in at most n iterations.

We now look at the complexity of running the above algorithm. In each iteration, we solve $A^T y \leq \lceil pc \rceil$. Note that

$$\|\lceil pc \rceil\|_\infty = \lceil \|pc\|_\infty \rceil = p,$$

and

$$\min_{\lceil pc_j \rceil \neq 0} |\lceil pc_j \rceil| \geq 1.$$

Therefore, by using the proof of Proposition 7.1 for the case $b = 0$, we have proven that if Assumption 2 holds, we can solve $A^T y \leq \lceil pc \rceil$ in at most $\text{poly}(n, \log(\bar{\chi}(A)), \log(p))$ elementary arithmetic operations. Here we use Proposition 2.4 repeatedly to conclude that $\bar{\chi}(A_J) \leq \bar{\chi}(A)$ in every iteration.

Proposition 7.4 *Suppose Assumption 2 holds and that we are given an integer* $p \geq 2n^{3/2}(\bar{\chi}(A))^2$*. Then in at most* $\mathrm{poly}(n, \log(\bar{\chi}(A)), \log(p))$ *elementary arithmetic operations, we can determine whether* $A^T y \leq c$ *has a solution, and if not, find a certificate of its infeasibility.*

Similarly we have the following result, in which we use the algorithm and the analysis in Subsection 5.1 and the relation (10).

Proposition 7.5 *Suppose we are given p, an integer power of 2, that is at least as large as* $2n^{3/2}(\bar{\chi}(A))^2$*. Then in at most* $\mathrm{poly}(n, |\log(\delta_\delta(A))|, \log(\Delta(A)/\delta(A)), \log(p))$ *elementary arithmetic operations, we can determine whether* $A^T y \leq c$ *has a solution, and if not, find a certificate of its infeasibility.*

7.3 Main Results

From now on, we assume that $c \in \mathbb{R}^n \setminus \{0\}$, and $b \in \mathbb{R}^m \setminus \{0\}$.

Proposition 7.6 *Suppose Assumption 2 holds,* (D) *is feasible and we are given an integer* $p \geq 2n^{3/2}(\bar{\chi}(A))^2$*. Then in at most* $\mathrm{poly}(n, \log(\bar{\chi}(A)), \log(p))$ *elementary arithmetic operations, we can either:*

(i) find z *such that* $A^T z = c$*, or*
(ii) detect that (D) *is unbounded, or*
(iii) find an inequality $a^T y \leq \gamma$ *in* $A^T y \leq c$ *such that* $a^T y^* < \gamma$ *for some optimal solution* y^* *of* (D)*.*

Proof
Let z be the (unique) minimizer of $\|A^T z - c\|$. z can be computed by solving $AA^T z = Ac$ using a good implementation of Gaussian elimination, in $\mathrm{poly}(n)$ elementary arithmetic operations. Let $c' := c - A^T z$. If $c' = 0$, then we have found z that satisfies condition (i) above. So we assume $c' \neq 0$. Let

$$c'' := \frac{p}{\|c'\|_\infty} c'.$$

Note that $A^T y \leq c''$ arises from $A^T y \leq c$ by a translation and a scaling. Hence maximizing $b^T y$ over $A^T y \leq c$ is equivalent to maximizing $b^T y$ over $A^T y \leq c''$ in the sense that y^* is an optimal solution of $\max\{b^T y : A^T y \leq c\}$ if and only if $(p/\|c'\|_\infty)(y^* - z)$ is an optimal solution of $\max\{b^T y : A^T y \leq c''\}$. Also note that $c'' \in \mathcal{N}(A)$, since c' is.

Now we solve the problem $(D') : \max\{b^T y : A^T y \leq \lceil c'' \rceil\}$. Note that (D') is feasible since (D) is and $\{y : A^T y \leq c''\} \subseteq \{y : A^T y \leq \lceil c'' \rceil\}$. Also, (D') is unbounded if and only if (D) is unbounded because each of these is true if and only if there exists $d \neq 0$ such that $A^T d \leq 0$ and $b^T d > 0$. Hence condition (ii) is satisfied. We can now assume both (D) and (D') are bounded. Let $(\bar{y}, \bar{s})$ be an optimal solution of (D'). We have by (17) that $\|\bar{s}\|_\infty \geq n\bar{\chi}(A)$. Corollary 6.4 implies that there exists an optimal solution $(\bar{y}', \bar{s}')$ of $\max\{b^T y : A^T y \leq c''\}$ such that

$$\|\bar{s} - \bar{s}'\|_\infty \leq \left[\sqrt{m}\bar{\chi}(A) + 1\right] \|c'' - \lceil c'' \rceil\|_\infty < \sqrt{m}\bar{\chi}(A) + 1.$$

Therefore, we pick the inequality with the largest $\bar{s}_j$ among the inequalities $A^T y \leq \lceil c'' \rceil$, and condition (iii) is satisfied.

We now look at the complexity of solving (D') using Proposition 7.1. We have

$$\|\lceil c''\rceil\|_\infty = \lceil \|c''\|_\infty \rceil = p.$$

Also, $\lceil c''\rceil \neq 0$ and $\min_{\lceil c''_j\rceil \neq 0} |\lceil c''_j\rceil| \geq 1$. So,

$$\log\left(\frac{\|\lceil c''\rceil\|_\infty}{\min_{\lceil c''_j\rceil\neq 0} |\lceil c''_j\rceil|}\right) \quad \leq \quad \log(p+1);$$

therefore, the required time bound is satisfied. □

Proposition 7.7 *Suppose Assumption 2 holds and that we are given an integer* $p \geq 2n^{5/2}m\left(\frac{\Delta(A)}{\delta(A)}\right)^2$. *Then we can find a solution of the system* $A^Ty \leq c$ *or a certificate of its infeasibility in at most*

$$\text{poly}\left(n, \log\left(\frac{\Delta(A)}{\delta(A)}\right), \log(p)\right)$$

elementary arithmetic operations.

Proof
Let $\hat{b} := A\left(p+1, (p+1)^2, \cdots, (p+1)^n\right)^T$. We apply Proposition 7.4 to test whether $(\hat{D}) : \max\{\hat{b}^Ty : A^Ty \leq c\}$ has a feasible solution, and if not, we obtain a certificate of its infeasibility. Therefore, we assume that $(\hat{D})$ is feasible. Since $(\hat{D})$ is not unbounded (by construction of $\hat{b}$), $(\hat{D})$ has optimal solution(s).

Suppose $\hat{b}$ is a linear combination of fewer than m columns of A. Then there exists an $m \times (m-1)$ submatrix C of A of rank $m-1$, so that the matrix $[C|\hat{b}]$ is singular. Hence,

$$\begin{aligned} 0 &= \det[C|\hat{b}] \\ &= (p+1)\det[C|A_1] + (p+1)^2\det[C|A_2] + \cdots + (p+1)^n\det[C|A_n], \end{aligned}$$

where A_j denotes the jth column of A. Suppose $\det[C|A_j] \neq 0$ for some j. Let k be the largest j such that $\det[C|A_j] \neq 0$. Then

$$\begin{aligned} 0 &= \sum_{j=1}^{k-1}\left[(p+1)^j(\pm\det[C|A_j])\right] + (p+1)^k|\det[C|A_k]| \\ &\geq -\Delta(A)\sum_{j=1}^{k-1}(p+1)^j + (p+1)^k\delta(A) \\ &= -\Delta(A)(p+1)\frac{(p+1)^{k-1}-1}{(p+1)-1} + (p+1)^k\delta(A) \\ &= (p+1)^k\left(\delta(A) - \frac{\Delta(A)}{p}\right) + \frac{\Delta(A)(p+1)}{p} > 0, \end{aligned}$$

since

$$p \geq \frac{\Delta(A)}{\delta(A)}.$$

This gives a contradiction. So $\det[C|A_j] = 0$ for all $j \in \{1, \ldots, n\}$, contradicting the fact that A has rank m. So $\hat{b}$ is not a linear combination of fewer than m columns of A. Therefore, $(\hat{D})$ is attained at a unique minimal face.

We now apply Proposition 7.6 to $(\hat{D})$. If it returns a z such that $A^T z = c$, we stop. Otherwise, we have an inequality $a^T y \leq \gamma$ in $A^T y \leq c$ such that $a^T y^* < \gamma$ for some optimal solution y^* of $(\hat{D})$. Let $\tilde{A} \in \mathbb{R}^{m \times (n-1)}$ be A with the column a removed, and $\tilde{c} \in \mathbb{R}^{(n-1)}$ be c with the corresponding entry γ removed. We then solve the more relaxed problem $\max\{\hat{b}^T y : \tilde{A}^T y \leq \tilde{c}\}$ and repeat the above. Note that $\tilde{A}$ must have full row rank in order to apply Proposition 7.6 to the new relaxed problem. So we perform the following procedures to reformulate this problem. We do Gaussian elimination to eliminate any redundant row of $[\tilde{A}|\hat{b}]$ to get $[\bar{A}|\bar{b}]$. Now,

$$\begin{aligned} \max\{\hat{b}^T y : \tilde{A}^T y \leq \tilde{c}\} &= \min\{\tilde{c}^T \tilde{x} : \tilde{A}\tilde{x} = \hat{b}, \tilde{x} \geq 0\} \\ &= \min\{\tilde{c}^T \tilde{x} : \bar{A}\tilde{x} = \bar{b}, \tilde{x} \geq 0\} \\ &= \max\{\bar{b}^T \bar{y} : \bar{A}^T \bar{y} \leq \tilde{c}\}. \end{aligned}$$

It is not hard to see that the first problem (and hence all of them) has an optimal solution (so the equations above are justified). Since the system $\bar{A}\tilde{x} = \bar{b}$ is consistent, $\bar{A}$ must have full row rank. So we apply Proposition 7.6 to the last problem above. If it returns a $\bar{z}$ such that $\bar{A}^T \bar{z} = \tilde{c}$, then $\tilde{A}^T z = \tilde{c}$, where z is obtained from $\bar{z}$ by adding a zero entry in the place that corresponds to the redundant row of $\tilde{A}$ being eliminated earlier. Otherwise, it returns an inequality $\bar{a}^T \bar{y} \leq \tilde{\gamma}$ in $\bar{A}^T \bar{y} \leq \tilde{c}$ such that $\bar{a}^T \bar{y}^* < \tilde{\gamma}$ for some optimal solution $\bar{y}^*$. Let y^* be obtained from $\bar{y}^*$ by adding a zero entry as before. Then y^* is an optimal solution of $\max\{\hat{b}^T y : \tilde{A}^T y \leq \tilde{c}\}$ because $\tilde{A}^T y^* = \bar{A}^T \bar{y}^* \leq \tilde{c}$ and $\hat{b}^T y^* = \bar{b}^T \bar{y}^*$. Also, $\tilde{a}^T y^* = \bar{a}^T \bar{y}^* < \tilde{\gamma}$.

Note that for each submatrix C of A, we have (using Proposition 2.14),

$$2n^{3/2}\,(\bar{\chi}(C))^2 \leq 2n^{5/2} m \left(\frac{\Delta(C)}{\delta(C)}\right)^2 \leq p.$$

Hence p satisfies the supposition of Proposition 7.6 *every* time it is being called.

By repeatedly applying Proposition 7.6, we obtain an ordering of the inequalities in $A^T y \leq c$, say, $\alpha_1^T y \leq \gamma_1$, $\alpha_2^T y \leq \gamma_2, \ldots, \alpha_n^T y \leq \gamma_n$, such that for some $r, 1 \leq r \leq n-1$, and some $z \in \mathbb{R}^m$:

- $\alpha_j^T z = \gamma_j$, for all $r+1 \leq j \leq n$,
- for each $1 \leq j \leq r$, $\alpha_j^T y^j < \gamma_j$ for some optimal solution y^j of $\max\{\hat{b}^T y : \alpha_k^T y \leq \gamma_k, \forall k \geq j\}$.

That is, we run Proposition 7.6 r times, by removing one inequality each time from $A^T y \leq c$ until we find a z that satisfies the remaining inequalities as equalities. Since the maximum is attained at a unique minimal face, the optimal solution set can be written as

$$\{y : A_=^T y = c_=\} = \{y : A_=^T y = c_=, A_<^T y < c_<\},$$

where $([A_<^T|c_<], [A_=^T|c_=])$ is a row-partition of $[A^T|c]$. It is easy to see that the rows of $A_<^T$ are precisely $\{\alpha_j^T : 1 \leq j \leq r\}$, whereas the rows of $A_=^T$ are precisely $\{\alpha_j^T : r+1 \leq j \leq n\}$. So $A_=^T z = c_=$, which implies $A_<^T z < c_<$. Therefore, z is a feasible solution of $(\hat{D})$.

We now look at the complexity of the above algorithm. We apply Proposition 7.4 once to $(\hat{D})$, which takes time

$$\text{poly}(n, \log(\bar{\chi}(A)), \log(p)) \leq \text{poly}\left(n, \log\left(\frac{\Delta(A)}{\delta(A)}\right), \log(p)\right),$$

by (10).

Afterwards, we apply Proposition 7.6 at most n times. At the kth time ($1 \le k \le r+1$), Proposition 7.6 takes at most $\text{poly}(n, \log(\bar{\chi}(A^{(k)})), \log(p))$ elementary arithmetic operations, where $A^{(1)} := A$ and for $k \ge 2$, $A^{(k)}$ is obtained by first removing some column of $A^{(k-1)}$, and then removing any redundant row. By Proposition 2.4, we have $\bar{\chi}(A^{(k)}) \le \bar{\chi}(A^{(k-1)}) \le \bar{\chi}(A)$, for all $k \ge 2$, and we can again use (10). □

Theorem 7.8 *If Assumption 2 holds, then we can solve the primal-dual LP problems*

$$(P) : \min\{c^T x : Ax = b, x \ge 0\} \text{ and } (D) : \max\{b^T y : A^T y \le c\}$$

in at most $\text{poly}\left(n, \log\left(\frac{\Delta(A)}{\delta(A)}\right)\right)$ *elementary arithmetic operations.*

Proof
Suppose we are given an integer $p \ge \bar{p}$, where $\bar{p}$ is defined in (12). We first describe an algorithm for solving the given LPs, and later explain how to obtain such a p. We apply Proposition 7.7 to test if $\{A^T y \le c\}$ and $\{Ax = b, x \ge 0\}$ are feasible, where the latter is the same as

$$\left\{ \begin{pmatrix} A \\ -A \\ -I \end{pmatrix} x \le \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix} \right\}.$$

(To use Proposition 7.7 for the above displayed data, we apply Propositions 2.14 and 2.4 to the matrix $[A^T | -A^T | -I]$ and note that $\bar{p}$ is large enough for the application of Proposition 7.7—and the results it uses—to this matrix too.) If one of them is infeasible, then we stop (having determined the status of each problem). Therefore, we may assume that both (P) and (D) are feasible.

By repeated application of Proposition 7.6 (as in the proof of Proposition 7.7, and we again have $2n^{3/2}(\bar{\chi}(C))^2 \le p$, for all submatrices C of A), we can split $\{A^T y \le c\}$ into $\{A_{(1)}^T y \le c_{(1)}, A_{(2)}^T y \le c_{(2)}\}$ and find a vector z, such that $A_{(2)}^T z = c_{(2)}$ and $A_{(1)}^T y^* < c_{(1)}$ for some optimal solution y^* of $\max\{b^T y : A^T y \le c\}$. Let $(x_{(1)}^T, x_{(2)}^T)^T$ be a partition of any primal solution x such that $x_{(1)}^T$ corresponds to $A_{(1)}^T$ and $x_{(2)}^T$ corresponds to $A_{(2)}^T$. Hence every primal optimal solution x satisfies $x_{(1)} = 0$. So,

$$\begin{aligned} \min\{c^T x : Ax = b, x \ge 0\} &= \min\{c_{(2)}^T x_{(2)} : A_{(2)} x_{(2)} = b, x_{(2)} \ge 0\} \\ &= \max\{b^T y : A_{(2)}^T y \le c_{(2)}\}. \end{aligned}$$

Using Proposition 7.7, we can find a feasible solution $x_{(2)}^*$ of

$$\left\{ \begin{pmatrix} A_{(2)} \\ -A_{(2)} \\ -I \end{pmatrix} x \le \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix} \right\}.$$

Then $c_{(2)}^T x_{(2)}^* = z^T A_{(2)} x_{(2)}^* = b^T z$, and by LP duality, $x_{(2)}^*$ is an optimal solution of $\min\{c_{(2)}^T x_{(2)} : A_{(2)} x_{(2)} = b, x_{(2)} \ge 0\}$. Let $x_{(1)}^* := 0$. Then x^* is an optimal solution of $\min\{c^T x : Ax = b, x \ge 0\}$.

Let $A_{(3)}^T y \le c_{(3)}$ be the subsystem of $A_{(2)}^T y \le c_{(2)}$ corresponding to the positive components of $x_{(2)}^*$. By complementary slackness, it follows that $\{y : A^T y \le c, A_{(3)}^T y = c_{(3)}\}$ is the set of optimal solutions of $\max\{b^T y : A^T y \le c\}$. We can use Proposition 7.7 to find such a solution.

As in the proof of Proposition 7.7, identifying the partition $[A_{(1)}|A_{(2)}]$ of A takes at most poly $\left(n, \log\left(\frac{\Delta(A)}{\delta(A)}\right), \log(p)\right)$ elementary arithmetic operations. Also, note that Δ/δ values for $[A^T| - A^T| - I]$, $[A^T_{(2)}| - A^T_{(2)}| - I]$, $[A|A_{(3)}| - A_{(3)}]$ are all bounded by $\frac{\Delta(A)}{\delta(A)}$. Therefore, the algorithm terminates in at most poly $\left(n, \log\left(\frac{\Delta(A)}{\delta(A)}\right), \log(p)\right)$ elementary arithmetic operations.

The correctness of the above algorithm is guaranteed by the assumption that $p \geq \bar{p}$. Without a prior knowledge of $\left(\frac{\Delta(A)}{\delta(A)}\right)$, we will use the following *"log-squaring trick"*. (Similar tricks have been used before for similar purposes; see [29].) Initially, we can guess n for the value of $\log\left(\frac{\Delta(A)}{\delta(A)}\right)$ and run the above algorithm so that our initial p is roughly $2(2m+n)^{3/2}(2mn+1)2^{2n}$. If the algorithm fails, we replace the current guess by its square, update p, and repeat the algorithm. We also check the output of the above algorithm. If it concludes that (P) (or (D)) is infeasible, we use the corresponding infeasibility certificate to ensure that (P) (or (D)) is indeed infeasible. Similarly, if the algorithm returns a primal-dual "optimal" solution pair, we use complementary slackness conditions to ensure it is indeed optimal. All of these can be done efficiently. If any of the output is false, we again square the most recent guess for $\log\left(\frac{\Delta(A)}{\delta(A)}\right)$, update p, and repeat the algorithm. It is easy to show that after

$$O\left(\log\left(\frac{\log\log\left(\frac{\Delta(A)}{\delta(A)}\right)}{\log(n)}\right)\right)$$

guesses, we have the current guess for p between $\bar{p}$ and $\tilde{p}$. (Here we assume that $\log\log\left(\frac{\Delta(A)}{\delta(A)}\right) \geq 2\log(n)$; otherwise, our first or second guess works and no additional iterations are necessary.) Also, clearly all the guesses for p is at most $\tilde{p}$; moreover, $\log(\tilde{p}) = O\left(\text{poly}\left(n, \log\left(\frac{\Delta(A)}{\delta(A)}\right)\right)\right)$. Therefore, the claimed overall complexity bound is established. □

Note that in the proof of the above theorem, one cannot increase the size of the guess significantly faster than we did, since the sizes of all the integers used by our algorithm must be bounded by a polynomial function of the sizes of the complexity measures we are using.

Theorem 7.9 *We can solve the primal-dual LP problems*

$$(P) : \min\{c^T x : Ax = b, x \geq 0\} \text{ and } (D) : \max\{b^T y : A^T y \leq c\}$$

by utilizing the LP solver subroutine of Subsection 5.1 $O(n^2)$ *times and therefore in at most* poly $\left(n, |\log(\delta_\delta(A))|, \log\left(\frac{\Delta(A)}{\delta(A)}\right)\right)$ *elementary arithmetic operations.*

Proof

We assume that we are given an integer $p \geq \bar{p}$. (We can remove this assumption as in the proof of Theorem 7.8, by applying a *log-squaring trick*.) First we check the feasibility of (P) and (D) using Proposition 7.5 and the underlying algorithm. If any of (P), (D) is infeasible, we have the certificates of such fact and we are done. So, we assume that both (P) and (D) have feasible solutions. Then we apply the proof of Proposition 7.6 to (D) and have the problem

$$(D') : \max\{b^T y : A^T y \leq \lceil c'' \rceil\}.$$

Our theorem in Subsection 5.1 cannot deal with this LP problem (since the objective function of (D') is arbitrary). We form the dual (call it (P')) of (D') and apply the proof of Proposition 7.6 to (P'). Now, the LP problems arising from the applications of Proposition 7.6 to (P') all satisfy the conditions needed in Subsection 5.1 (namely, condition (ii) of the subsection for b and c). So, calling this subroutine $O(n)$ times, as in the proof of Theorem 7.8, we can compute optimal solutions of (P') and (D'). (At some point, during this process, inside the proof of Theorem 7.8, the method in the proof of Proposition 7.7 is used. This requires the LP solver subroutine to be called with data satisfying condition (i)—potentially not satisfying condition (ii)—of Subsection 5.1.) Now, we have an optimal solution of (D') and we can keep applying this technique in using the proof of Theorem 7.8 to solve (P) and (D). This clearly requires no more than $O(n)$ problems of the type (D') to be solved. Since each such problem can be solved with $O(n)$ calls to the LP solver subroutine, the $O(n^2)$ bound follows. □

7.4 Overall Complexity Bounds

Suppose we have an interior-point algorithm satisfying Assumption 2, with an $O\left(n^{\alpha}\left(\log\left(\bar{\chi}(A)\right)\right)^{\beta}\right)$ iteration bound, for some $\alpha \geq 0$, $\beta \geq 0$. Then Theorem 7.8 implies an iteration bound of

$$O\left(n^{1+\alpha}\left[\log\left(\frac{\Delta(A)}{\delta(A)}\right)+\log(n)\right]^{\beta}\left(\log\left(\frac{\log\log\left(\frac{\Delta(A)}{\delta(A)}\right)}{\log(n)}\right)\right)\right).$$

On the other hand, using the methods of Subsection 5.1 and Theorem 7.9, we obtain the iteration bound

$$O\left(n^{2.5}\left(|\log(\delta_\delta(A))| + n\log\left(\frac{\Delta(A)}{\delta(A)}\right)+n\log(n)\right)\left(\log\left(\frac{\log\log\left(\frac{\Delta(A)}{\delta(A)}\right)}{\log(n)}\right)\right)\right).$$

The above bound is not better than Vavasis-Ye's and can be much worse in general. However, in the case that A is totally unimodular, it becomes the same. In this very special case, we can omit the factor of $(\log\log(\bar{\chi}(A)))$ (caused by a *log-squaring* type trick) in the iteration bound of Vavasis-Ye algorithm. See, for instance, Proposition 7.10 and the discussion following it. In the case that A is integral, the bounds can be considered close. See below.

7.5 Integer Data and Network Flow Problems

- Integer Data:

  When the data is integer, $\delta(A) = 1$, $\delta_\delta(A) \geq \frac{1}{\Delta(A)}$ and $\log(\Delta(A)) \leq n\log(n) + \text{size}(A)$. Therefore, we have Tardos' theorem as a special case. Also, in this case it is very easy to get upper bounds (whose sizes are bounded by polynomial functions of the input size) for $\tilde{p}$ so that the multiplicative factor $\left(\log\left(\frac{\log\log\left(\frac{\Delta(A)}{\delta(A)}\right)}{\log(n)}\right)\right)$ in the complexity bound can be removed.

- Totally Unimodular Matrix A:

  Recall that a matrix is *totally unimodular* if all of its square submatrices have determinants -1, 0 or 1. That is, $\delta_\delta(A) = \delta(A) = \Delta(A) = 1$. The following is special case of Proposition 2.14.

Proposition 7.10 *(Ho [11]) Let $A \in \Re^{m \times n}$ be a full row rank totally unimodular matrix. Then $\bar{\chi}(A) \leq \sqrt{mn}$.*

Proof
Take any basis B of A. It is elementary to show that $A_B^{-1}A$ is also totally unimodular. Then for all x such that $\|x\|_2 = 1$,

$$\begin{aligned} \|A_B^{-1}Ax\|_2 &= \sqrt{\sum_{i=1}^{m}\left(\sum_{j=1}^{n}(A_B^{-1}A)_{ij}x_j\right)^2} \leq \sqrt{\sum_{i=1}^{m}\left(\sum_{j=1}^{n}|x_j|\right)^2} \\ &= \sqrt{\sum_{i=1}^{m}(\|x\|_1)^2} \leq \sqrt{mn}, \end{aligned}$$

because $\max_{\|x\|_2=1} \|x\|_1 = \sqrt{n}$ when $x = \frac{1}{\sqrt{n}}e$. Therefore $\bar{\chi}(A) \leq \sqrt{mn}$ by Proposition 2.3. □

In fact we can exhibit a totally unimodular matrix A with $\bar{\chi}(A) = \Theta(\sqrt{mn})$. Consider the complete graph on vertices $\{1, \ldots, m+1\}$, with arcs ij if $i < j$. Let A be its node-arc incidence matrix, with any one row deleted. Then A is a totally unimodular $m \times n$ full row rank matrix, where $n = m(m+1)/2$. It can be easily shown that if we choose $x = e$ and B such that the columns of A_B correspond to a spanning tree that is also a path, i.e., a Hamiltonian path with the correspoding incidence matrix:

$$A_B = \begin{pmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \\ & & & & 1 \end{pmatrix},$$

then $(A_B^{-1}Ax)_j = j(m-j+1)$. Therefore

$$\bar{\chi}(A) \geq \sqrt{\frac{\sum_{j=1}^{m} j^2(m-j+1)^2}{\frac{m(m+1)}{2}}} = \Theta(m^{1.5}) = \Theta(\sqrt{mn}).$$

Therefore, the upper bound proven in Proposition 7.10 is tight up to the order.

Note that we used above, the fact that A_B^{-1} is the all ones upper-triangular matrix. As it is well-known, for every $B \in \mathcal{B}(A)$, there exist permutations of the rows and the columns of A_B such that the resulting matrix is upper-triangular. Since A_B^{-1} is also totally unimodular, it can only have $-1, 0, 1$ entries. Therefore, in this special setting, $B \in \mathcal{B}(A)$, corresponding to Hamiltonian paths, maximize $\|A_B^{-1}\|$.

- Minimum Cost Flow Problems:

Consider the minimum cost flow problem with the constraints $Ax = b$ and $0 \leq x \leq u$, where A is the node-arc incidence matrix of a given directed graph with any one row deleted (so that it has full row rank). By introducing the slack variables v, we convert the constraints into standard equality form:

$$\hat{A}\begin{pmatrix} x \\ v \end{pmatrix} = \begin{pmatrix} b \\ u \end{pmatrix}, \begin{pmatrix} x \\ v \end{pmatrix} \geq 0,$$

where

$$\hat{A} := \begin{pmatrix} A & 0 \\ I & I \end{pmatrix}.$$

This structure arises whenever we convert an upper bounded LP problem to the standard equality form. Vavasis and Ye [29] prove that $\bar{\chi}(\hat{A}) = O(mn)$. Using Propositions 2.7 and 7.10 (and the arguments following that), we have $\bar{\chi}(\hat{A}) = \Theta(\sqrt{mn})$ when A is totally unimodular.

Acknowledgement. Part of this work was done while the second author was visiting the Department of Mathematical and Computing Sciences of Tokyo Institute of Technology on a sabbatical leave from the University of Waterloo (October 1997–March 1998). Support and hospitality of Masakazu Kojima and T.I.T. are gratefully acknowledged.

References

[1] I. Adler and P.A. Beling, Polynomial algorithms for linear programming over the algebraic numbers. *Algorithmica* 12 (1994) 436–457.

[2] L. Blum, F. Cucker, M. Shub and S. Smale, *Complexity and Real Computation.* Springer-Verlag, New York, NY, U.S.A., 1998.

[3] W. Cook, A. M. H. Gerards, A. Schrijver and É. Tardos, Sensitivity theorems in integer programming. *Mathematical Programming A* 34 (1986) 251-264.

[4] I. I. Dikin, On the speed of an iterative process. *Upravlaemye Sistemy* 12 (1974) 54–60.

[5] M. Epelman, Complexity, conditions numbers, and conic linear systems. Ph. D. Thesis, Massachusetts Institute of Technology, 1999.

[6] M. Epelman and R. M. Freund, Pre-conditioners and relations between different measures of conditioning for conic linear systems. *Working Paper OR–344–001*, Operations Research Center, M.I.T., 2000.

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

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

[9] O. Güler, A. J. Hoffman and U. G. Rothblum, Approximations to solutions to systems of linear inequalities. *SIAM J. Matrix Anal. Appl.* 16 (1995) 688–696.

[10] O. Güler ad Y. Ye, Convergence behavior of interior-point algorithms. *Mathematical Programming A* 60 (1993) 215–228.

[11] J. C. K. Ho, Structure of the central path and related complexity issues for linear programming. M. Math. Thesis, University of Waterloo, December 1998.

[12] A. J. Hoffman, On approximate solutions of systems of linear inequalities. *J. Res. Nat. Bur. Stand.* 49 (1952) 263–265.

[13] N. Karmarkar, A new polynomial time algorithm for linear programming. *Combinatorica* 4 (1984) 373–395.

[14] L. Khachiyan, On the complexity of approximating extremal determinants in matrices. *Journal of Complexity* 11 (1995) 138–153.

[15] N. Megiddo, S. Mizuno and T. Tsuchiya, A modified layered-step interior point algorithm for linear programming. *Mathematical Programming A* 82 (1998) 339-355.

[16] S. Mehrotra and Y. Ye, Finding an interior point in the optimal face of linear programs. *Mathematical Programming A* 62 (1993) 497–515.

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

[18] J. Renegar, Linear programming, complexity theory and elementary functional analysis. *Mathematical Programming A* 70 (1995) 279–351.

[19] C. Roos, T. Terlaky and J.-Ph. Vial, *Theory and algorithms for linear optimization. An interior point approach.* Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley and Sons, Ltd., Chichester, 1997.

[20] A. Schrijver, *Theory of Linear and Integer Programming.* John Wiley and Sons, 1986.

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

[22] É. Tardos, A strongly polynomial algorithm to solve combinatorial linear programs. *Operations Research* 34 (1986) 250–256.

[23] É. Tardos, A strongly polynomial mimimum cost circulation algorithm. *Combinatorica* 5 (1985) 247–255.

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

[25] L. Tunçel, Approximating the complexity measure of Vavasis-Ye algorithm is NP-hard. *Mathematical Programming A* 86 (1999) 219–223.

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

[27] L. Tunçel, A pseudo-polynomial complexity analysis for interior-point algorithms. *Dept. of Comb. and Opt. Research Report* 93-16 (June 1993, Revised November 1993).

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

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

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

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

[32] Y. Ye, M. J. Todd and S. Mizuno, An $O(\sqrt{n}L)$-iteration homogeneous and self-dual linear programming algorithm. *Mathematics of Operations Research* 19 (1994) 53-67.

[33] S. Zhang, Global error bounds for convex conic problems. *SIAM J. Optimization* 10 (2000) 836–851.