

SCHOOL OF OPERATIONS RESEARCH  
AND INDUSTRIAL ENGINEERING  
COLLEGE OF ENGINEERING  
CORNELL UNIVERSITY  
ITHACA, NEW YORK

TECHNICAL REPORT NO. 595

November 1983

POLYNOMIAL EXPECTED BEHAVIOR OF A PIVOTING  
ALGORITHM FOR LINEAR COMPLEMENTARITY  
AND LINEAR PROGRAMMING PROBLEMS

by

Michael J. Todd\*

\*Research partially supported by a fellowship from the Alfred P. Sloan  
Foundation and by NSF Grant ECS82-15361.

Typist: Kathy King

POLYNOMIAL EXPECTED BEHAVIOR OF A PIVOTING  
ALGORITHM FOR LINEAR COMPLEMENTARITY  
AND LINEAR PROGRAMMING PROBLEMS

by

Michael J. Todd

School of Operations Research  
and Industrial Engineering  
College of Engineering  
Cornell University  
Ithaca, New York

Abstract

We show that a particular pivoting algorithm, which we call the lexicographic Lemke algorithm, takes an expected number of steps that is bounded by a quadratic in  $n$ , when applied to a random linear complementarity problem of dimension  $n$ . We present two probabilistic models, both requiring some nondegeneracy and sign-invariance properties. The second distribution is concerned with linear complementarity problems that arise from linear programming. In this case we give bounds that are quadratic in the smaller of the two dimensions of the linear programming problem, and independent of the larger. Similar results have been obtained by Adler and Megiddo.

Key Words: Computational Complexity, Average Running Time, Simplex Algorithm, Linear Programming, Linear Complementarity Problem.

Abbreviated title: Polynomial Expected Behavior

Acknowledgement of Support: Research partially funded by a fellowship from the Alfred Sloan Foundation and by NSF Grant ECS82-15361.

# ABSTRACT

We show that a particular pivoting algorithm, which we call the lexicographic Lemke algorithm, takes an expected number of steps that is bounded by a quadratic in  $n$ , when applied to a random linear complementarity problem of dimension  $n$ . We present two probabilistic models, both requiring some nondegeneracy and sign-invariance properties. The second distribution is concerned with linear complementarity problems that arise from linear programming. In this case we give bounds that are quadratic in the smaller of the two dimensions of the linear programming problem, and independent of the larger. Similar results have been obtained by Adler and Megiddo.

## 1. Introduction

This paper presents an analysis of the expected behavior of a particular algorithm (closely related to those of Lemke [9] and Van der Heyden [21]) for the linear complementarity problem: given  $M \in \mathbb{R}^{n \times n}$  and  $q \in \mathbb{R}^n$ , find  $w$  and  $z$  with

$$w = Mz + q \quad (1)$$

$$w \geq 0, z \geq 0, w^T z = 0. \quad (2)$$

We show that the expected number of steps taken is polynomial in  $n$ , under two probabilistic models. First, we make fairly strong nondegeneracy assumptions, together with a sign-invariance property on the distribution generating  $(q, M)$ , and prove that the expected number of steps is at most  $n(n+1)/4$ . Unfortunately, the nondegeneracy assumptions rule out several of the most important applications of the linear complementarity problem. Our second model therefore addresses those problems that arise from linear programming.

Consider the linear programming problem

$$\begin{aligned} \max \quad & c^T x \\ \text{subject to} \quad & Ax \leq b \\ & x \geq 0, \end{aligned} \quad (3)$$

where there are  $m$  constraints in  $p$  variables, so that  $A$  is  $m \times p$ .

This is related to the linear complementarity problem (1)-(2) with

$n = m+p$  and data

$$q = \begin{pmatrix} b \\ -c \end{pmatrix}, \quad M = \begin{pmatrix} 0 & -A \\ A^T & 0 \end{pmatrix}. \quad (4)$$

Any solution to (1)-(2) yields an optimal solution to (3); in addition, the algorithm that we shall analyze either finds such a solution or

demonstrates that (3) is infeasible or unbounded. Thus we assume that a problem of the form (3) is generated probabilistically. Again, we require fairly strong nondegeneracy assumptions for technical reasons, together with a sign-invariance property. We then show that the expected number of steps taken by our algorithm on the resulting linear complementarity problem with data given in (4) is at most  $\min\{(m^2+5m+11)/2, (2p^2+5p+5)/2\}$ . Thus the expected number of steps is bounded by a quadratic function of the minimum of the two dimensions  $m$  and  $p$ , and this bound is independent of the other dimension. The different functions of  $m$  and  $p$  in the bound arise from the asymmetry of the algorithm, which first determines feasibility of the problem, and then whether it has an optimal solution or is unbounded.

Most previous work on the expected behavior of pivoting algorithm has considered (in some guise) Lemke's algorithm, which introduces an artificial vector  $d \in R^n$  and generates a sequence of basic solutions to

$$\begin{aligned} w &= dz_0 + Mz + q \\ w &\geq 0, z_0 \geq 0, z \geq 0 \end{aligned} \tag{5}$$

with  $w^T z = 0$ .

Borgwardt [6,7] considers a linear programming problem of the form  $\max c^T x$ ,  $Ax \leq e$ , where  $e$  here and below denotes a vector of ones of appropriate dimension. He computes a polynomial bound in [7] on the expected number of steps for a parametric method to obtain an optimal solution for the objective  $c^T x$  from the optimal solution for  $\bar{c}^T x$ . This algorithm is very closely related to Lemke's algorithm with  $b \geq 0$ ,  $\bar{c} \leq 0$  and  $d^T = (0, (\bar{c}-c)^T)$ . He also derives a polynomial bound on the expected number of steps for a complete pivoting method that does not start with a feasible vertex. Interestingly, his "phase I" procedure is inductive, and very similar to the algorithm we consider. The only drawback to

Borgwardt's results is that his model always produces feasible problems with a known feasible solution; Haimovich [8] discusses this probabilistic model in comparison with a number of others.

Smale [17,18] addresses the linear programming problem (3) with  $m$  fixed and  $p$  approaching infinity. His algorithm is Dantzig's self-dual parametric method, which is precisely Lemke's algorithm with  $(q,M)$  as in (4) and the artificial vector  $d = e$ . Smale shows that the expected number of steps is bounded by  $C_m(\log p)^{m^2+m}$ . Unfortunately, the constant  $C_m$  is an exponential function of  $m$ . See also Megiddo [12], where it is shown that a bound depending only on  $m$  exists, and Blair [5], for a simplified model and analysis.

Haimovich [8] considers a number of probabilistic models, obtaining linear bounds on the expected number of steps for a parametric objective or parametric right hand side algorithm to solve a linear programming problem. We describe only one of these models, which was also addressed by Adler [1], who independently obtained similar results. Instead of assuming each entry of  $A$ ,  $b$  and  $c$  in (3) is drawn independently from the standard normal distribution, equivalent to Smale's model, Haimovich and Adler assume the distribution is arbitrary except for satisfying certain nondegeneracy assumptions with probability one and a certain sign-invariance property; essentially each constraint is equally likely to be of " $\geq$ " or of " $\leq$ " form. This model has also been considered for generating random polytopes, apparently first by Motzkin [14], and later by Prekopa [15], Adler and Berenguer [2], and May and Smith [10]. The model that we consider is also of this form. Both Haimovich and Adler obtain linear bounds on the expected number of steps of a parametric algorithm, conditioning on the event that the problem has an optimal solution for some value of the parameter. Thus their results are potentially relevant to

the "phase 2" problem, given an initial feasible vertex. The "phase 1" problem of obtaining such a vertex is considered in neither [8] nor [1], although Haimovich (private communication) has indicated some results with quadratic bounds on the expected number of steps, using a variable-dimension approach.

Finally, Megiddo [11] and Saigal [16] have considered the general linear complementarity problem. Megiddo shows that, under a probabilistic model similar to that of Smale, the expected number of basic feasible solutions to (5) with  $w^T z = 0$  grows exponentially with  $n$ , when the artificial vector is  $d = e$ . Thus one might expect Lemke's algorithm to take an exponential number of steps. Saigal adopts a probabilistic model with sign-invariance properties similar to those of Adler and Haimovich. He shows that the expected number of basic feasible solutions to (5) with  $w^T z = 0$  is linear in  $n$  when the artificial vector  $d$  is chosen randomly, with a sign-invariant distribution. On the other hand, when  $d = e$ , the expected number is shown to be between  $(n+2)/2$  and  $(n+6)(\frac{3}{2})^n/6$ . Thus there is a striking contrast between results for a particular positive  $d$  and those for a random  $d$ ; a positive  $d$  allows the algorithm to start from an infeasible basis, and yields possibly exponential growth (Smale [17,18], Megiddo [11,12], Saigal [16]), while a random  $d$  requires a "random" feasible basis to start, and yields polynomial growth (Borgwardt [7], Haimovich [8], Adler [1] and Saigal [16]). Moreover, Saigal's proof demonstrates the difficulty of obtaining a polynomial bound with  $d = e$ ; this vector without sign-invariance intrudes on the calculations of the probability that a particular potential basis is feasible. These probabilities can be bounded so that the potential growth of  $2^n$  is pulled down to about  $1.5^n$  (Saigal) or  $1.15^n$  (Megiddo), but it is still exponential.

Our method can be viewed as Lemke's algorithm with an artificial vector  $d = (\delta^n, \delta^{n-1}, \dots, \delta)^T$  for sufficiently small positive  $\delta$ . Because  $d > 0$ , the method can start from an infeasible basis. But because, during certain iterations of the algorithm, the first few components of  $d$  can be taken as essentially zero, the vector  $d$  does not cause so many problems with estimating the probability that a given basis is feasible--loosely speaking,  $d$  is "more sign-invariant" than  $e$ . Thus we are able to prove a polynomial bound on the expected number of steps.

Section 2 describes in more detail this lexicographic Lemke algorithm, and states how it can be implemented for linear programming problems. In section 3 we analyze the expected behavior of the algorithm on general linear complementarity problems, while section 4 is concerned with those that arise in linear programming. Section 5 discusses an extension to oriented matroids.

Similar results have been obtained independently by Megiddo [13] for the general linear complementarity problem and by Adler and Megiddo [3] for linear programming. Adler and Megiddo have since [4] obtained a quadratic lower bound in the latter case under a strengthened probabilistic model.

## 2. The lexicographic Lemke algorithm

In this section we describe the method to solve the linear complementarity problem that we will analyze in sections 3 and 4. When the linear system (1) is nondegenerate, i.e., every solution has at least  $n$  nonzeros, and all principal minors of  $M$  are positive, this is the algorithm introduced by Van der Heyden [21], and corresponds to Lemke's algorithm with an artificial vector  $d = (\delta^n, \delta^{n-1}, \dots, \delta)^T$ , for all



sufficiently small positive  $\delta$ . Todd [19] showed this equivalence in extending the algorithm to the combinatorial setting of oriented matroids, but Van der Heyden's algorithm can terminate unsuccessfully for some important problems where  $M$  has a nonpositive principal minor. Unfortunately, unless  $b > 0$ , this happens immediately when the linear complementarity problem arises from a linear program.

We shall therefore employ the variant of Lemke's algorithm with  $d = (\delta^n, \delta^{n-1}, \dots, \delta)^T$  for any sufficiently small positive  $\delta$ , as described in [18]. To avoid dealing with a particular  $\delta$ , we shall use equivalent lexicographic rules, to be described below. This algorithm can process arbitrary linear or convex quadratic programming problems, as shown in the context of oriented matroids in [20]. We assume the reader is familiar with Lemke's algorithm in its usual form.

The following notation is convenient. Let  $P$  be a matrix with rows indexed by (elements of the set)  $R$  and columns by  $C$ . Then its transpose, and its inverse if it is square and nonsingular, have rows indexed by  $C$  and columns by  $R$  in the natural manner. For any  $J \subseteq R$  and  $K \subseteq C$ ,  $P_{JK}$  denotes the submatrix of  $P$  with rows indexed by  $J$  and columns by  $K$ . Moreover, we assume these rows and columns retain their original indices in  $P_{JK}$ . We write  $P_{JK}^T$  and  $P_{JK}^{-1}$  (if  $P_{JK}$  is square and nonsingular) for  $(P_{JK})^T$  and  $(P_{JK})^{-1}$  respectively; these are not to be confused with  $(P^T)_{JK}$  or  $(P^{-1})_{JK}$ . It is convenient also to denote by  $e_k$  any unit vector of appropriate dimension and row indices whose unit entry appears in the row indexed  $k$ . Thus  $e_k$  is not a definite vector; its meaning depends on the context. However, if  $k \in K$ , then  $Pe_k$  is the column of  $P$  indexed by  $k$ .

Because rows and columns carry their indices with them, we may write down a matrix in partitioned form in any convenient order. When we use

lexicographic rules, however, we need a specific order; thus we always list the order precisely. A nonzero row vector,  $v^T = (v_i, v_j, \dots, v_k)$ , with its entries explicitly listed, is called lexicographically positive if its first nonzero entry is positive. We say an explicitly listed vector  $u^T$  is lexicographically smaller than an explicitly listed vector  $v^T$  if  $v^T - u^T$  is lexicographically positive. A matrix  $P = (p_i, p_j, \dots, p_k)$ , with its columns explicitly listed, is lexicographically positive if each of its rows, with the order inherited from this list, is lexicographically positive.

Each iteration of the algorithm we use corresponds to a basic feasible solution to (5) with  $w^T z = 0$ . If the algorithm has not yet terminated,  $z_0$  is positive. Let  $J$  index the basic  $w_j$ 's and  $K$  the basic  $z_k$ 's with  $k > 0$ . Then  $J, K$  and some unique index  $\lambda$  partition  $N = \{1, 2, \dots, n\}$ . The usual basis matrix would be  $B' = [I_{NJ}, -d, -M_{NK}]$ , with  $B'$  nonsingular and  $q$  a nonnegative combination of its columns. However, since  $d$  is a vector of powers of some small positive  $\delta$ , it is simpler to switch the roles of  $q$  and  $d$ . Since the combination of the columns of  $B'$  yielding  $q$  uses a positive multiple of  $-d$ , the matrix  $B = [I_{NJ}, -q, -M_{NK}]$  is nonsingular and  $d$  is a nonnegative combination of its columns.

Let us write  $F = [q, M]$ , with rows indexed 1 to  $n$  and columns 0 to  $n$ . Thus  $q$  carries the index 0. Let  $L = K \cup \{\lambda\} = N \setminus J$  and  $H = K \cup \{0\}$ . Then the matrix  $B$  above can be written as

$$B = \begin{bmatrix} I_{JJ} & -F_{JH} \\ 0 & -F_{LH} \end{bmatrix}, \quad (6)$$

and we will call such a matrix a basis matrix if it is nonsingular. In order for this basis matrix to be encountered by the algorithm,  $d = (\delta^n, \delta^{n-1}, \dots, \delta)^T$

must be a nonnegative combination of its columns for all sufficiently small positive  $\delta$ . Thus  $B^{-1}d \geq 0$ , where

$$B^{-1} = \begin{bmatrix} I_{JJ} & -F_{JH}F_{LH}^{-1} \\ 0 & -F_{LH}^{-1} \end{bmatrix}. \quad (7)$$

Since  $B^{-1}d = (B^{-1}e_n)\delta + (B^{-1}e_{n-1})\delta^2 + \dots + (B^{-1}e_1)\delta^n$ , this occurs iff the matrix

$$(B^{-1}e_n, B^{-1}e_{n-1}, \dots, B^{-1}e_1) \quad (8)$$

is lexicographically positive. In this case we say that  $B$  is feasible. Most of our analysis is concerned with estimating the probability that the basis corresponding to a given partition  $N = J \cup K \cup \{\lambda\}$  is feasible. If we do not count as a pivot step obtaining the initial feasible basis, but count the step in which  $z_0$  becomes zero, then the expected number of steps of the algorithm is bounded by the expected number of feasible bases.

Now we describe how the algorithm starts and proceeds from feasible basis to feasible basis. If  $q$  is nonnegative, the algorithm stops immediately with the solution  $w = q, z = 0$ . Thus assume  $q$  has a negative component, and let  $\lambda$  be the index of the first one. Set  $J = N \setminus \{\lambda\}$  and  $K = \emptyset$ . Then it is easy to check that the resulting  $B$  is feasible. We say  $w_\lambda$  has just left the basis.

At a general step we have a feasible basis corresponding to a partition  $N = J \cup K \cup \{\lambda\}$ , and either  $w_\lambda$  or  $z_\lambda$  has just left the basis; we choose its complement (the other in this pair) to enter the basis. The entering

column is then  $a = -Me_\lambda$  if  $z_\lambda$  is entering or  $a = e_\lambda$  if  $w_\lambda$  is entering the basis. We compute the updated vector  $y = B^{-1}a$ . (Note that the components of  $y$ , like the columns of  $B$ , are indexed by  $J \cup H$ .) If  $y \leq 0$  the algorithm terminates, in one of two ways. If  $y_0 = 0$ , then the column  $-q$  is not involved in the linear dependence  $a + B(-y) = 0$ ; the algorithm stops without having solved the linear complementarity problem, and this corresponds to termination on a secondary ray in Lemke's algorithm. On the other hand, if  $y_0 < 0$ , then dividing the equation  $a + B(-y) = 0$  by  $-y_0 > 0$  gives a solution to the linear complementarity problem.

Now suppose  $y$  has a positive component. Then we may make a pivot, introducing the vector  $a$  to replace some column of  $B$  so that the matrix  $(B^{-1}e_n, B^{-1}e_{n-1}, \dots, B^{-1}e_1)$  remains lexicographically positive. Thus we make a minimum ratio test on  $B^{-1}e_n$  and  $y$ , then (if necessary) on  $B^{-1}e_{n-1}$  and  $y$ , and continue until a unique leaving index  $\lambda' \in J \cup K \cup \{0\}$  is determined. If  $\lambda' = 0$  the algorithm has again failed; this corresponds to termination on a secondary ray. Otherwise, the vector  $a$  replaces the column indexed  $\lambda'$  in  $B$ . If  $a = e_\lambda$  we set  $J' = (J \cup \{\lambda\}) \setminus \{\lambda'\}$  and  $K' = K \setminus \{\lambda'\}$ ; if  $a = -Me_\lambda$  we set  $J' = J \setminus \{\lambda'\}$  and  $K' = (K \cup \{\lambda\}) \setminus \{\lambda'\}$ . We now move to the next iteration. If  $\lambda' \in J$  then  $w_{\lambda'}$  has just left the basis, while if  $\lambda' \in K$  then  $z_{\lambda'}$  has just left the basis.

Suppose the linear complementarity problem arises from a linear programming problem (3) as in (4). Then termination on a secondary ray exhibits primal or dual infeasibility. Moreover, in this application of the algorithm it is unnecessary to work with the  $n \times n$  basis matrix  $B$ ; a smaller

simplex tableau can be used. Here we describe the resulting pivot rules--for more detail see [20].

We introduce a variable  $f$  to represent the objective function and slack variables  $u_1, u_2, \dots, u_m$ , so that the initial tableau is

$$f - c^T x = 0$$

$$u + Ax = b.$$

At any iteration, let  $t(v', v)$  denote the entry of the current tableau in the row corresponding to basic variable  $v'$  and the column corresponding to variable  $v$  (or the right hand side if  $v = b$ ). If  $v'$  is nonbasic, we let  $t(v', v) = -1$  if  $v = v'$ , 0 otherwise. While the lexicographic Lemke algorithm is applied just once to the linear complementarity problem, the structure of  $M$  makes it convenient to split the process into two phases.

The phase I procedure finds a feasible tableau with

$$(t(v', b), t(v', u_m), \dots, t(v', u_1))$$

lexicopositive for all basic  $v'$ . At each iteration, we choose first the leaving variable  $v'$  from those with  $t(v', b) < 0$  to minimize lexicographically

$$(t(v', u_m), \dots, t(v', u_1)) / (-t(v', b)).$$

Having chosen  $v'$  to leave, we stop with an indication of infeasibility if  $t(v', v) \geq 0$  for all  $v \neq b$ ; otherwise we choose the entering variable  $v$

from those with  $t(v',v) < 0$  to minimize lexicographically

$$(t(x_p, v), \dots, t(x_1, v)) / t(v', v).$$

Once feasibility is achieved, we proceed to phase II. Here, we maintain lexico-feasibility, i.e.,

$$(t(v', b), t(v', u_m), \dots, t(v', u_1))$$

is lexicopositive for all basic  $v'$ . We choose the entering variable also by a lexicographic rule. Let  $x_r$  be the last basic  $x_j$  (if there are none, let  $r = 0$ ). If  $t(f, v) \geq 0$  for  $v = u_1, \dots, u_m, x_1, \dots, x_{r-1}$ , then choose the entering variable  $v$  to be the first  $x_j$  with  $t(f, x_j) < 0$  -- if there are none, of course, we are already optimal. Otherwise, choose  $v$  from  $u_1, \dots, u_m, x_1, \dots, x_{r-1}$  so that  $t(f, v) < 0$  and  $v$  minimizes lexicographically

$$(t(x_r, v), \dots, t(x_1, v)) / t(f, v).$$

With  $v$  chosen, we stop with an indication of unboundedness if  $t(v', v) \leq 0$  for all basic  $v'$ , and otherwise choose the leaving variable  $v'$  to maintain lexico-feasibility.

It should be apparent that this algorithm can also be implemented in a revised simplex framework, by generating the required parts of the tableau as needed.

### 3. The analysis for general linear complementarity problems

We assume here that the distribution on the data  $(q,M)$  satisfies the following two conditions:

(a) (Nondegeneracy) With probability one, statements (i) and (ii) below hold:

(i) Every square submatrix of  $M$  whose sets of row indices and column indices differ in at most one element is nonsingular.

(ii) There is no almost-complementary solution to  $w = Mz + q$  with fewer than  $n$  components of  $w$  and  $z$  nonzero. (Here almost-complementary means that  $w_i z_i$  is nonzero for at most one index  $i$ .)

(b) (Sign-invariance) The distributions of  $(q,M)$  and  $(Sq, SMS)$  are identical for all sign matrices  $S$ , i.e. all diagonal matrices with all diagonal entries  $\pm 1$ .

Theorem 1. When the probability distribution satisfies assumptions (a) and (b), the expected number of steps of the lexicographic Lemke algorithm applied to the linear complementarity problem (1)-(2) is at most  $n(n+1)/4$ .

We can confine our considerations to problems whose data satisfy the nondegeneracy properties (i) and (ii). Moreover, by (b), it is sufficient to calculate an upper bound on the expected number of steps taken for  $(Sq, SMS)$ , where  $(q,M)$  is fixed with (i), (ii) holding, and all possible sign matrices  $S$  are equally likely. Thus for the rest of this section, we assume this situation.

Consider the basis matrix

$$B = \begin{bmatrix} I_{JJ} & -F_{JH} \\ 0 & -F_{LH} \end{bmatrix} \quad (6)$$

that might arise during the course of the algorithm applied to  $(q, M)$ . Here  $J \cup K \cup \{l\}$  is a partition of  $N$ ,  $L = K \cup \{l\}$  and  $H = K \cup \{0\}$ . Note that (i) implies that all such matrices  $B$  are nonsingular; the proof follows that of lemma 1 below.

Let  $i$  denote the last index in  $L$ .

Lemma 1. All components of  $B^{-1}e_i$  are nonzero.

Proof. Suppose  $e_0^T B^{-1}e_i = 0$ , i.e., that  $q$  is not involved in the expression of  $e_i$  in terms of the columns of  $B$ . Then there is a dependence among the columns of  $[I_{N, J \cup \{i\}}, M_{NK}]$ , so that  $M_{L \setminus \{i\}, K}$  is singular, contradicting (i). Thus  $e_0^T B^{-1}e_i \neq 0$ . Now if  $e_j^T B^{-1}e_i = 0$  for some  $j \in J$ , there would be a dependence among the columns of  $[I_{N, J \cup \{i\} \setminus \{j\}}, M_{NK}]$  and  $q$  with a nonzero weight on the column  $q$ , and this contradicts (ii). A similar contradiction arises if  $e_k^T B^{-1}e_i = 0$  for any  $k \in K$ .

Now note that

$$B^{-1} = \begin{bmatrix} I_{JJ} & -F_{JH}F_{LH}^{-1} \\ 0 & -F_{LH}^{-1} \end{bmatrix}; \quad (7)$$

if  $(q, M)$  is replaced by  $[Sq, SMS]$ , then  $F_{JH}$  is replaced by  $S_{JJ}F_{JH}S_{HH}$  and  $F_{LH}$  by  $S_{LL}F_{LH}S_{HH}$ ; here  $S_{JJ}$  and  $S_{LL}$  are the appropriate submatrices of the sign matrix  $S$ , while

$$S_{HH} \equiv \begin{pmatrix} S_{KK} & 0 \\ 0 & s_{00} = 1 \end{pmatrix}.$$

Thus,  $B^{-1}$  is replaced by

$$\tilde{B}^{-1} \equiv \begin{bmatrix} I_{JJ} & -S_{JJ}F_{JH}F_{LH}^{-1}S_{LL} \\ 0 & -S_{HH}F_{LH}^{-1}S_{LL} \end{bmatrix}. \quad (9)$$



We wish to bound the probability that  $\tilde{B}$  is feasible, i.e. that  $(\tilde{B}^{-1}e_n, \dots, \tilde{B}^{-1}e_1)$  is lexico-positive. To help understand the argument, let us write this matrix with the columns in the listed order, and the rows in the natural order (0 first, and then indices in  $J \cup K$  in increasing order--there is no row indexed  $\varnothing$ ):

$$(\tilde{B}^{-1}e_n, \dots, \tilde{B}^{-1}e_1) = \begin{bmatrix} & & & & \\ & & & & \\ & & & & \\ i+1 \rightarrow & & & 1 & | & \tilde{B}^{-1}e_i & | & ? \\ & & & & 1 & | & & \\ n \rightarrow & 1 & & & & 1 & | & \end{bmatrix}$$

Lemma 2.  $\tilde{B}$  is feasible with probability equal to  $(1/2)^i$  if  $\ell = i$ , and at most  $(1/2)^{i-1}$  if  $\ell < i$ .

Proof. Suppose first  $\ell = i$ . For each  $j \in J$  with  $j > i$ , the first nonzero in row  $j$  of  $(\tilde{B}^{-1}e_n, \dots, \tilde{B}^{-1}e_1)$  is 1, which is always positive. For each  $j \in J$  with  $j < i$ , the first nonzero entry in row  $j$  of  $(\tilde{B}^{-1}e_n, \dots, \tilde{B}^{-1}e_1)$  is  $e_j^T \tilde{B}^{-1}e_i$ , which is  $-e_j^T S_{JJ} F_{JH} F_{LH}^{-1} S_{LL} e_i = -s_{jj} s_{ii} (e_j^T F_{JH} F_{LH}^{-1} e_i)$ . Here we have used lemma 1 to assure that the term in parentheses is nonzero. This entry is positive with probability 1/2, switching sign with  $s_{jj}$ . For each  $k \in K$  the first nonzero entry in row  $k$  of  $(\tilde{B}^{-1}e_n, \dots, \tilde{B}^{-1}e_1)$  is similarly  $e_k^T \tilde{B}^{-1}e_i = -e_k^T S_{HH} F_{LH}^{-1} S_{LL} e_i = -s_{kk} s_{ii} (e_k^T F_{LH}^{-1} e_i)$ , which is again positive with probability 1/2, switching sign with  $s_{kk}$ . Finally, the first nonzero entry in row 0 of  $(\tilde{B}^{-1}e_n, \dots, \tilde{B}^{-1}e_1)$  is  $e_0^T \tilde{B}^{-1}e_i = -e_0^T S_{HH} F_{LH}^{-1} S_{LL} e_i = -s_{ii} (e_0^T F_{LH}^{-1} e_i)$ , which is positive with probability 1/2, switching sign with  $s_{ii}$ . Since the  $s_{jj}$ 's,  $s_{kk}$ 's and  $s_{ii}$  are independent, the probability that  $\tilde{B}$  is feasible is  $(1/2)^i$ .

Next suppose  $\lambda < i$ . Then by a similar argument, the first nonzero entry in row  $j$  of  $(\tilde{B}^{-1}e_n, \dots, \tilde{B}^{-1}e_l)$ , where  $j \in J$  with  $j < i$ , is positive with

probability  $1/2$ , switching sign with  $s_{jj}$ , while if  $j > i$ , it is always positive. The first nonzero entry in row  $k$  of  $(\tilde{B}^{-1}e_n, \dots, \tilde{B}^{-1}e_1)$ , where  $k \in K$ ,  $k \neq i$ , is also positive with probability  $1/2$ , switching sign with  $s_{kk}$ . Finally, the first nonzero entry in row  $0$  of  $(\tilde{B}^{-1}e_n, \dots, \tilde{B}^{-1}e_1)$  is positive with probability  $1/2$ , switching sign with  $s_{ii}$ . (Note that we can say nothing about the first nonzero entry in row  $i$  of  $(\tilde{B}^{-1}e_n, \dots, \tilde{B}^{-1}e_1)$ , which is  $-e_i^T F_{LH}^{-1} e_i$ , independent of  $S$ .) Since all  $s_{jj}$ 's,  $s_{kk}$ 's, and  $s_{ii}$  are independent,  $(\tilde{B}^{-1}e_n, \dots, \tilde{B}^{-1}e_1)$  is lexicographically positive with probability at most  $(1/2)^{i-1}$ .

Proof of Theorem 1. We merely sum all triples  $J, K, \ell$  that may occur with the probabilities given by lemma 2. The index  $i$  can range from  $1$  to  $n$ . For a given  $i$ , if  $\ell = i$ , then  $K$  can be any of the  $2^{i-1}$  subsets of  $\{1, \dots, i-1\}$ . If  $\ell < i$ , then there are  $i-1$  choices for  $\ell$ , and then  $K$  consists of  $i$  together with any of the  $2^{i-2}$  subsets of  $\{1, \dots, i-1\} \setminus \{\ell\}$ . Thus the expected number of steps is at most

$$\sum_{i=1}^n (2^{i-1}(1/2)^i + (i-1)2^{i-2}(1/2)^{i-1}) = \sum_{i=1}^n i/2 = n(n+1)/4.$$

To conclude this section, let us consider a slightly more restrictive model, with assumption (b) strengthened so that  $(q, M)$  and  $(Sq, SMS')$  have the same distribution for any sign matrices  $S$  and  $S'$ . This probabilistic model has been considered by Saigal [16], and also embraces that of Megiddo [11]. The analysis above can be repeated with no changes, except that the probability in lemma 2 becomes  $(1/2)^i$  even when  $\ell < i$ . In the proof of this lemma, we find that the first nonzero entry in row  $i$  of  $(\tilde{B}^{-1}e_n, \dots, \tilde{B}^{-1}e_1)$  is  $-s_{ii}s'_{ii}(e_i^T F_{LH}^{-1} e_i)$ , which switches sign with  $s_{ii}$ , while the first nonzero entry in its  $0$ th row switches sign with  $s'_{ii}$ . The probability

therefore becomes  $(1/2)^i$  in this case also. The resulting bound on the number of iterations is  $n(n+3)/8$ . This bound can also be obtained by a consideration of volumes of simplicial cones, as in Megiddo's analysis, if all entries of  $M$  and of  $q$  are assumed to have the standard normal distribution and to be independent.

We remark that, even in this model where lemma 2 gives the exact probabilities, theorem 1 or its modification only yields an upper bound, since there may be bases which would be feasible if reached, except that the algorithm terminates without encountering them.

Finally, we give our reasons for only requiring the weaker assumption (b). First, it allows us to insist that  $M$  have positive diagonal, or all principal minors positive, or be positive (semi-) definite, with probability one, without violating sign-invariance. This allows considerably more freedom in applying our result to particular classes of problems, while incurring a trivial penalty in the upper bound attained. Most importantly, it allows us to consider problems arising from linear programming, which we address in the next section.

#### 4. The analysis for linear programming

Now we consider the problem

$$\begin{aligned} \max \quad & c^T x \\ \text{subject to} \quad & Ax \leq b \\ & x \geq 0 \end{aligned} \tag{10}$$

where  $A$  is  $m \times p$ , so that there are  $m$  general constraints in  $p$  variables. The resulting linear complementarity problem has  $n = m+p$ , with data

$$q = \begin{pmatrix} b \\ -c \end{pmatrix}, \quad M = \begin{pmatrix} 0 & -A \\ A^T & 0 \end{pmatrix}. \quad (11)$$

In order that our subscripting conventions make sense, we index the rows of  $A$  by  $N_1 = \{1, \dots, m\}$  and its columns by  $N_2 = \{m+1, \dots, n\}$ .

We make the following assumptions on the probability distribution generating the data  $(A, b, c)$  of (10) (and hence the data  $(q, M)$  via (11)).

(c) (Nondegeneracy) With probability one, statements (i)-(iii) below hold:

- (i) Every square submatrix of  $A$  is nonsingular.
- (ii) The linear system  $u + Ax = b$  is nondegenerate, in that every solution has at least  $m$  nonzero variables.
- (iii) The linear system  $A^T y - v = c$  is nondegenerate, in that every solution has at least  $p$  nonzero variables.

(d) (Sign-invariance) The distributions of  $(A, b, c)$  and  $(S_1 A S_2, S_1 b, S_2 c)$  are identical for all sign matrices  $S_1$  and  $S_2$ .

Note that (d) is equivalent to requiring that the induced distribution of  $(q, M)$  be sign-invariant in the sense of (b) in section 3.

Theorem 2. When the probability distribution on  $(A, b, c)$  satisfies assumptions (c) and (d), the expected number of steps of the lexicographic Lemke algorithm for problem (10) is at most  $\min\{(m^2+5m+11)/2, (2p^2+5p+5)/2\}$ .

As in section 3, we consider a particular instance  $(A, b, c)$  of (10) which satisfies (i), (ii) and (iii) of (c), and then compute the expected number of steps when the data are  $S_1 A S_2$ ,  $S_1 b$ , and  $S_2 c$ , where all possible sign matrices  $S_1$  and  $S_2$  are equally likely. For the rest of this section, we assume this situation.

We consider the basis matrix  $B$  as in (6), but now use the special structure of  $M$ .

Let  $J_p = N_p \cap J$ ,  $K_p = N_p \cap K$ ,  $L_p = N_p \cap L$ ,  $p = 1, 2$ . Then we have

$$B = \begin{bmatrix} I_{J_1 J_1} & 0 & 0 & A_{J_1 K_2} & -b_{J_1} \\ 0 & I_{J_2 J_2} & -A_{K_1 J_2}^T & 0 & c_{J_2} \\ 0 & 0 & 0 & A_{L_1 K_2} & -b_{L_1} \\ 0 & 0 & -A_{K_1 L_2}^T & 0 & c_{L_2} \end{bmatrix}. \quad (12)$$

If the partition  $J \cup K \cup \{\lambda\}$  arises in the algorithm, we must have  $B$  nonsingular, and hence the lower "2 x 3" block in (12) is nonsingular. Thus

$A_{L_1 K_2}$  has full column rank, so that  $|L_1| \geq |K_2|$ , and  $[A_{L_1 K_2}, -b_{L_1}]$  has full row rank, so  $|L_1| \leq |K_2| + 1$ . Similarly,  $|K_1| \leq |L_2| \leq |K_1| + 1$ .

Hence either  $|K_1| = |L_2|$  (case 1) or  $|L_1| = |K_2|$  (case 2). In either case we denote by  $i_1$  the last index in  $L_1$ , with  $i_1 = 0$  if  $L_1 = \emptyset$ , and by  $i_2$  the last index in  $L_2$ , with  $i_2 = m$  if  $L_2 = \emptyset$ .

We know that the basis inverse has the form

$$B^{-1} = \begin{bmatrix} I_{JJ} & -F_{JH} F_{LH}^{-1} \\ 0 & -F_{LH}^{-1} \end{bmatrix}$$

but we wish to investigate in detail the zero structure of the columns of  $B^{-1}$  indexed by  $L$ .

Lemma 3. If  $|K_1| = |L_2|$  then  $B$  in (12) is nonsingular. Further, each component of  $B^{-1}e_{i_1}$  is nonzero, and for each  $i \in K_2$ ,  $e_r^T B^{-1}e_i$  is nonzero for each  $r \in J_2 \cup K_1$  and zero for each  $r \in J_1 \cup K_2 \cup \{0\}$ .

(For a pictorial representation of the important structure of  $B^{-1}$ , see the first of the two matrices above lemma 5.)

Proof. For the first part we must show that  $A_{K_1 L_2}^T$  and  $[A_{L_1 K_2}, -b_{L_1}]$  are nonsingular. Since  $|K_1| = |L_2|$ , the first matrix is nonsingular by c(i). If there were a dependence among the first block of columns in the second matrix, then  $A_{K_1 L_2}$  would be singular, a contradiction. Moreover, if there were a dependence involving the final column  $-b_{L_1}$ , then there would be a solution to  $u + Ax = b$  involving only  $u_{J_1}$  and  $x_{K_2}$ ; since  $|J_1| + |K_2| < |N_1 \setminus K_1| + |L_2| = m$ , this would contradict c(ii). Hence  $B$  is nonsingular.

Consider the combination of the columns of  $B$  yielding  $e_{i_1}$ . If this combination did not involve column 0 (the column with the  $b$ 's and  $c$ 's) then  $[A_{L_1 K_2}, e_{i_1}]$  would be singular, so that  $A$  would have a singular square submatrix, contradicting c(i). Next, the combination must include all columns indexed by  $J_1 \cup K_2$ , for otherwise there would be a dependence among all but one column of

$$\begin{bmatrix} I_{J_1 J_1} & A_{J_1 K_2} & -b_{J_1} & 0 \\ 0 & A_{L_1 K_2} & -b_{L_1} & e_{i_1} \end{bmatrix}$$

involving column 0, and this contradicts c(ii). Finally, the combination must include all columns indexed by  $J_2 \cup K_1$ , for otherwise there would be a

dependence among all but one column of

$$\begin{bmatrix} I_{J_2 J_2} & -A_{K_1 J_2}^T & c_{J_2} \\ 0 & -A_{K_1 K_2}^T & c_{L_2} \end{bmatrix}$$

involving column 0, contradicting c(iii).

Now let  $i \in K_2$ , and consider  $B^{-1}e_i$ . From (12), this has the form

$$\begin{bmatrix} 0 \\ -A_{K_1 J_2}^T A_{K_1 L_2}^{-T} e_i \\ -A_{K_1 L_2}^{-T} e_i \\ 0 \\ 0 \end{bmatrix}.$$

Suppose there were a zero component in  $-A_{K_1 J_2}^T A_{K_1 L_2}^{-T} e_i$  or in  $-A_{K_1 L_2}^{-T} e_i$ .

Then  $e_i$  could be expressed as a linear combination of all but one column of

$$\begin{bmatrix} I_{J_2 J_2} & -A_{K_1 J_2}^T \\ 0 & -A_{K_1 L_2}^T \end{bmatrix}$$

and this would imply that  $A$  had a singular square submatrix, contradicting our nondegeneracy assumption c(i). This completes the proof of the lemma.

Similarly we can prove

Lemma 4. If  $|L_1| = |K_2|$  then  $B$  in (12) is nonsingular. Further, each component of  $B^{-1}e_{i_2}$  is nonzero.

(For a pictorial representation of the important structure of  $B^{-1}$ , see the matrix just above lemma 5.)

Now recall that, if  $(q, M)$  is replaced by  $(Sq, SMS)$ , then

$$B^{-1} = \begin{bmatrix} I_{JJ} & -F_{JH}F_{LH}^{-1} \\ 0 & -F_{LH}^{-1} \end{bmatrix}$$

is replaced by

$$\tilde{B}^{-1} = \begin{bmatrix} I_{JJ} & -S_{JJ}F_{JH}F_{LH}^{-1}S_{LL} \\ 0 & -S_{HH}^{-1}F_{LH}^{-1}S_{LL} \end{bmatrix},$$

where, as in section 3,  $s_{00} \equiv 1$ . We now calculate the probability that  $\tilde{B}$  is feasible.

Again, to understand the argument below it may be helpful to see the structure of  $(\tilde{B}^{-1}e_n, \dots, \tilde{B}^{-1}e_1)$  when its rows are ordered naturally. In case 1, lemma 3 shows that we have



$$(\tilde{B}^{-1}e_n, \dots, \tilde{B}^{-1}e_1) = \begin{bmatrix} 0 \rightarrow & 0 & 0 & 0 & 0 & & \\ & & & & & & \\ & & & & & & \\ i_1+1 \rightarrow & & 0 & 0 & & 1 & \\ & & & & & \cdot & \\ n \rightarrow & & & & 1 & & \tilde{B}^{-1}e_{i_1} \quad ? \\ & & x \dots x & & & & \\ & & x \dots x & 1 & & & \\ & & & & & & \\ i_2+1 \rightarrow & & 1 & x \dots x & & & \\ & & & \cdot & \cdot & & \\ n \rightarrow & 1 & \cdot & \cdot & \cdot & & \\ & & & x \dots x & & & \end{bmatrix}$$

with all components of  $\tilde{B}^{-1}e_{i_1}$  nonzero and "x's" marking nonzeros in rows with corresponding unit vectors. Case 2 is simpler: from lemma 4 we have

$$(\tilde{B}^{-1}e_n, \dots, \tilde{B}^{-1}e_1) = \begin{bmatrix} & & & & & & \\ & & & & & & \\ & & & & & & \\ i_2+1 \rightarrow & & & & 1 & & \tilde{B}^{-1}e_{i_2} \quad ? \\ & & & 1 & & & \\ & & 1 & & & & \\ n \rightarrow & 1 & & & & & \end{bmatrix}$$

and  $\tilde{B}^{-1}e_{i_2}$  has all entries nonzero.

Lemma 5. The probability that  $\tilde{B}$  is feasible is at most

$$(1/2)^{i_1+i_2-m} \quad \text{in case 1 if } \ell = i_1 \text{ or } \ell = i_2;$$

$$(1/2)^{i_1+i_2-m-1} \quad \text{in case 1 if } \ell \neq i_1 \text{ and } \ell \neq i_2;$$

$(1/2)^{i_2}$  in case 2 if  $\lambda = i_2$ ; and  
 $(1/2)^{i_2-1}$  in case 2 if  $\lambda \neq i_2$ .

Proof. We consider each row of  $(\tilde{B}^{-1}e_n, \dots, \tilde{B}^{-1}e_1)$  and determine its first nonzero entry. For row  $j$ , where  $j \in J_2$  with  $j > i_2$ , this entry is one and occurs in column  $j$ , for all  $S$ . For row  $r$ , where  $r \in J_2$  with  $r < i_2$  or  $r \in K_1$ , this entry lies in column  $i_2$  and is positive with probability  $1/2$ , switching sign with  $s_{rr}$  and with  $s_{i_2 i_2}$ . The situation for the other rows depends on the case.

Consider first case 1. Then the first nonzero entry in row  $j$ ,  $j \in J_1$ , is one for all  $S$  if  $j > i_1$ , and otherwise occurs in column  $i_1$  and is positive with probability  $1/2$ , switching sign with  $s_{jj}$  and  $s_{i_1 i_1}$ . The first nonzero entry in row  $k$ ,  $k \in K_2$ , occurs in column  $i_1$  and is positive with probability  $1/2$ , switching sign with  $s_{kk}$  and  $s_{i_1 i_1}$ . Finally the first nonzero entry in row  $0$  occurs in column  $i_1$  and is positive with probability  $1/2$ , switching signs with  $s_{i_1 i_1}$ . Suppose  $\lambda = i_1$  or  $\lambda = i_2$ . Then each row  $r$ , where  $r$  lies in  $J_1$  with  $r < i_1$ , in  $J_2$  with  $r < i_2$ , in  $K_1$ , in  $K_2$  or in  $0$ , has first nonzero entry positive with probability  $1/2$ , and all these events are independent. The probability of feasibility is thus  $(1/2)^{i_1+i_2-m}$ . The independence follows since the relevant entry changes sign with  $s_{i_1 i_1}$  ( $r = 0$ ),  $s_{i_2 i_2}$  ( $r = i_1, \lambda = i_2$ ), and otherwise  $s_{rr}$ , and all these  $s$ 's are independent. Now suppose  $\lambda \neq i_1$  and  $\lambda \neq i_2$ . Then the first nonzeros in rows  $i_1$  and  $i_2$  both change sign together, with  $s_{i_1 i_1} s_{i_2 i_2}$ . Hence two of the events above are dependent. We therefore ignore one of them, and conclude that the probability of feasibility is at most  $(1/2)^{i_1+i_2-m-1}$ .

Now consider case 2. Then the first nonzero entry in row  $r$ ,  $r \in J_1 \cup K_2 \cup \{0\}$ ,  $r \neq i_2$ , occurs in column  $i_2$ , and is positive with

probability  $1/2$ , switching sign with  $s_{rr}$  and  $s_{i_2 i_2}$  ( $r \neq 0$ ) or with  $s_{i_2 i_2}$  ( $r = 0$ ). Suppose  $\lambda = i_2$ . Then since all  $s_{rr}$ 's,  $r < i_2$ , and  $s_{i_2 i_2}$ , switch sign independently, the probability that  $\tilde{B}$  is feasible is exactly  $(1/2)^{i_2}$ . However, if  $\lambda \neq i_2$ , then the first nonzero entry in row  $i_2$  occurs in column  $i_2$  and is independent of  $S$ . Thus we must ignore the event that this entry is positive, and we conclude that the probability that  $\tilde{B}$  is feasible is at most  $(1/2)^{i_2-1}$ . This completes the proof of the lemma.

We are now ready for the

Proof of Theorem 2. It is convenient to further subdivide our two cases, according as  $\lambda \in N_1$  (cases  $1_\alpha$  and  $2_\alpha$ ) or  $\lambda \in N_2$  (cases  $1_\beta$  and  $2_\beta$ ). We compute a bound on the expected number of feasible bases for each case, denoted  $E_\rho(1_\alpha)$ , etc. Below we use the convention that all binomial coefficients  $\binom{r}{k}$  are zero for  $r < 0$  or  $k < 0$  or  $k > r$ .

Case  $1_\alpha$ . Here  $i_1$  runs from 1 to  $m$  and  $i_2$  from  $m$  (meaning  $J_2 = N_2$ ) to  $n$ . If  $\lambda = i_1$ ,  $K_1$  can be any subset of  $\{1, \dots, i_1-1\}$  and  $K_2$  any subset of  $\{m+1, \dots, i_2\}$  containing  $i_2$  (if  $i_2 > m$ ), with  $|K_1| = |K_2|$ . Thus the number of choices for  $K_1, K_2$  and  $\lambda = i_1$  is

$$\sum_k \binom{i_1-1}{k} \binom{i_2-m-1}{k-1} = \binom{i_1+i_2-m-2}{i_1-2},$$

except for the exceptional case with  $i_1 = 1$  and  $i_2 = m$ , when the number is 1. (When  $i_1 > 1$  and  $i_2 = m$ , the right hand side above is correct, while the left is not.) To see the identity above, for  $i_1 > 1$ ,  $i_2 > m$ , note that we can make any choice for the  $i_1-2$  elements  $(\{1, \dots, i_1-1\} \setminus K_1) \cup (K_2 \setminus \{i_2\})$  from the  $i_1+i_2-m-2$  elements  $\{1, \dots, i_1-1, m+1, \dots, i_2-1\}$ . We will be using several similar identities below, without elaboration. If  $\lambda < i_1$  (so that  $i_1 \geq 2$  and  $i_2 > m$ ), we can make

$i_1 - 1$  choices for  $\lambda$ , and then  $K_1$  can be any subset of  $\{1, \dots, i_1\} \setminus \{\lambda\}$  containing  $i_1$  and  $K_2$  any subset of  $\{m+1, \dots, i_2\}$  containing  $i_2$ , with  $|K_1| = |K_2|$ . Thus the number of choices for  $K_1, K_2$  and  $\lambda$  is

$$(i_1 - 1) \sum_k \binom{i_1 - 2}{k - 1} \binom{i_2 - m - 1}{k - 1} = (i_1 - 1) \binom{i_1 + i_2 - m - 3}{i_1 - 2}.$$

Using the probabilities from lemma 5, we find that the expected number of feasible bases in case  $1_\alpha$  is

$$E_p(1_\alpha) \leq 1/2 + \sum_{i_1=1}^m \sum_{i_2=m}^n (1/2)^{i_1 + i_2 - m} \left[ \binom{i_1 + i_2 - m - 2}{i_1 - 2} + 2(i_1 - 1) \binom{i_1 + i_2 - m - 3}{i_1 - 2} \right],$$

where the  $1/2$  comes from the exceptional case  $i_1 = 1, i_2 = m$ . The summand for  $i_2 = m$  is just  $(1/2)^{i_1}$ . Thus, removing this part of the sum and writing  $s$  for  $i_1 - 1$  and  $t$  for  $i_2 - m - 1$ , we find

$$E_p(1_\alpha) \leq 3/2 + \sum_{s=0}^{m-1} \sum_{t=0}^{p-1} (1/2)^{s+t+2} \left[ \binom{s+t}{s-1} + 2s \binom{s+t-1}{s-1} \right]. \quad (13)$$

Case  $1_\beta$ . Here  $i_1$  runs from 1 to  $m$  and  $i_2$  from  $m+1$  to  $n$ . If  $\lambda = i_2$ ,  $K_2$  can be any subset of  $\{m+1, \dots, i_2 - 1\}$  and  $K_1$  any subset of  $\{1, \dots, i_1\}$  containing  $i_1$ , with  $|K_1| = |K_2| + 1$ . So the number of choices for  $K_1, K_2$  and  $\lambda = i_2$  is

$$\sum_k \binom{i_1 - 1}{k} \binom{i_2 - m - 1}{k} = \binom{i_1 + i_2 - m - 2}{i_2 - m - 1},$$

If  $\lambda < i_2$ , there are  $i_2 - m - 1$  choices for  $\lambda$ , and then  $K_2$  can be any subset of  $\{m+1, \dots, i_2\} \setminus \{\lambda\}$  containing  $i_2$  and  $K_1$  any subset of  $\{1, \dots, i_1\}$  containing  $i_1$ , with  $|K_1| = |K_2| + 1$ . So the number of choices for  $K_1, K_2$  and  $\lambda < i_2$  is

$$(i_2 - m - 1) \sum_k \binom{i_1 - 1}{k} \binom{i_2 - m - 2}{k - 1} = (i_2 - m - 1) \binom{i_1 + i_2 - m - 3}{i_2 - m - 1},$$

Using the probabilities from lemma 7, we obtain

$$E p(1_\beta) \leq \sum_{i_1=1}^m \sum_{i_2=m+1}^n (1/2)^{i_1 + i_2 - m} \left[ \binom{i_1 + i_2 - m - 2}{i_2 - m - 1} + 2(i_2 - m - 1) \binom{i_1 + i_2 - m - 3}{i_2 - m - 1} \right].$$

Again writing  $s$  for  $i_1 - 1$  and  $t$  for  $i_2 - m - 1$  we find

$$E p(1_\beta) \leq \sum_{s=0}^{m-1} \sum_{t=0}^{p-1} (1/2)^{s+t+2} \left[ \binom{s+t}{t} + 2t \binom{s+t-1}{t} \right]. \quad (14)$$

Case 2 <sub>$\alpha$</sub> . Here  $i_1$  runs from 1 to  $m$  and  $i_2$  from  $m+1$  to  $n$ . If  $\lambda = i_1$ ,  $K_1$  can be any subset of  $\{1, \dots, i_1 - 1\}$  and  $K_2$  any subset of  $\{m+1, \dots, i_2\}$  containing  $i_2$ , with  $|K_1| = |K_2| - 1$ . Thus the number of choices for  $K_1$  and  $K_2$  and  $\lambda = i_1$  is

$$\sum_k \binom{i_1 - 1}{k} \binom{i_2 - m - 1}{k} = \binom{i_1 + i_2 - m - 2}{i_1 - 1},$$

If  $\lambda < i_1$ , we can make  $i_1 - 1$  choices for  $\lambda$  and then  $K_1$  can be any subset of  $\{1, \dots, i_1\} \setminus \{\lambda\}$  containing  $i_1$  and  $K_2$  any subset of

$\{m+1, \dots, i_2\}$  containing  $i_2$ , with  $|K_1| = |K_2| - 1$ . Thus the number of choices for  $\lambda$ ,  $K_1$  and  $K_2$  is

$$(i_1 - 1) \sum_k \binom{i_1 - 2}{k - 1} \binom{i_2 - m - 1}{k} = (i_1 - 1) \binom{i_1 + i_2 - m - 3}{i_1 - 1},$$

Using the probabilities from lemma 5, we find

$$E_p(2_\alpha) \leq \sum_{i_1=1}^m \sum_{i_2=m+1}^n (1/2)^{i_2-1} \left[ \binom{i_1 + i_2 - m - 2}{i_1 - 1} + (i_1 - 1) \binom{i_1 + i_2 - m - 3}{i_1 - 1} \right].$$

Thus, again using  $s$  for  $i_1 - 1$  and  $t$  for  $i_2 - m - 1$ , we get

$$E_p(2_\alpha) \leq \sum_{s=0}^{m-1} \sum_{t=0}^{p-1} (1/2)^{m+t} \left[ \binom{s+t}{s} + s \binom{s+t-1}{s} \right]. \quad (15)$$

Case 2 <sub>$\beta$</sub> . Now  $i_1$  runs from 0 to  $m$  and  $i_2$  from  $m+1$  to  $n$ . If  $\lambda = i_2$ ,  $K_2$  can be any subset of  $\{m+1, \dots, i_2 - 1\}$  and  $K_1$  any subset of  $\{1, \dots, i_1\}$  containing  $i_1$  (if  $i_1 > 0$ ), with  $|K_1| = |K_2|$ . So the number of choices for  $K_1$ ,  $K_2$  and  $\lambda = i_2$  is

$$\sum_k \binom{i_1 - 1}{k - 1} \binom{i_2 - m - 1}{k} = \binom{i_1 + i_2 - m - 2}{i_2 - m - 2},$$

except for the case with  $i_1 = 0$  and  $i_2 = m+1$ , when the number is 1.

(When  $i_1 = 0$  and  $i_2 > m+1$ , the right hand side above is correct, while the left is not.) If  $\lambda < i_2$  (so that  $i_2 \geq m+2$  and  $i_1 > 0$ ), we can make  $i_2 - m - 1$  choices for  $\lambda$ , and then  $K_2$  can be any subset of

$\{m+1, \dots, i_2\} \setminus \{\lambda\}$  containing  $i_2$  and  $K_1$  any subset of  $\{1, \dots, i_1\}$  containing  $i_1$ , with  $|K_1| = |K_2|$ . So the number of choices of  $K_1, K_2$  and  $\lambda < i_2$  is

$$(i_2 - m - 1) \sum_k \binom{i_1 - 1}{k - 1} \binom{i_2 - m - 2}{k - 1} = (i_2 - m - 1) \binom{i_1 + i_2 - m - 3}{i_2 - m - 2},$$

Using the probabilities from lemma 5, we find that

$$E_p(2_\beta) \leq 1/2 + \sum_{i_1=0}^m \sum_{i_2=m+1}^n (1/2)^{i_2} \left[ \binom{i_1 + i_2 - m - 2}{i_2 - m - 2} + 2(i_2 - m - 1) \binom{i_1 + i_2 - m - 3}{i_2 - m - 2} \right].$$

The summand for  $i_1 = 0$  is just  $(1/2)^{i_2}$ . Removing this part of the sum we find

$$E_p(2_\beta) \leq 1 + \sum_{s=0}^{m-1} \sum_{t=0}^{p-1} (1/2)^{m+t+1} \left[ \binom{s+t}{t-1} + 2t \binom{s+t-1}{t-1} \right]. \quad (16)$$

Now let  $E_p$  denote the expected total number of feasible bases. From (13) to (16) we find

$$\begin{aligned} E_p \leq 5/2 + \sum_{s=0}^{m-1} \sum_{t=0}^{p-1} \{ & (1/2)^{s+t+2} \left[ \binom{s+t+1}{s} + 2(s+t) \binom{s+t-1}{s-1} \right] \\ & + (1/2)^{m+t} \left[ \binom{s+t}{s} + (1/2) \binom{s+t}{s+1} + (s+t) \binom{s+t-1}{s} \right] \}. \end{aligned} \quad (17)$$

To complete the proof, we approximate the right hand side of (17) in two ways. We use the following identities:

$$(s+t) \binom{s+t-1}{s-1} = s \binom{s+t}{s} = (t+1) \binom{s+t}{t+1};$$

$$(s+t) \binom{s+t-1}{s} = (s+1) \binom{s+t}{s+1} = t \binom{s+t}{t}; \text{ and}$$

$$\sum_{k=0}^{\infty} \binom{k+\ell}{\ell} (1/2)^k = 2^{\ell+1}.$$

Thus

$$\begin{aligned} E_p &\leq 5/2 + \sum_{s=0}^{m-1} \sum_{t=0}^{p-1} \{ (1/2)^{s+1} \binom{(t+1)+s}{s} (1/2)^{t+1} + (1/2)^{s+1} s \binom{t+s}{s} (1/2)^t \\ &\quad + (1/2)^m \binom{t+s}{s} (1/2)^t + (1/2)^{m+2} \binom{(t-1)+(s+1)}{s+1} (1/2)^{t-1} \\ &\quad + (1/2)^{m+1} (s+1) \binom{(t-1)+(s+1)}{s+1} (1/2)^{t-1} \} \\ &\leq 5/2 + \sum_{s=0}^{m-1} \{ 1 + s + (1/2)^{m-s-1} + (1/2)^{m-s} + (1/2)^{m-s-1} (s+1) \} \\ &\leq 5/2 + \sum_{s=0}^{m-1} \{ 1 + s + 3(1/2)^{m-s} + m(1/2)^{m-s-1} \} \\ &\leq 5/2 + \frac{m(m+1)}{2} + 3 + 2m = (m^2 + 5m + 11)/2. \end{aligned} \tag{18}$$

In addition,



$$\begin{aligned}
 E_p &\leq 5/2 + \sum_{t=0}^{p-1} \sum_{s=0}^{m-1} \{ (1/2)^{t+2} \binom{s+(t+1)}{t+1} (1/2)^s + (1/2)^{t+2} (t+1) \binom{(s-1)+(t+1)}{t+1} (1/2)^{s-1} \\
 &\quad + (1/2)^{t+1} \binom{s+t}{t} (1/2)^s + (1/2)^{t+1} \binom{(s+1)+(t-1)}{t-1} (1/2)^{s+1} + (1/2)^{t+1} t \binom{s+t}{t} (1/2)^s \} \\
 &\leq 5/2 + \sum_{t=0}^{p-1} \{ 1 + (t+1) + 1 + 1/2 + t \} \\
 &= 5/2 + p(p+1) + 3p/2 = (2p^2 + 5p + 5)/2.
 \end{aligned} \tag{19}$$

The inequalities (18) and (19) establish the theorem.

We suspect that, under more restrictive assumptions on the probability model, a quadratic lower bound on the expected number of steps could also be proved.<sup>1</sup> First, in the linear programming context, the expected number of steps taken by the algorithm is exactly  $E_p$ . Every feasible basis of the required form is in fact encountered--this follows from linear complementarity theory. Second, we suspect that the probability bounds given in lemma 5 are close to being tight, under a suitable strengthening of the probability model assumptions--the problem being two nonzeros that switch sign together, or a nonzero that does not switch signs, with changes in  $S$ . Finally, we have not made any gross overestimates in our calculations of bounds on the right hand side of (18).

We also suspect that the nondegeneracy assumption  $(c(i))$  can be somewhat relaxed. It is used in the proof to identify the position of the first nonzero entry in various rows of  $(\tilde{B}^{-1}e_n, \dots, \tilde{B}^{-1}e_1)$ . As long as the first nonzero entry lies in a column indexed by  $k \in K \cup \{l\}$  different from the row index, it will be positive with probability  $1/2$ , switching

---

<sup>1</sup>Such a result has recently been established by Adler and Megiddo [4].

signs with the diagonal entries of  $S$  corresponding to the row and the column index. Thus we have the technical problem of identifying nondegeneracy assumptions sufficient to ensure that a large enough number of rows have their first nonzeros switching signs independently. We have chosen the simplest but most restrictive solution to this technical problem.

To conclude this section, let us consider the linear programming problem in equality form,

$$\begin{aligned} \max \quad & \tilde{c}^T \tilde{x} \\ & \tilde{A} \tilde{x} = \tilde{b} \\ & \tilde{x} \geq 0, \end{aligned} \tag{20}$$

where  $\tilde{A}$  is  $m \times n$ . Let us assume that the probability distribution for  $(\tilde{A}, \tilde{b}, \tilde{c})$  is such that, with probability one, all square  $m \times m$  submatrices of  $\tilde{A}$  are nonsingular, and the linear systems  $\tilde{A} \tilde{x} = \tilde{b}$  and  $\tilde{A}^T y - v = \tilde{c}$  are nondegenerate, and such that  $(\tilde{A}, \tilde{b}, \tilde{c})$  and  $(\tilde{A}\tilde{S}, \tilde{b}, \tilde{S}\tilde{c})$  have the same distribution for any sign matrix  $\tilde{S}$ . Then let us partition  $\tilde{A}$  into  $[\tilde{A}_1, \tilde{A}_2]$ ,  $\tilde{c}^T$  into  $(\tilde{c}_1^T, \tilde{c}_2^T)$ , and  $\tilde{x}^T$  into  $(\tilde{x}_1^T, \tilde{x}_2^T)$ , where  $\tilde{A}_1$  is the first  $m$  columns of  $\tilde{A}$ . Then with probability one,  $\tilde{A}_1$  is nonsingular, and (20) can be rewritten as (10), with  $A = \tilde{A}_1^{-1} \tilde{A}_2$ ,  $b = \tilde{A}_1^{-1} \tilde{b}$ ,  $c = \tilde{c}_2 - \tilde{A}_2^T \tilde{A}_1^{-T} \tilde{c}_1$  and  $x = \tilde{x}_2$ . Moreover, the assumptions placed on the distribution of  $(\tilde{A}, \tilde{b}, \tilde{c})$  imply that (c) and (d) hold for the induced distribution of  $(A, b, c)$ . Thus theorem 2 remains true if we solve (20) by first reformulating it as above in the form (10), i.e., we start with the first  $m$  variables basic.

## 5. Extension to oriented matroids

The analyses in sections 3 and 4 are also valid in the general context of oriented matroids, where the same algorithm can be applied [19,20].

The following brief description assumes familiarity with these papers.

For the linear complementarity problem, we assume that all oriented matroids  $\hat{M}$  on the set  $S \cup T \cup \{r\}$ , with  $S$  a base and with no almost complementary circuit of size smaller than  $n+1$ , are equally likely.

Here  $S$  and  $T$  are disjoint subsets of size  $n$  corresponding to the variables  $w$  and  $z$ , and  $r \notin S \cup T$  corresponds to the right hand side.

The conclusion is again the same quadratic bound on the expected number of steps of the lexicographic Lemke algorithm. The proof is completely

analogous; choosing a sign matrix  $S$  corresponds to reorienting certain elements of the matroid, and all such reorientations are equally likely.

Recall that the analysis of section 3 could be applied when the probability distribution was concentrated on matrices  $M$  having positive diagonal, or positive principal minors, or being symmetric and positive (semi-) definite. Analogous properties can be defined for oriented matroids, and we can obtain a quadratic bound on the expected number of steps, conditioning on the event that one of these properties holds.

For the linear programming problem, we assume that all oriented matroids  $M_X$  on  $U \cup X \cup \{f, g\}$ , with  $U \cup \{f\}$  a base and with no circuit of size smaller than  $m+2$  or cocircuit of size smaller than  $p+2$ , are equally likely. Here  $U$ ,  $X$ ,  $\{f\}$  and  $\{g\}$  are disjoint,  $|U| = m$  and  $|X| = |p|$ ;  $U$  corresponds to the slack variables of (3),  $X$  to the original variables,  $f$  to the objective function and  $g$  to the right hand side. Again we obtain the quadratic bound on theorem 2 on the expected number of steps for the lexicographic Lemke algorithm, and again the proof is completely analogous.

References

1. I. Adler, "The expected number of pivots needed to solve parametric linear programs and the efficiency of the self-dual simplex method," manuscript, Department of Industrial Engineering and Operations Research, University of California, Berkeley, California (May 1983).
2. I. Adler and S.E. Berenguer, "Random linear programs," Operations Research Center Report No. 81-4, University of California, Berkeley, California (1981).
3. I. Adler and N. Megiddo, "A simplex-type algorithm solves linear programs of order  $m \times n$  in only  $O((\min(m,n))^2)$  steps on the average," manuscript, Department of Industrial Engineering and Operations Research, University of California, Berkeley, and Department of Computer Science, Stanford University, Stanford, California (November 1983).
4. I. Adler and N. Megiddo, "A simplex algorithm whose average number of steps is bounded between two quadratic functions of the smaller dimension," manuscript, Department of Industrial Engineering and Operations Research, University of California, Berkeley, and Department of Computer Science, Stanford University, Stanford, California (December 1983).
5. C. Blair, "Random linear programs with many variables and few constraints," Faculty Working Paper No. 946, College of Commerce and Business Administration, University of Illinois at Urbana-Champaign, Illinois (April 1983).
6. K.H. Borgwardt, "Some distribution-independent results about the asymptotic order of the average number of steps of the simplex method," Mathematics of Operations Research 7 (1982) 441-462.
7. K.H. Borgwardt, "The average number of pivot steps required by the simplex-method is polynomial," Zeitschrift fur Operations Research 26 (1982) 157-177.
8. M. Haimovich, "The simplex algorithm is very good!--on the expected number of pivot steps and related properties of random linear programs," manuscript, Columbia University, New York, New York (April 1983).
9. C.E. Lemke, "Bimatrix equilibrium points and mathematical programming," Management Science 11 (1965) 681-689.
10. J.H. May and R.L. Smith, "Random polytopes: their definition, generation, and aggregate properties," Mathematical Programming 24 (1982) 39-54.
11. N. Megiddo, "The probabilistic analysis of Lemke's algorithm for the linear complementarity problem," manuscript, Department of Computer Science, Stanford University, Stanford, California (September 1983).

12. N. Megiddo, "Improved asymptotic analysis of the average number of steps performed by the self-dual simplex algorithm," manuscript, Department of Computer Science, Stanford University, Stanford, California (September 1983).
13. N. Megiddo, "On the expected number of linear complementarity cones intersected by random and semi-random rays," manuscript, Department of Computer Science, Stanford University, Stanford, California (November 1983).
14. T.S. Motzkin, "The probability of solvability of linear inequalities," in: H.A. Antosiewicz, ed., Proceedings of the second symposium in linear programming (National Bureau of Standards and Directorate of Management Analysis, USAF, 1955) pp. 607-611.
15. A. Prekopa, "On the number of vertices of random convex polyhedra," Periodica Mathematica Hungarica 2 (1972) 259-282.
16. R. Saigal, "On some average results for linear complementarity problems," manuscript, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois (August 1983).
17. S. Smale, "On the average number of steps of the simplex method of linear programming, Mathematical Programming 27 (1983) 241-262.
18. S. Smale, "The problem of the average speed of the simplex method," in: A. Bachem, M. Grotschel and B. Korte, eds., Mathematical programming: the state of the art (Springer-Verlag, Berlin-Heidelberg-New York-Tokyo, 1983) pp. 530-539.
19. M.J. Todd, "Complementarity in oriented matroids," to appear in SIAM Journal on Algebraic and Discrete Methods.
20. M.J. Todd, "Linear and quadratic programming in oriented matroids," Technical Report No. 565, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York (March 1983).
21. L. Van der Heyden, "A variable dimension algorithm for the linear complementarity problem," Mathematical Programming 19 (1980) 328-346.