# Probabilistic analysis of the simplex method

**Karl-Heinz Borgwardt**

**Angaben zur Veröffentlichung / Publication details:**

Borgwardt, Karl-Heinz. 1990. "Probabilistic analysis of the simplex method." In *Mathematical Developments Arising from Linear Programming*, edited by Jeffrey C. Lagarias and Michael J. Todd, 1097863. Providence, Rhode Island: American Mathematical Society. https://doi.org/10.1090/conm/114.



# Probabilistic Analysis of the Simplex Method

**Survey-Talk on**

**Probabilistic Analysis of the Simplex Method**

**given by Karl Heinz Borgwardt**

University of Augsburg, Department of Mathematics

8900 Augsburg, West Germany

## Abstract

We are interested in the average number of pivot steps required for solving LPs. It is a common experience shared by almost all Operations Research experts that the Simplex Method is a very good tool for solving linear programming problems occurring in practice. On the other hand we know that all the usual variants of the method have an exponential worst-case behaviour. To understand and to explain that discrepancy had meant a great challenge for mathematicians for over three decades.

The best way to close that gap and to give a formal description of practical efficiency seemed to be a probabilistic analysis of the average behaviour. For that purpose one has to specify a stochastic model defining a (fictive) distribution of linear programming problems occurring in practice. Then a chosen fixed variant of the Simplex Method has to be analyzed under that model.

The talk is a report on three major streams and successful approaches for such an analysis. It is interesting that all these approaches are based on parametric variants.

## 1. Introduction

What I am telling here is a rather old story. Most of the mathematical results mentioned here were derived during the years 1981 – 1984.

Throughout the talk we consider the following type of linear programming problems

$$\begin{array}{ll} \text{maximize} & v^T x \\ \text{subject to} & a_1^T x \leq b^1, \ldots, a_m^T x \leq b^m \\ \text{where } v, x, a_1, \ldots, a_m \in \mathbf{R}^n,\ b \in \mathbf{R}^m \text{ and } m \geq n. \end{array} \tag{1.1}$$

The matrix $A = \begin{pmatrix} a_1^T \\ \vdots \\ a_m^T \end{pmatrix} \in \mathbf{R}^{m \times n}$ consists of the row vectors $a_i^T$ and $b = \begin{pmatrix} b^1 \\ \vdots \\ b^m \end{pmatrix}$.

So we can describe the above inequality system by $Ax \leq b$. The feasible set $X = \{x \in \mathbf{R}^n \mid Ax \leq b\}$ is a convex polyhedron. We prefer that type, because we do not need slack variables, which would increase the dimension of the space and would destroy our ability to imagine what is going on geometrically. The Simplex Method solves such problems by running two Phases.

(1.2) Phase I: Calculation of a vertex $x_0 \in X$. If this is impossible, then STOP.

Phase II: Construction of a sequence $x_0, \ldots, x_s$ of vertices of $X$ such that successive vertices are adjacent and the objective strictly increases from vertex to vertex.
STOP as soon as the optimal vertex $x_s$ is reached or
STOP as soon as the nonexistence of an optimal solution becomes obvious at a vertex $x_s$.

Since we want to study the complexity of that procedure, we are mainly interested in the number $s$. This is the crucial figure of complexity, because it is easy to verify that the single pivot step requires $0(mn)$ arithmetic operations for all usual variants and because Phase I works in a very similar way — only a modification of the problem is required — to Phase II. In order to simplify our considerations, it makes sense to concentrate on Phase II for the moment. We will return to questions concerning the complete algorithm later. So we are going to study the value of $s$ (the number of vertex exchanges in Phase II), which turns out to coincide with the number of pivot steps in the case of nondegenerate problems. Nondegeneracy will be assumed throughout our analysis.

Our definition of Phase II is not yet complete, because we have to give a rule for choosing the successor vertex if there are more than one adjacent and improving vertices. Since we are interested in a probabilistic analysis, we want to answer the following question:

How great is $s$ "on the average" for fixed dimensions $m$ and $n$?

In order to state clearly what we mean by "on the average" we have to introduce a stochastic model for the real-world distribution of the problems.

It is clear now that the average number of steps will depend highly on

- the variant which is used,
- the chosen stochastic model.

Throughout the following we shall rely on the following **nondegeneracy assumptions** under the stochastic models discussed here, which do not influence our calculated

expectation values:

(1.3) Every submatrix of $(A, b)$ and of $\binom{A}{v^T}$ is of full rank.

One of the consequences of (1.3) is the fact that every feasible $X$ is pointed, because there cannot be a straight line which is feasible everywhere. For understanding the geometry of the problems, we introduce the following two terms.

**Basic solution $x_\Delta$:** This is a solution of a system of equations

$$\begin{aligned} a_{\Delta^1}^T x &= b^{\Delta^1} \\ &\vdots \qquad \text{with } \Delta = \{\Delta^1, \ldots, \Delta^n\} \subset \{1, \ldots, m\} \text{ and } \Delta^i < \Delta^j \text{ for } i < j. \\ a_{\Delta^n}^T x &= b^{\Delta^n} \end{aligned}$$

A basic solution becomes a vertex, if the remaining restrictions $a_i^T x \leq b^i$ $(i = 1, \ldots, m;$ $i \notin \Delta)$ are satisfied at $x_\Delta$.

**Line:** This is a solution set of a system of equations

$$\begin{aligned} a_{\Delta^1}^T x &= b^{\Delta^1} \\ &\vdots \\ a_{\Delta^{n-1}}^T x &= b^{\Delta^{n-1}}. \end{aligned}$$

A line may contain an edge of $X$, if there is a section satisfying all the restrictions.

It is a consequence of nondegeneracy, that these systems have unique solutions, that every vertex of $X$ corresponds to a set of exactly $n$ active restrictions, that every pivot step leads to a new vertex and that an existing optimal point will be unique.

In the following, we will describe three successful approaches which are based on different stochastic models. We shall emphasize on the reasons why the analysis could be done rather than showing how it was done.

## 2. Smale's Work

In 1982 Steve Smale [10] did an analysis where the linear programming problem was imbedded into the more general linear complementarity problem. The linear programming problem he dealt with was of the type

$$\begin{aligned} &\text{maximize} && v^T x \\ &\text{subject to} && a_1^T x \leq b^1, \ldots, a_m^T x \leq b^m \text{ and } x \geq 0 \\ &\text{where } x, v, a_1, \ldots, a_m \in \mathbf{R}^n,\ b \in \mathbf{R}^m. \end{aligned} \tag{2.1}$$

This problem was regarded as a special case of the linear complementarity problem

$$\begin{aligned} &\text{find } w, z \text{ such that for given } q, M \\ &w - Mz = q,\ w^T z = 0,\ w \geq 0,\ z \geq 0 \\ &\text{where } w, z, q \in \mathbf{R}^{m+n},\ M \in \mathbf{R}^{(m+n)\times(m+n)}. \end{aligned} \tag{2.2}$$

When we specialize $M$ to $\begin{pmatrix} 0 & -A \\ A^T & 0 \end{pmatrix}$ and $q$ to $\begin{pmatrix} b \\ -v \end{pmatrix}$, a solution of (2.2) will give the solution of (2.1) and of the corresponding dual problem.

Geometrically, the task is to represent $q$ as a conical combination of $m+n$ columns (alternatively taken out from $-M$ or the identity matrix $I$), i. e.,

$$q = Iw - Mz \text{ and } w^T z = 0,\ w \geq 0, z \geq 0.$$

We solve the problem by Lemke's algorithm. Usually, we do not know the representation for $q$, but for $e = \mathbf{1}$ (the vector of $m+n$ ones). There we have $e = Ie - M0$.

Now imagine a straight line between $e$ and $q$ and parametrize a movement from $e$ to $q$ by setting

$$q_\lambda = \lambda q + (1-\lambda)e \text{ for } \lambda \in [0,1]. \tag{2.3}$$

We know the conical representation for $\lambda = 0$, but we want it for $\lambda = 1$. When we move with increasing $\lambda$, we try to maintain $q_\lambda = w_\lambda - Mz_\lambda$, $w_\lambda \geq 0$, $z_\lambda \geq 0$ by adapting $w_\lambda$ and $z_\lambda$. Everytime the set of positive $w$-entries has to change, we enter a new cone (generated by $m+n$ column vectors out of $-M$ or $I$). We can calculate the corresponding change of representation by a pivot step in Lemke's algorithm. Consequently, we get an upper bound for the number of pivot steps in Lemke's algorithm by counting the number of cones which had been intersected by $[e, q]$.

The stochastic assumptions used in the analysis of Lemke's algorithm were as follows.

(2.4)
1) Let the distribution of $(A, b, v)$ be absolutely continuous.
2) Let $A, b, v$ be distributed independently, and let particularly the columns of $A$ be distributed independently.
3) Let the probability measure of $(A, b, v)$ be invariant under coordinate permutations (in columns of $A$ resp. $b$).

Smale's analysis came to the following result about $E_{m,n}(s^L)$, the expected number of pivot steps in Lemke's algorithm for $(m, n)$-problems.

**(2.5) Theorem.** *$E_{m,n}(s^L) \leq C(n)\ (1 + ln(m+1))^{n(n+1)}$ under conditions 1) – 3) where $C(n)$ is a constant depending on $n$ but not on $m$.*

This result showed that $E_{m,n}(s^L)$ is polynomial in $m$, but the dependence in $n$ could still be exponential.

## 3. The shadow-vertex-algorithm and other parametric variants

Lemke's algorithm can be interpreted as an application of the principle of parametric programming. Even more direct is the use of parametric variants under the next two approaches.

For that reason we explain how the variants work. As stated before, we consider Phase I to be done and a start vertex $x_0$ to be given. Now $x_0$ itself is an optimal vertex with respect to a certain objective direction $u \in \mathbf{R}^n$.

Having the optimal vertex with respect to $u^T x$, we desire to have that for $v^T x$. Now we project the polyhedron $X$ onto the two-dimensional plane spanned by $u$ and $v$ and we obtain a two-dimensional polyhedron $\Gamma(X, u, v)$. In $\Gamma(X, u, v)$, some of the vertices of $X$ have disappeared in the interior, the others have become vertices of $\Gamma(X, u, v)$ (the shadow of $X$). The latter vertices of $X$ will be called shadow-vertices. As a result of their definition, the vertices $x_0$ and $x_s$ are such shadow-vertices. And there is a Simplex-Path connecting both touching only shadow-vertices. This path is even unique as a consequence of nondegeneracy if we run through the smaller angle between $u$ and $v$. Hence the number of shadow-vertices $S$ is an upper bound for the number of pivot steps for realizing the **shadow-vertex path**. This realization is done with little effort, we only need a representation for a second objective $u$ in the Simplex-Tableau, as we had it for $v$ before. Running along the Simplex-Path from $x_0$ to $x_s$ can now be regarded as walking through the vertices maximizing an objective $(1 - \lambda)u + \lambda v$ with $\lambda \in [0, 1]$ while $\lambda$ increases.

And also very useful for the following is the conception of **cooptimality**. A point $y \in X$ is called $(u, v)$-cooptimal, if $y$ is $u$-optimal ($u^T y$ is maximal) among all points of $X \cap \{x \mid v^T x = v^T y\}$. That means that $u^T y$ cannot be improved without changing $v^T x$.

Obviously, every cooptimal vertex is optimal with respect to an objective $(1 - \lambda)u + \lambda v$ with $\lambda \in [0, 1]$ and it is clear, that every cooptimal edge lies on the boundary of the shadow of $X$ and that every cooptimal vertex is a shadow vertex. Hence we have

$$s \leq C \leq S \tag{3.1}$$

($C$ = number of $(u, v)$-cooptimal vertices, $S$ = number of shadow vertices.)

## 4. Results under the Sign-Invariance-Model

The following stochastic model — the Sign-Invariance-Model — could be analyzed and lead to a much more optimistic judgement on the average behaviour.

(4.1) Let the distribution of $(A, b, v)$ and of $(S_1AS_2, S_1b, S_2v)$ be identical for all sign matrices $S_1 \in \mathbf{R}^{m\times m}$ and $S_2 \in \mathbf{R}^{n\times n}$.
Here a sign matrix is a diagonal matrix with $+1$ or $-1$ in the diagonal entries.

In order to get an impression on the quality of a Phase II-algorithm under that model, it suffices to consider a relaxed version of sign-invariance, the so-called flipping model. Here we consider again our data set for the linear programming problem

(4.2) maximize $v^Tx$
subject to $a_1^Tx\ b^1, \ldots, a_m^Tx\ b^m$
but the omitted directions of the inequalities ($\leq$ or $\geq$) are regarded as random variables, which are chosen independently for the $m$ inequalities. $\leq$ and $\geq$ shall have probability $\frac{1}{2}$.

By the way, we generate $2^m$ problems out of one data-set. The idea is now to solve all these problems, to count the pivot steps and to divide by the number of problems. The feasible regions of these problems will be called cells. Note that cells may be empty.

The following result of Haimovich (1983) [7] (and Adler [12]) gives an impression, why Phase II runs very quickly.

**(4.3) Theorem.** *$E_{m,n}$ ($C$ | a cooptimal path exists) $\leq n\frac{m-n+2}{m+1}$ under the "flipping model".*

**Sketch of a proof.**

Only $\binom{m}{0}+\binom{m}{1}+\ldots+\binom{m}{n}$ problems out of the $2^m$ generated have nonempty cells (partition of space). There are $\binom{m}{n}$ basic solutions and $\binom{m}{n-1}$ lines.
Every cooptimal path ends up in a basic solution or in a ray which is a part of a line.
Every basic solution is $v$-optimal in exactly one cell.
Every edge (line segment) is cooptimal in exactly one cell.
Every cell has at most one cooptimal path.
Every line is divided by the remaining $m-n+1$ restrictions into $m-n+2$ segments. One of them gives the ray mentioned above.

Hence the number of cells with cooptimal path is $\binom{m}{n} + \binom{m}{n-1}$ and the number of traversed segments (edges) for running through all cooptimal paths is $\binom{m}{n-1}(m-n+2)$.

In 1983 three different and independently written papers showed that the whole work (including Phase I and Phase II) could be done in quadratic time (on the average). The authors were 1) Todd; 2) Adler/Megiddo; 3) Adler/Karp/Shamir [11], [2], [1].

**(4.4) Theorem.** *$E_{m,n}(s^C) \leq 2(n+1)^2$ under the Sign-Invariance-Model* (4.1) *where $s^C$ is the number of pivot steps required for the complete method.*

This result was derived in the papers of Todd and Adler/Megiddo by a very skilfull choice of the start vector in Lemke's algorithm (compare Section 2). They chose a vector $(\delta, \delta^2, \delta^3, \ldots, \delta^{m+n}) \in \mathbf{R}^{m+n}$ with $\delta > 0$ arbitrarily small to have an initial representation. So they realized the so-called "Lexicographic-Lemke-Algorithm". The advantage of this algorithm comes from the fact that we have good control, on which sides a cone is entered and left.

Closer to our geometrical interpretation is the method used in the paper of Adler/ Karp/Shamir, the so called **"Lexicographic Constraint-By-Constraint Method"**. This algorithm works directly in the original space and works as follows:

**Initialization:** Determine a basic solution $x_\Delta$ with $\Delta = \{1, \ldots, n\}$ and choose $u$ (the objective making $x_\Delta$ optimal) as $u = \delta a_1 + \delta^2 a_2 + \ldots + \delta^n a_n$ with $\delta > 0$ arbitrarily small. Then $x_\Delta$ is an optimal vertex on $X^{(n)} = \{x \mid a_1^T x \leq b^1, \ldots, a_n^T x \leq b^n\}$.

**Typical Step:** Start at $\overline{x}$, the optimal vertex for $u^T x$ on $X^{(n+k-1)} = \{x \mid a_1^T x \leq b^1, \ldots, a_{n+k-1}^T x \leq b^{n+k-1}\}$. If $a_{n+k}^T \overline{x} \leq b^{n+k}$ then $\overline{x}$ is also $u$-optimal on $X^{(n+k)} = \{x \mid a_1^T x \leq b^1, \ldots, a_{n+k}^T x \leq b^{n+k}\}$. We proceed to the next Step.

If $a_{n+k}^T \overline{x} > b^{n+k}$, then use the shadow vertex algorithm to minimize $a_{n+k}^T x$ on $X^{(n+k-1)}$ until $a_{n+k}^T x \leq b^{n+k}$ is achieved. All path-points are $(u, -a_{n+k})$ cooptimal, hence the first point on our path with $a_{n+k}^T x = b^{n+k}$ maximizes $u^T x$ on $X^{(k+n)}$. If we achieve $a_{n+k}^T x \leq b^{n+k}$ we proceed to the next step.

If it is impossible to achieve $a_{k+n}^T x \leq b^{k+n}$, then the original problem was infeasible. We STOP.

**Final Step:** We start at $\overline{x}$ which is $u$-maximal on $X^{(m)} = X$. We apply the shadow-vertex algorithm to maximize $v^T x$ on $X$.

This amounts to an $m - n + 1$-fold application of the shadow-vertex-algorithm. But the very tricky choice of $u$ at the initial step enabled the authors to show that the expected number of pivot steps is not only $0(mn)$ but $0(n^2)$.

In the paper of Adler/Megiddo, it was shown that slight additional conditions on the distributions of the entries of $A$ suffice to establish also a lower bound on the average behaviour such as $Cn^2$ with $C > 0$.

After these three papers had been written, Megiddo observed that although the used algorithms were different, they realized exactly the same Simplex-Paths [13].

As observed often in probabilistic analysis of algorithms, one must be very careful in the interpretation of the results and be aware of the influence of the stochastic assumptions.

This is also a very important point for the Sign-Invariance-Model.

We have already mentioned that many cells will be empty. The quotient

$$\frac{\text{number of feasible problems}}{\text{number of generated problems}} = \frac{\binom{m}{0} + \binom{m}{1} + \ldots + \binom{m}{n}}{2^m}$$

tends to 0 for $m \to \infty$ and $n$ fixed (asymptotically).

Also the expected number of vertices per generate problem

$$\frac{2^n \binom{m}{n}}{2^m}$$

tends to 0 asymptotically.

Only **conditioning on feasible problems** can avoid averaging over a lot of infeasible and easy problems. Here we have as expected number of vertices per nonempty cell

$$\frac{2^n \binom{m}{n}}{\binom{m}{0} + \ldots + \binom{m}{n}} \to 2^n \text{ asymptotically.}$$

But the most important influence comes from the **average redundancy rate** (share of redundant constraints). Here we have (conditioned on **nonempty cells**).

$$\text{Average Redundancy Rate} = \frac{\binom{m-1}{n}}{\binom{m}{0} + \ldots + \binom{m}{n}} \to 1 \text{ asymptotically.}$$

Adler and Megiddo showed in [2] that for $m = 0(2n)$ $E_{m,n}(s^C \mid \text{problem is feasible and has an optimal solution}) = 0(n^{2.5})$. (Compare also [1] and [9].) The average redundancy rate for $m \leq 2n$ is $\leq \frac{1}{2}$. This shows that the results are reliable and realistic for $m = 0(2n)$, but that the model makes problems easy when $m \gg n$. And this holds even in the absence of infeasible problems.

So we must be aware of the danger, that the results reflect the model rather than the quality of the algorithm.

## 5. Results under the Rotation-Symmetry-Model

The danger of having a dramatic share of infeasible problems or of redundant constraints could be avoided in the stochastic model analyzed by the author in several papers [3], [4], [5]. We deal with problems of the type

$$\begin{array}{ll} \text{maximize} & v^T x \\ \text{subject to} & a_1^T x \leq 1, \ldots, a_m^T x \leq 1 \\ \text{were } v, x, a_1, \ldots, a_m \in \mathbf{R}^n, & m \geq n. \end{array} \tag{5.1}$$

Here the origin is guaranteed to be feasible in any case.

The stochastic model is

(5.2) Let $a_1, \ldots, a_m$, $v$ and $u$ (the direction determining the start vertex) be distributed

- independently
- identically
- symmetrically under rotations of $\mathbf{R}^n$.

Note that (5.1) can easily be generalized to having positive right sides $b^i > 0$, when these right sides are independently and identically distributed over $(0, \infty)$. ($b^i$ independent of all $a_1, \ldots, a_m$). Normalization of the inequalities will then lead back to (5.2).

Now remember our inequality $s \leq S$ when we decide to apply the shadow-vertex-algorithm in Phase II. Candidates for becoming shadow-vertices are only the $\binom{m}{n}$ basic solutions $x_\Delta$ solving $\begin{array}{l} a_{\Delta^1}^T x = 1 \\ \quad \vdots \\ a_{\Delta^n}^T x = 1 \end{array}$.

We observe the following one-to-one correspondence.

$$x_\Delta \longleftrightarrow \Delta = \{\Delta^1, \ldots, \Delta^n\} \longleftrightarrow \mathrm{CH}(a_{\Delta^1}, \ldots, a_{\Delta^n}) \text{ (convex hull).} \tag{5.3}$$

In addition to the primal polyhedron $X = \{x \mid Ax \leq b\}$ we introduce the "polar polyhedron" $Y = \mathrm{CH}(0, a_1, \ldots, a_m)$.

The following equivalencies enable us to derive the average number of shadow vertices directly from the input data.

**(5.4) Lemma.**

*1) $x_\Delta$ is a vertex of $X \iff \mathrm{CH}(a_{\Delta^1}, \ldots, a_{\Delta^n})$ is a facet of $Y$.*

*2) Let $x_\Delta$ be a vertex of $X$. Then*

$$x_\Delta \text{ is a shadow vertex} \iff \mathrm{CH}(a_{\Delta^1}, \ldots, a_{\Delta^n}) \cap \mathrm{Span}(u, v) \neq \emptyset.$$

By the addition theorem for expectation values we get the following integral expression for $E_{m,n}(S)$.

(5.5) $E_{m,n}(S) = \binom{m}{n} \int\limits_{\mathbf{R}^n} \cdots \int\limits_{\mathbf{R}^n} P(\mathrm{CH}(a_1,\ldots,a_n) \text{ defines a facet of } Y$
and is intersected by $\mathrm{Span}(u,v))dF(a_1)\ldots dF(a_n)$.

Here $F$ denotes the distribution function under consideration.

It took several years to evaluate this integral. From 1977 till 1984 we could derive some asymptotic results ($m \to \infty$, $n$ fixed) under special distributions for the $a_i$.

**(5.6) Theorem** (1977 – 1984). *There is a function $\varepsilon(m,n)$ with $\varepsilon(m,n) \to 0$ asymptotically such that for*

**Gaussian distribution on $\mathbf{R}^n$:**

$$E_{m,n}(S) \le \sqrt{\ln m}\, n^{\frac{3}{2}}\, 2^{\frac{3}{2}}\, \frac{20}{9}(1+\varepsilon(m,n))$$
$$E_{m,n}(S) \ge \sqrt{\ln m}\, n^{\frac{3}{2}}\, 2^{\frac{3}{2}}\, \frac{3}{5}(1-\varepsilon(m,n)).$$

**Uniform distribution on unit ball:**

$$E_{m,n}(S) \le m^{\frac{1}{n+1}}\, n^2\, 2\Big(1+\frac{1}{\sqrt{2}}\Big)(1+\varepsilon(m,n))$$
$$E_{m,n}(S) \ge m^{\frac{1}{n+1}}\, n^2\, 2\frac{1}{3}(1-\varepsilon(m,n)).$$

**Uniform distribution on unit sphere:**

$$E_{m,n}(S) \le m^{\frac{1}{n-1}}\, n^2\, 2\Big(1+\frac{1}{\sqrt{2}}\Big)(1+\varepsilon(m,n))$$
$$E_{m,n}(S) \ge m^{\frac{1}{n-1}}\, n^2\, 2\frac{1}{7}(1-\varepsilon(m,n)).$$

**General distributions with bounded support:**

$$E_{m,n}(S) \le m^{\frac{1}{n-1}}\, n^2\, \sqrt{2\pi}\, (1+\varepsilon_F(m,n)).$$

**(5.7) Theorem** (Borgwardt 1979–1984). *For $m \to \infty$, $n$ fixed.*

1) *There exist distributions according to* (5.2) *such that $E_{m,n}(S)$ converges to a constant $C(n)$ in $m$.*
*Example: radial distribution* $\tilde{F}(r) = \begin{cases} 0 & r < 1 \\ 1-\frac{1}{r} & r \ge 1 \end{cases}$.
2) *There is a distribution with the above property and $C(n) \le n^{\frac{5}{2}}$.*

*Example:* $\tilde{F}(r) = \begin{cases} 0 & r < 1 \\ 1 - \frac{1}{r^{n^2}} & r \geq 1 \end{cases}$.

The dramatic difference in the order of growth for different distributions results from different redundancy rates. We can realize a redundancy rate of 0 by choosing the uniform distribution on the sphere, whereas the examples in (5.7) establish redundancy rates very close to 1. So it is not astonishing, that these examples simulate the effect observed at the Sign-Invariance-Model. Asymptotically, the influence of $m$ disappears. And this holds here while every instance is feasible and almost all are bounded!

In 1981 I succeeded in proving polynomiality in both dimensions.

**(5.8) Theorem** (1981). *For all distributions according to* (5.2)

$$E_{m,n}(S) \leq e\pi(\frac{\pi}{2} + \frac{1}{e})m^{\frac{1}{n-1}}n^3.$$

We now come back to the inclusion of Phase I. We could introduce a (rather lengthy) method for doing the task of Phase I and Phase II by several applications of the shadow-vertex-algorithm. This enables us to exploit our known results, since in every application the stochastic requirements (5.2) are still satisfied. By the way all our trouble with the auxiliary objective $u^T x$ disappear. We do not need it any longer.

The complete method works as follows.

**Initialization:**

1) Starting from the origin, find a vertex of the polyhedron $X^{(2)} = \{x \mid Ax \leq b, x^3 = \ldots = x^n = 0\}$ (two-dimensional).

   This is easily done by running along a coordinate axis, until a restricting hyperplane is hit. In case the axis does not hit, we invert the search direction. If we still do not hit, the original problem is unbounded. We can STOP.

   If a hitting point is found, it is located on an edge of $X^{(2)}$. Find a vertex of that edge. If the edge does not have vertices, the original problem is unbounded, we can STOP.

2) Apply the shadow-vertex-algorithm starting from the given vertex in order to maximize $v^T x$ on $X^{(2)}$.

   If $v^T x$ is unbounded, we can STOP (unbounded on $X$, too).

**Typical Step:**

Use the solution point $(x^1, \ldots, k^{k-1}, 0, \ldots, 0)^T$ of $X^{(k-1)}$. It is located on an edge of $X^{(k)}$.

1) Find an adjacent vertex in $X^{(k)}$ to that edge.
2) Apply the shadow vertex algorithm by using $e_k^T x$ and $v^T x$ as objective and auxiliary objective.

If $v^T x$ is unbounded on $X^{(k)}$ we can STOP.

3) If $k < n$ we go to the next step and use the solution point with $k := k + 1$.
If $k = n$ we STOP.

This method allowed evaluation, since it was in principle an $n - 1$ fold application of the shadow-vertex-algorithm. Our main result was

**(5.9) Theorem** (1981/1984). *For all distributions satisfying* (5.2) *our complete method solves the LP in not more than*

$$m^{\frac{1}{n-1}}(n+1)^2 n^2 \frac{e\pi}{4}\left(\frac{\pi}{2} + \frac{1}{e}\right)$$

*steps on the average.*

By this method we were also able to verify that result for problems including sign constraints $x \geq 0$.

So long our result had a certain drawback. We had to know a feasible point. But how should one handle problems of the form

$$\begin{array}{ll} \text{maximize} & v^T x \\ \text{subject to} & a_1^T x \leq b^1, \ldots, a_m^T x \leq b^m \\ \text{with arbitrary } b^i \text{ (not necessarily positive)?} & \end{array} \tag{5.10}$$

We call the polyhedron of (5.10) $P_n$.
We reformulate our restrictions

$$\begin{array}{c} a_1^T x \leq b \\ \vdots \\ a_m^T x \leq b^m \end{array} \longleftrightarrow \begin{array}{l} \text{set } b^i = 1 - \tilde{b}^i \\ \text{resp. } \tilde{b}^i = 1 - b^i \end{array} \begin{array}{c} a_1^T x \leq 1 - \tilde{b}^1 \\ \vdots \\ a_m^T x \leq 1 - \tilde{b}^m \end{array} \longleftrightarrow \begin{array}{c} a_1^T x + \tilde{b}^1 \leq 1 \\ \vdots \\ a_m^T x + \tilde{b}^m \leq 1 \end{array}$$

Now we imbed our polyhedron $P_n$ in $\mathbf{R}^{n+1}$ by writing

$$(a_i^1, a_i^2, a_i^3, \ldots, a_i^n, \tilde{b}^i) \begin{pmatrix} x^1 \\ \vdots \\ x^n \\ x^{n+1} \end{pmatrix} \leq 1 \iff a_i^T x + \tilde{b}^i x^{n+1} \leq 1$$

for $i = 1, \ldots, m$.

This inequality system defines a new polyhedron $P_{n+1} \subset \mathbf{R}^{n+1}$.

In level $x_{n+1} = 1$ we find our original polyhedron $P_n$ (augmented by 1's in the additional component), because $a_i^T x + \tilde{b}^i x_{n+1} \leq 1$ and $x_{n+1} = 1$ means that $a_i^T x + \tilde{b}^i \leq 1$.

In level $x_{n+1} = 0$ we find a polyhedron $\tilde{P}_n$ (augmented by 0's in the additional component), because $a_i^T x + \tilde{b}^i x_{n+1} \leq 1$ and $x_{n+1} = 0$ means that $a_i^T x \leq 1$.

If all the vectors $a_i$ satisfy the conditions of (5.2), we can immediately solve the $\tilde{P}_n$-problem and our known results hold. Now we solve what we can solve and proceed to level 1.

1) Solve the problem of $\tilde{P}_n$ (by our old complete method, 0 is feasible). If this problem has no solution, we can STOP (unboundedness of $v^T x$ leads to unboundedness of $v^T x$ also in $P_n$ — if $P_n$ is nonempty).

2) Consider the solution of the $\tilde{P}_n$-problem in level 0. It is located on an edge of $P_{n+1}$. Now apply the shadow vertex algorithm to

$$\text{(5.11)} \qquad \begin{array}{ll} \text{maximize} & x^{n+1} \\ \text{subject to} & a_i^T x + \tilde{b}^i x^{n+1} \leq 1 \ (1 = 1, \ldots, m). \end{array}$$

Every point on the realized path is $(v, e_{n+1})$-cooptimal. Hence a path point of level $x^{n+1} = 1$ will be $v$-optimal in level $x_{n+1} = 1$. Then the truncated vector $(x^1, \ldots, x^n)$ is the optimal solution.

If we cannot reach level $x_{n+1} = 1$, then this proves that $P_n$ is empty and that our original problem was infeasible.

Since this is only a prolongation of our known complete method, the results hold, if only the stochastic assumptions still hold.

The effort for Step 1 is known from (5.9). We have to consider the additional effort for Step 2.

We have a first obvious result:

**(5.12) Theorem.** *If* $\binom{a_1}{\tilde{b}^1}, \ldots, \binom{a_m}{\tilde{b}^m}$ *are distributed on* $\mathbf{R}^{n+1}$ *according to (5.2), then*

$$E_{m,n}(s^t) \leq m^{\frac{1}{n}}(n+2)^4 \mathcal{C} \quad \textit{(for the total algorithm).}$$

**A special result.**

**(5.13) Theorem.** *If* $a_1, \ldots, a_m, \tilde{b}^1, \ldots, \tilde{b}^m$ *are independent and all components of* $A$ *and* $\tilde{b}$ *are Gaussian distributed, then*

$$E_{m,n}(s^t) \leq m^{\frac{1}{n}}(n+2)^4 \mathcal{C}.$$

For more general distributions of the right hand sides $1 - \tilde{b}^i$ a lot of additional work has to be done. Up to now we know

**(5.14) Theorem.** *If $a_1, \ldots, a_m, \tilde{b}^1, \ldots, \tilde{b}^m$ are independent, the $a_i$ are distributed according to (5.2) and if the $\tilde{b}^i$ are uniformly distributed over an interval $[-q, q]$, then Step 2 does not destroy the size of effort in Step 1, which is $0(m^{\frac{1}{n-1}} n^4)$.*

So we have an algorithm which admits a probabilistic analysis solving all types of LP-problems. Its average behaviour is still polynomial, although we work under the "hard" rotational symmetry model.

## References


[1] Adler, I., Karp, R. & Shamir, R. [1987]: *A Simplex Variant Solving an $m \times d$ Linear Program in $0(\min(m^2, d^2))$ Expected Number of Pivot Steps,* Journal of Complexity 3 (1987).

[2] Adler, I. & Meggido, N. [1985]: *A Simplex Algorithm where the Average Number of Steps is Bounded Between two Quadratic Functions of the Smaller Dimension,* Journal of the ACM 32 (1985), 871–895.

[3] Borgwardt, K. H. [1977]: *Untersuchungen zur Asymptotik der mittleren Schrittzahl von Simplexverfahren in der linearen Optimierung,* Dissertation Universität Kaiserslautern.

[4] Borgwardt, K. H. [1982a]: *Some Distribution-Independent Results About the Asymptotic Order of the Average Number of Pivot Steps of the Simplex Method,* Mathematics of Operations Research 7 (1982), 441–462.

[5] Borgwardt, K. H. [1982b]: *The Average Number of Pivot Steps Required by the Simplex-Method is Polynomial,* Zeitschrift für Operations Research 26 (1982), 157–177.

[6] Borgwardt, K. H. [1987]: *The Simplex Method — A Probabilistic Analysis,* Springer Verlag Berlin, Heidelberg, New York, 1987.

[7] Haimovich, M. [1983]: *The Simplex Algorithm is Very Good! — On The Expected Number of Pivot Steps and Related Properties of Random Linear Programs,* 415 Uris Hall, Columbia University, New York, April 1983.

[8] May, J. H. & Smith, R. L. [1982]: *Random Polytopes: Their Definition, Generation and Aggregate Properties,* Mathematical Programming 24 (1982), 39–54.

[9] Shamir, R. [1987]: *The Efficiency of the Simplex Method: A Survey,* Management Science 33 (1987), 301–334.

[10] Smale, S. [1983]: *On the Average Speed of the Simplex Method,* Mathematical Programming 27 (1983), 241–262.

[11] Todd, M. J. [1983]: *Polynomial Expected Behavior of a Pivoting Algorithm for Linear Complementarity and Linear Programming Problems,* Mathematical Programming 35 (1986), 173–192.

[12] Adler, I. [1983]: *The Expected Number of Pivots Needed to Solve Parametric Linear Programs and the Efficiency of the Self-Dual Simplex Method,* Manuscript, Dept. of Industrial Engineering and Operations Research, University of California, Berkeley, California, 1983.

[13] Megiddo, N. [1985]: *A Note on the Generality of the Selfdual Algorithm with Various Starting Points,* Methods of Operations Research 49 (1985), 271–275.

[14] Adler, I. & Berenguer, S. E. [1983]: *Generating Random Linear Programs,* Manuscript, Dept. of Industrial Engineering and Operations Research, University of California, Berkeley, California, 1983.

[15] Adler, I. & Berenguer, S. E. [1981]: *Random Linear Programs,* Operations Research Center Report No. 81-4, Dept. of Industrial Engineering and Operations Research, University of California, Berkeley, California, 1981.