Probabilistic Analysis of the Simplex Method

Karl Heinz Borgwardt

University of Augsburg

1. Introduction

Since the invention of the Simplex Method in 1947 by George B. Dantzig it had been an open problem for mathematicians why this algorithm works so efficiently in practice. This question became even more interesting when in the beginning of the seventies Klee, Minty, Jeroslow and other authors recognized that the usual variants of the Simplex Method are nonpolynomial in the sense of worst-case complexity theory. That means that there is no polynomial in the dimension of linear programming problems which acts as an upper bound for the respective numbers of pivot steps required for solving all the possible problems.

Obviously it was a great advantage for the Simplex Method that people had not known this negative and bad fact from the beginning. So the users got the impression that this is a very reliable, fast and efficient tool for solving their problems. They did not care very much about the reasons for that behaviour. On the other hand an immediate nonpolynomiality proof would very likely have led to a very cautious and detached judgement on the quality of the method as we have learned from many other cases (especially in combinatorial optimization). The persuasive power of polynomiality proofs can be observed very clearly at the example of the ellipsoid method (which is not efficient in practice at all) and at the boom and the public attention for the algorithm of Karmarkar.

Since the beginning of the seventies a great effort has been made to clarify the "normal" quality of the Simplex Method and in 1977 — ten years ago, the first breakthrough was achieved by the author. In the following time several authors tried to improve the results, find sharper bounds in different directions studying different stochastic models. This paper is to give a brief survey of these approaches and explain the advantages and drawbacks of the main results in this field of research.

Throughout the paper we will deal with linear programming problems of the following type. (Note that different types can easily be transformed into that form.)

$$\begin{array}{lll} \text{Maximize} & v^T x & \text{(a linear objective)} \\ \text{where} & a_1^T x \le b^1 & \\ & \vdots & \text{resp. } Ax \le b \text{ with } A = \begin{pmatrix} a_1^T \\ \vdots \\ a_m^T \end{pmatrix} \\ & a_m^T x \le b^m & \end{array}$$

with $v, x, a_1, \ldots, a_m \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ and $m \ge n$.

We call this an (m, n)-problem to characterize the two dimensions of the problem.

The restrictions $Ax \leq b$ define a convex polyhedron. In the vertices (edges) of that polyhedron at least n $(n-1)$ restrictions $a_i^T x \leq b^i$ are active. And we know that if there are vertices and if there are optimal points, there is a vertex among the optimal points. Finally, we know that every nonoptimal vertex is incident to an edge which improves the objective. These facts are exploited in the Simplex Method which works as follows.

Phase I: Find a vertex of $X = \{x \in \mathbb{R}^n \mid Ax \leq b\}$.
If this is impossible, STOP.

Phase II: Construct a sequence of vertices $x_0, x_1, \ldots, x_s$ starting with the vertex x_0 (found in Phase I) such that successive vertices are adjacent and that $v^T x$ is strictly increasing.
Stop as soon as an optimal vertex is reached or as soon as it becomes clear that optimal points do not exist.

The procedure for Phase I is closely related to that for Phase II. Only a slight modification of the problem is necessary. So it makes sense to concentrate on Phase II for the moment. The number s characterizes the number of vertex exchanges which have to be performed. Under nondegeneracy assumptions (which are valid in our analysis) this number coincides with the number of pivot steps for solving the Phase II-task.

Our definition of Phase II is not yet complete, because we should declare a rule for determining the successor vertex if more than one improving adjacent vertices are available. Such a rule characterizes the **variant** of the algorithm.

Throughout the paper we use the following nondegeneracy assumptions.

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

This guarantees that each vertex corresponds to a set of exactly n active restrictions, that an optimal point is unique (if existing) and that every pivot step leads to a new vertex.

Carrying out a probabilistic analysis of the Simplex Method requires the fixation of a variant and the decision for a certain stochastic model simulating the real-world-distribution of linear programming problems.

It should be mentioned that the number of steps highly depends

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

Studying the average behaviour means calculating the expected number of steps under the given stochastic model. In the following we will demonstrate and explain the two main stochastic models, the main ideas for evaluation of the expected values and the main results, which showed the polynomiality of the average behaviour.

2. The Sign-Invariance-Model

We begin with the most simple stochastic model which could be analyzed, the so-called Sign-Invariance-Model (MAY & SMITH (1982)). Let $A = \begin{pmatrix} a_1^T \\ \vdots \\ a_m^T \end{pmatrix}$ and $b = \begin{pmatrix} b^1 \\ \vdots \\ b^m \end{pmatrix}$ be given. To determine the direction of the inequalities ($a_i^T x \leq b^i$ or $a_i^T x \geq b^i$) we perform m Bernoulli-experiments (we flip a coin m times) with probability $\frac{1}{2}$. So the m directions are independent, $\geq$ and $\leq$ are equally likely.

Consequently, a given data set (A, b) induces 2^m different linear programming problems. By multiplication with (-1) it is possible to transform the restrictions into the standard form $Ax \leq b$. Averaging over the probability space now means to solve all these 2^m problems, to count the numbers of required pivot steps and to divide the sum by 2^m. The feasibility regions of the individual problems will be called cells. Of course, some of the cells will be empty. Then the algorithm has the task to recognize infeasibility. A first analysis for the average Phase II-complexity was done by HAIMOVICH (1983) and ADLER. They investigated and used a parametric Simplex variant which works as follows.

Let v be the objective of the problem and let u be an auxiliary objective such that $\begin{pmatrix} A \\ v \\ u \end{pmatrix}$ has only full-rank submatrices. We call a point $y \in X$ (the feasible region) (u, v)-cooptimal if $u^T z \leq u^T y$ for all $z \in X$ with $v^T z = v^T y$.

Now we can find a Simplex-Path connecting all (u, v) cooptimal points according to the parametric algorithm of GASS & SAATY (1955). That means that we successively calculate the solution-vertices of the problem

$$\begin{aligned} &\text{Maximize} && (u + \rho v)^T x \\ &\text{subject to} && x \in X \\ &\text{with } \rho \text{ running from } -\infty \text{ to } +\infty. \end{aligned}$$

We can assume that Phase I is already done and that a start vertex x_0 has been found such that x_0 is a maximal vertex with respect to a given auxiliary objective u. Constructing a vertex-sequence according to the algorithm described above touches only (u, v)-cooptimal vertices. So the number of cooptimal vertices can be used as an upper bound for the number of pivot steps.

Haimovich comes to the following conclusion.

Theorem (Haimovich 1983). *The expectation value of the numbers of pivot steps used on cooptimal paths in the sign invariance model and under the condition that a cooptimal path exists, is not greater than* $n\frac{m-n+2}{m+1}$.

Proof. There are 2^m possible cells and $\binom{m}{n}$ intersection points of n restriction hyperplanes. We call the intersection sets of $n-1$ restriction hyperplanes **lines**.

Every such line itself is intersected by $m-n+1$ restriction-hyperplanes. So we obtain $m-n+2$ segments on each line. Each segment belongs to a different cell and every segment is cooptimal in exactly one cell. Now we know that nondegeneracy guarantees that

1) Every cell has at most one $v^T x$-maximal vertex and at most one (u,v)-cooptimal path.
2) Each intersection point of n hyperplanes and every line segment is optimal in exactly one cell.
3) We have $\binom{m}{n} + \binom{m}{n-1}$ cells with (u,v)-cooptimal paths.
3) There are $\binom{m}{n-1}(m-n+2)$ line segments.

So he concludes $E_{m,n}$ (number of line segments on a cooptimal path/cooptimal path exists) $= n\frac{m-n+2}{m+1}$.

Here an analysis of Phase I is not yet done. We have a fictive start vertex based on the auxiliary objective u.

That necessary generalization was independently done in papers of TODD (1983), ADLER & MEGIDDO (1983), ADLER, KARP & SHAMIR (1983). Their methods were based on work of SMALE (1982) who had analyzed a similar model and algorithm but had not achieved polynomiality. We refer to the paper of Adler, Karp and Shamir, because its notation is closest to our geometrical concept. Again, we want to solve the problem

$$\begin{array}{ll} \text{Maximize} & v^T x \\ \text{subject to} & a_1^T x \leq b^1, \ldots, a_m^T x \leq b^m. \end{array}$$

But now we use a Lexicographic Constraint-By-Constraint algorithm for doing the task of Phase I and Phase II together. For this purpose we use relaxed feasibility regions $X^{(k)} = \{x \in \mathbb{R}^n \mid a_1^T x \leq b^1, \ldots, a_k^T x \leq b^k\}$ for $n \leq k \leq m$. Note that $X^{(m)} = X$. Our algorithm runs through $m-n+1$ stages.

Stage 0: Determine the unique vertex $\overline{x} \in X^{(n)}$ and choose u as $u = \varepsilon^1 a_1 + \varepsilon^2 a_2 + \ldots + \varepsilon^n a_n$ with $\varepsilon > 0$ sufficiently small.
Go to Stage 1.

Stage k: $(1 \leq k \leq m-n)$
Start at $\overline{x}$, the maximal vertex for $u^T x$ on $X^{(k-1+n)}$. If $\overline{x} \in X^{(k+n)}$, then go to Stage $k+1$. Else use the parametric algorithm to maximize $(-a_{k+n}^T x)$ until $a_{k+n}^T x \leq b^{k+n}$ is achieved. The last traversed edge contains a point x with $a_{k+n}^T x = b^k$. All path-points are $(u, -a_k)$-cooptimal, hence x maximizes $u^T x$ on $X^{(k+n)}$. If it is impossible to achieve $a_{k+n}^T x \leq b^{k+n}$ then the original problem is infeasible. STOP. Else set $\overline{x} = x$ and go to Stage $k+1$.

Stage $m-n+1$: Start at $\bar{x}$ which is $u^T x$-maximal on $X^{(m)} = X$. Apply the parametric algorithm and find the optimal point or find out that there is no solution.

Adler, Karp and Shamir sum up over all 2^{n+k} instances in Stage $k+1$ and over all the basic cones. A fixed vertex is cooptimal in exactly $n+1$ of 2^{n+k} instances, because it is incident to $2n$ segments. Either it is optimal or two of its incident segments are cooptimal. The authors come to an expected value of at most $2^{n+1}(n+1)$ for arbitrary choice of u. But the lexicographic choice of u in Stage 0 enables them to prove that

$$E_{m,n}(s^{CBC}) \leq 2(n+1)^2.$$

Theorem (ADLER, KARP & SHAMIR (1983); ADLER, MEGIDDO (1983); TODD (1983)). *For solving the complete (m,n)-problem, the expected number of pivot steps is not greater than $2(n+1)^2$ under the sign invariance model.*

3. Drawbacks of the Sign-Invariance-Model

We already mentioned that not every problem instance has a feasible cell. We know that m restrictions (nondegenerate) partition the n-dimensional space into $\binom{m}{0}+\binom{m}{1}+\ldots+\binom{m}{n}$ cells. On the other hand the number of generated problems is 2^m. So for $m \gg n$ the quotient

$$\frac{\text{number of feasible problems}}{\text{number of generated problems}} \quad \text{tends to 0 very fast.}$$

Only **conditioning** on feasible problems can avoid averaging over a lot of infeasible (and much less difficult) problems. Also, the expected number of vertices is low. We have

$$E_{m,n}\ (\text{vertices per generated problem}) = \frac{2^n\binom{m}{n}}{2^m}$$

$$E_{m,n}\ (\text{vertices per feasible problem}) = \frac{2^n\binom{m}{n}}{\binom{m}{0}+\ldots+\binom{m}{n}}$$

which tends to 2^n for $m \to \infty$, n fixed and is similar to the vertex number of an n-dimensional cube (a very simple polyhedron for linear programming).

The most important aspect is the average redundancy rate. If we introduce the m-th restriction, the number of cells increases from

$$\binom{m-1}{0}+\ldots+\binom{m-1}{n} \text{ to } \binom{m}{0}+\ldots+\binom{m}{n}.$$

That means that some cells have been reproduced and some have been divided by the last restriction. We can calculate the number of divided cells. It is

$$\binom{m}{0}-\binom{m-1}{0}+\binom{m}{1}-\binom{m-1}{1}+\ldots+\binom{m}{n}-\binom{m-1}{n}=\binom{m-1}{0}+\ldots+\binom{m-1}{n-1}.$$

Consequently the new restriction is relevant only in $2[\binom{m-1}{0}+\ldots+\binom{m-1}{n-1}]$ of 2^m generated and $\binom{m}{0}+\ldots+\binom{m}{n}$ feasible cells. Hence both the pure and the conditional redundancy rate tends to 1 very quickly.

Also the expected number of nonredundant constraints is low. It is calculated by multiplication of m with the nonredundancy rate. Whereas the pure expected number again tends to 0, the conditional expected number of nonredundant constraints behaves like

$$\frac{m2\{\binom{m-1}{0}+\ldots+\binom{m-1}{n-1}\}}{\binom{m}{0}+\ldots+\binom{m}{n}} \quad \rightarrow \quad 2n \text{ for } m \rightarrow \infty \text{ and } n \text{ fixed.}$$

This is again typical for a cube in $\mathbb{R}^n$. These results show that the Sign-Invariance-Model hardly generates problems with more than $2n$ nonredundant constraints. So the expected values for s and s^{CBC} do not give reliable explanations for the behaviour of the Simplex Method in the asymptotic case ($m \gg n$). Instead of proving the quality of the Simplex Method, they reflect the special properties of the chosen stochastic model.

4. The Rotation-Symmetry-Model

The above mentioned danger can be avoided by using a different model, the rotation-symmetry-model. The author analyzed that model in several papers (BORGWARDT (1977, 1982a, 1982b etc.)). We deal with problems of the type

$$\begin{array}{ll} \text{Maximize} & v^T x \\ \text{subject to} & a_1^T x \leq b^1 \\ & \vdots \\ & a_m^T x \leq b^m \\ \text{where } b^i > 0 \text{ for } i = 1, \ldots, m. & \end{array}$$

Here we have feasibility in any case because $0 \in X$. So we are sure to solve "hard" problems. Our stochastic assumptions are as follows:

> Let $a_1, \ldots, a_m$, v, $b^1, \ldots, b^m$ be independent random variables, let $a_1, \ldots, a_m$, v be distributed identically and symmetrically under rotations of $\mathbb{R}^n$ and let the components of b be distributed identically over $(0, \infty)$.

Without loss of generality we can simplify these assumptions by "normalization". That means that we define X by $a_1^T x \leq 1, \ldots, a_m^T x \leq 1$ and that we demand $a_1, \ldots, a_m$, v to be distributed independently, identically and symmetrically under rotations. It

should be noted that under this model the following nondegeneracy assumption is valid with probability 1: Every submatrix of $\binom{A}{v^T}$ and of $(A\ \mathbb{1})$ has full rank.

The variant is again a parametric algorithm, the so-called "shadow vertex algorithm". Here we start again at a vertex maximizing an auxiliary objective $u^T x$ and construct a sequence of (u,v)-cooptimal vertices in order to reach the $v^T x$-maximum. All these cooptimal vertices will be mapped into vertices of a two dimensional polyhedron, if we project X orthogonally onto the plane spanned by u and v. That means that they become vertices of "the shadow" of X, which explains the term "shadow vertex". Since every (u,v)-cooptimal vertex is also a shadow vertex, we know that S, the number of shadow vertices, is an upper bound for s.

Now we observe the following one-to-one correspondence. Every vertex of X is the intersection point of n restriction hyperplanes. Let us denote these restrictions $a_{\Delta^1}^T x = 1, \ldots, a_{\Delta^n}^T x = 1$ and let x_Δ be the intersection point. Hence we have a one-to-one correspondence between the vertex x_Δ and an index set $\Delta = \{\Delta^1, \ldots, \Delta^n\} \subset \{1, \ldots, m\}$ with $\Delta^1 < \Delta^2 < \ldots < \Delta^n$. On the other hand Δ corresponds uniquely to the n-element set $\{a_{\Delta^1}, \ldots, a_{\Delta^n}\}$. For our analysis of the average number of shadow vertices the following geometrical facts are extremely useful. Let us denote by Y the convex hull of $(0, a_1, \ldots, a_m)$.

Lemma. *1. x_Δ is a vertex of X if and only if the convex hull of $\{a_{\Delta^1}, \ldots, a_{\Delta^n}\}$ lies completely on the boundary of Y.*

2. Let x_Δ be a vertex of X. Then it is even a shadow vertex if and only if the convex hull of $\{a_{\Delta^1}, \ldots, a_{\Delta^n}\}$ is intersected by $\operatorname{span}(u,v)$.

The following integral expression results from that lemma.

$E_{m,n}(S) = \binom{m}{n} \int_{\mathbb{R}^n} \ldots \int_{\mathbb{R}^n} P\ (\mathrm{CH}(a_1, \ldots, a_n)$ is a boundary simplex of Y and $\mathrm{CH}(a_1, \ldots, a_n)$ is intersected by $\operatorname{span}(u,v))\ dF(a_1) \ldots dF(a_n)$.

Here F denotes the distribution function under consideration. This integral expression can be simplified because of rotational symmetry. So the boundary event and intersection event become independent for fixed $a_1, \ldots, a_n$. The first depends only on the distance of the hyperplane through $a_1, \ldots, a_n$ to 0 and the second is proportional to the spherical measure of the convex cone spaned by $a_1, \ldots, a_n$. Nevertheless it took several years to evaluate this integral sufficiently well. In the following I want to list some of the main results for our Phase I-analysis.

Asymptotic Results for $m \to \infty$ and fixed n

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

Gaussian distribution on $\mathbb{R}^n$:

$$E_{m,n}(S) \leq \sqrt{\ln m}\, n^{\frac{3}{2}}\, 2^{\frac{3}{2}}\, \frac{20}{9}(1+\varepsilon(m,n))$$

$$E_{m,n}(S) \geq \sqrt{\ln m}\, n^{\frac{3}{2}}\, 2^{\frac{3}{2}}\, \frac{3}{5}(1-\varepsilon(m,n)).$$

Uniform distribution on unit ball:

$$E_{m,n}(S) \leq m^{\frac{1}{n+1}}\, n^2\, 2\Big(1+\frac{1}{\sqrt{2}}\Big)(1+\varepsilon(m,n))$$

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

Uniform distribution on unit sphere:

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

$$E_{m,n}(S) \geq m^{\frac{1}{n-1}}\, n^2\, 2\frac{1}{7}(1-\varepsilon(m,n)).$$

General distributions with bounded support:

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

(ε depends on F here).

The difference in the order of growth for different distributions results from different redundancy rates. For instance uniform distribution on unit sphere leads to redundancy rate 0, whereas Gaussian distribution induces a redundancy rate rather close to 1. Note that:
The restriction $a_i^T x \leq 1$ is redundant if and only if $a_i \in \operatorname{Int} Y = \operatorname{Int} \operatorname{CH}(0, a_1, \ldots, a_m)$. This can never occur for a distribution on the unit sphere.
Now we could try to find the minimal possible asymptotic growth.

Theorem (Borgwardt 1979–1984). *All results for $m \to \infty$ and fixed n.*

1) *For all distributions where $P(\|a\| \geq r)^{-1}$ is of polynomial order for $r \to \infty$ we know that $E_{m,n}(S)$ is bounded by a constant $C(n)$ not depending on m. (Example $C(n) = n^{\frac{5}{2}}$ can be realized).*
2) *For all distributions with bounded support $E_{m,n}(S) \to \infty$.*
3) *For every $\delta > 0$ there is a distribution with bounded support such that $E_{m,n}(S) = 0(m^\delta)$ as a function of m.*

Note the similarity of 1) with the result for the Sign-Invariance-Model. It shows that such a small growth can even be achieved by analyzing only feasible problems when the redundancy rate is kept small.

This is a great advantage of our model. We can vary the redundancy rate from almost 1 to 0 and observe the behaviour, whereas the Sign-Invariance-Model has a fixed redundancy rate very close to 1. In 1981 we achieved a polynomial upper bound.

Theorem (Borgwardt 1981). *For all distributions satisfying our conditions independence, identity, rotational symmetry we have* $E_{m,n}(S) \leq e\pi(\frac{\pi}{2} + \frac{1}{e})n^3 m^{\frac{1}{n-1}}$.

5. The Inclusion of Phase I

As in the analysis of Phase II under the Sign-Invariance-Model we had some trouble with the auxiliary objective function $u^T x$. These complications disappear completely by our construction of a probabilistically analyzable method for inclusion of Phase I.

For demonstrating that algorithm we introduce the following problems with additional restrictions

Problem	I_k	$(k = 2, \ldots, n)$
Maximize	$v^T x$	
subject to	$Ax \leq b$	and $x^{k+1} = \ldots = x^n = 0$.

That means we fix components $k+1, \ldots, n$ for a while to 0. Note that I_n is our original problem.

Initialization:

Set $k = 2$. Find a vertex of I_2. Solve I_2 by means of the parametric algorithm.
If this is impossible STOP.
Else store the solution vertex $(x^1, x^2, 0, \ldots, 0)^T$.

Typical Step:

$(k = 3, \ldots, n)$.
Use the solution point $(x^1, \ldots, x^{k-1}, 0, \ldots, 0)^T$ for I_{k-1}. It is located on an edge of the I_k-polyhedron.

1) Find an adjacent vertex $(\overline{x}^1, \ldots, \overline{x}^k, 0, \ldots, 0)^T$ of the I_k-polyhedron to that edge.
2) Solve I_k by means of the parametric algorithm by setting $u = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ (1 in kth-component).
   If the problem has unbounded objective STOP.
   Else store the maximal vertex.
3) If $k = n$ STOP. Else set $k = k + 1$. Go to 1).

Here there is no doubt about the choice of u. The algorithm is by far not the most efficient one but it has the tremendous advantage that we can do a probabilistic analysis by $n - 1$-times application of the shadow-vertex algorithm.

So we get our main result.

Theorem (Borgwardt 1981). *For every distribution satisfying our three conditions our complete method does not require more than* $m^{\frac{1}{n-1}} n^4 \cdot C$ *steps on the average.*

Based on that method we could also prove that problems with sign constraints, i. e.,

$$\begin{array}{ll} \text{Maximize} & v^T x \\ \text{subject to} & Ax \leq b \;\; x \geq 0 \text{ and } b > 0 \end{array}$$

can be solved in $m^{\frac{1}{n-1}} n^4 C$ steps on the average.

But still our analysis had one drawback. For our algorithm it is necessary to know a feasible point (to start with), here we have the origin. If we knew a different feasible point, it would work too, we only had to transform the coordinate system. But how should we handle problems of the form

$$\begin{array}{ll} \text{Maximize} & v^T x \\ \text{subject to} & a_1^T x \leq b^1, \ldots, a_m^T x \leq b^m \text{ (general form)} \\ & \text{with arbitrary } b^i? \\ & \text{We call that problem } P_n. \end{array}$$

If there is a negative b^i we do not know a feasible point.
Now we reformulate our restrictions as

$$a_i^T x \leq b^i \iff a_i^T x \leq 1 - \tilde{b}^i \text{ with } \tilde{b}^i = 1 - b^i.$$

We embed our polyhedron in $\mathbb{R}^{n+1}$ by postulating

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

In level $x^{n+1} = 0$ we have a polyhedron and a problem which satisfies our requirements about the distribution of the restrictions, because the a_i are distributed independently, identically and symmetrically under rotations. That means our analysis would hold for the solution of those problems $\tilde{P}_n$. But we must be aware of the fact that we have to find the $v^T x$-optimum in level $x^{n+1} = 1$, where $a_i^T x + \tilde{b}_i x^{n+1} \leq 1$ means that $a_i^T x + \tilde{b}_i \leq 1 \iff a_i^T x \leq 1 - \tilde{b}_i$. Hence we rediscover our problem in level $x^{n+1} = 1$.

Now we proceed as follows.

1) Solve the problem $\tilde{P}_n$ (by means of our complete method)

$$\begin{array}{ll} \text{Maximize} & v^T x + 0 x^{n+1} \\ \text{subject to} & a_i^T x + \tilde{b}_i x^{n+1} \leq 1 \text{ and } x^{n+1} = 0 \; (i = 1, \ldots, m). \end{array}$$

If this problem has no solution STOP (unboundedness of the objective is guaranteed also in P_n).
Else store the solution vertex and go to 2).

2) Apply the shadow vertex algorithm in order to

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

Every point on the path is (v, e_{n+1})-optimal. If the maximal value of x^{n+1} is less than 1 then our problem is infeasible $\rightarrow$ STOP.
Else: As soon as level $x^{n+1} \geq 1$ is reached, find a point of level $x^{n+1} = 1$ on the last edge. Now $(x^1, \ldots, x^n)^T$ is the optimal point for P_n. STOP.

We ask for the average complexity of that method. It is clear that the solution of $\tilde{P}_n$ requires not more than $m^{\frac{1}{n-1}} n^4 C$ on the average.

We need to know the effort for Step 2. And we can prove the following

Theorem (Borgwardt 1987). *Problems of the kind*

$$\begin{array}{ll} \textit{Maximize} & v^T x \\ \textit{subject to} & a_1^T x \leq b^1, \ldots, a_m^T x \leq b^m \end{array}$$

where $v, a_1, \ldots, a_m$ are distributed independently, identically and symmetrically under rotations, where the b^i's are independent real variables, also independent of the a_i's and uniformly distributed over an interval $[-q, +q] \subset \mathbb{R}$, then the method described above does not require more than $m^{\frac{1}{n-1}} n^4 C_2$ pivot steps on the average.

So we have an algorithm — by far not the most efficient one — which admits a probabilistic analysis, which solves all LP-problems and which is polynomial in the average number of steps. And the analysis is done under the "hard" rotational symmetry model.

References

[1] Adler, I., Karp, R. & Shamir, R. [1983b]: *A Simplex Variant Solving an $m \times d$ Linear Program in $0(\min(m^2, d^2))$ Expected Number of Pivot Steps*, University of California, Computer Science Division, Berkeley, December 1983.

[2] Adler, I. & Meggido, N. [1983]: *A Simplex Algorithm where the Average Number of Steps is Bounded Between two Quadratic Functions of the Smaller Dimension*, Department of Industrial Engineering and Operations Research, University of California, Berkeley, California, December 1983.

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

[3] 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.