Ideas Leading to a Better Bound on the Average Number of Pivot Steps for Solving an LP

Karl Heinz Borgwardt

Department of Mathematics, University of Augsburg, 86135 Augsburg, Germany

1 Introduction

This paper deals with the complexity of the solution-process of linear programming problems of the following type:

$$\begin{array}{ll} \text{maximize} & v^T x \\ \text{subject to} & a_1^T x \le 1, \ldots, a_m^T x \le 1 \\ \text{where} & x, v, a_1, \ldots, a_m \in \mathbb{R}^n \text{ and } m \ge n. \end{array} \tag{1}$$

Note that the lower dimension n is the number of variables, and that the larger dimension m gives the number of restrictions defining the feasible polyhedron

$$X = \{x \mid a_1^T x \le 1, \ldots, a_m^T x \le 1\}. \tag{2}$$

We are interested in the expected numbers of iteration steps of solution algorithms. So we have to base our considerations on a stochastic model on the distribution of the input data of (1) and choose the Rotation-Symmetry-Model (RSM):

$a_1, \ldots, a_m, v$ and an auxiliary vector u are distributed on $\mathbb{R}^n \setminus \{0\}$ identically, independently and symmetrically under rotations. (3)

Linear programming problems generated according to (3) satisfy with probability 1 (almost surely) the Nondegeneracy-Condition:

Any n vectors out of $\{a_1, \ldots, a_m, u, v\}$ are linearly independent and any $n+1$ points out of $\{a_1, \ldots, a_m\}$ are in general position. (4)

Hence the concentration on nondegenerate cases does not affect the calculation of the expected value of bounded random variables, as e.g. the number of required iterations in a special variant of the Simplex-Method, whose general description is given als follows:

Phase I: Decide whether X possesses a vertex. STOP if the answer is NO. If the answer is YES, calculate such a vertex $x_0 \in X$. (5)

Phase II: Construct a sequence of vertices $x_0, \ldots, x_s \in X$, such that for $i = 0, \ldots, s-1$ the vertices x_i and x_{i+1} are adjacent and $v^T x_i < v^T x_{i+1}$. The final vertex x_s is either the optimal vertex or a vertex where the nonexistence of an optimum becomes obvious. (6)

Since Phase I can be organized similarly to Phase II, it makes sense to concentrate on Phase II at the moment. We shall discuss Phase I in section VI. Let us first ask for the number s. The natural probabilistic question is: How large is s *on the average* for a fixed dimension-pair (m, n)? Note that in (6) there is no unique rule for determining the successor vertex. Such a rule will characterize a variant of the Simplex-Algorithm. Different variants may lead to completely different paths.

Figure 1: A polyhedron X, vectors v and u, shadow-vertices $x_0, \ldots x_9$, a shadow-vertex path $x_0, \ldots x_5$ from the u-optimum to the v-optimum, and an alternative (gray) Simplex-Path in the background.

One of these variants, the so-called *shadow-vertex-algorithm*, will be the object of our investigation, because it admits a simple geometric interpretation. Suppose that a vertex x_0 has been derived in Phase I. Then x_0 is an extremal point of X and there are some objectives, which are maximized exactly at x_0. Let u be such an objective direction, i.e. x_0 maximizes $u^T x$ on X. Remember that our original objective was $v^T x$. Now project $X \subset \mathbb{R}^n$ on $Span(u, v)$, the two-dimensional plane spanned by u and v. This projection induces a classification of the vertices of X:

> Some vertices are mapped on vertices of the two-dimensional image of X and can still be identified after the projection. These vertices will be called **shadow-vertices**.
>
> The other vertices are mapped into the interior of the image of X.

Now, there is a path from x_0 to x_s visiting only shadow-vertices, which can easily be realized by implementing the shadow-vertex-algorithm. Hence the number of shadow vertices is a natural upper bound for the number of vertices visited on this path from x_0 to x_s. And, it turns out that with the RSM the expected number of necessary vertex-exchanges $E_{m,n}(s)$ is just a quarter of the expected number of shadow vertices $E_{m,n}(S)$. So it suffices to derive upper (and lower) bounds $E_{m,n}(S)$, as done in [1], [2], [3] . The most important result was given in [3]:

$$E_{m,n}(S) \leq \text{Const.}\ m^{\frac{1}{n-1}} \cdot n^3 \quad \text{for all } (m, n) \text{ and for all RSM-distributions.} \qquad (7)$$

But from [1] we knew an upper bound for all RSM-distributions with bounded support, which only applies to a subset of dimension pairs, namely the so called "asymptotic case" $m \to \infty, n$ fixed, (i. e. n is fixed and m must be larger than an unspecified value $\overline{m}(n)$). Then

$$E_{m,n}(S) \leq \text{Const. } m^{\frac{1}{n-1}} \cdot n^2 \quad \text{for } m \to \infty \text{ and } n \text{ fixed.} \tag{8}$$

Numerical experiments and some crude estimations in the derivation of (7) made it plausible that (8) might hold for all dimension-pairs. The proof for that (10 years open) conjecture was given in [4] and is very long and technical. Now we know

$$E_{m,n}(S) \leq \text{Const. } m^{\frac{1}{n-1}} \cdot n^2 \quad \text{for all } (m,n) \text{ and for all RSM-distributions.} \tag{9}$$

The existence of a derivation for a lower bound in [3] for $E_{m,n}(S)$, when a special RSM-distribution is applied, shows that our new result is in a certain sense sharp: For uniform distribution on ω_n – the unit sphere in $\mathbb{R}^n$ – we know that

$$E_{m,n}(S) \geq \text{Const. } m^{\frac{1}{n-1}} \cdot n^2 \quad \text{for } m \to \infty \text{ and } n \text{ fixed.} \tag{10}$$

So it will not be possible to improve (9) by giving lower orders of growth for m resp. n without deteriorating the growth in the other dimension.

It is the aim of this paper to give background information, why the new approach in [4] could deliver the better (upper) bound and why the old approach in [3] had to fail in getting absolute precision. For this purpose, and as [4] contains the formal proof, we concentrate on giving illustrative, plausible arguments for the comparison of both approaches. In addition to [4], we show (in section VI) the impact of the new proof on the average-case analysis of the number s_t of pivot steps in the complete Simplex-Method (including Phase I). Here, the bound can now be reduced to

$$E_{m,n}(s_t) \leq \text{Const. } m^{\frac{1}{n-1}} \cdot n^3 \quad \text{for all } (m,n) \text{ and for all RSM-distributions.} \tag{11}$$

2 Analysis of the Primal and the Polar Polyhedron

Since we are to count a certain subset of the vertices of X and since the impact of the random data $a_1, \ldots, a_m$ on X is rather indirect, it is recommended to shift our analysis to the corresponding polar (resp. dual) polyhedron

$$Y := CH(0, a_1, \ldots, a_m), \qquad \text{where } CH \text{ stands for convex hull.} \tag{12}$$

The key to our translation lies in the following one-to-one correspondence.

$$\text{Let } \triangle \text{ be an } n\text{-element index set } \{\triangle^1, \triangle^2, \ldots, \triangle^n\} \subset \{1, \ldots, m\}. \tag{13}$$

This index set $\triangle$ uniquely defines a point $x_\triangle$ in the primal space as the solution point of the system of equations

$$a_{\triangle^1}^T x = 1, \ldots, a_{\triangle^n}^T x = 1. \tag{14}$$

In $x_\triangle$ exactly n restrictions are active, hence $x_\triangle$ is one of $\binom{m}{n}$ such *basic solutions* of (1). On the other side, $\triangle$ uniquely defines a simplex $CH(a_{\triangle^1}, \ldots, a_{\triangle^n})$ in the dual space. Only the $\binom{m}{n}$ basic solutions are candidates for being shadow vertices of X. For that purpose, $x_\triangle$ must simultaneously satisfy two conditions:

$$1.\ \text{it must be a vertex of } X, \text{which is equivalent to } a_i^T x_\triangle \;\le\; 1\; \forall i \notin \triangle \tag{15}$$

$$2.\ \text{it must optimize some objective } w^T x \text{ on } X, \text{where } w \in Span(u,v) \cap \omega_n \tag{16}$$

Due to the Lemma of Farkas this can equivalently be expressed in the polar space. $CH(a_{\triangle^1}, \ldots, a_{\triangle^n})$ corresponds to a shadow vertex $x_\triangle$, iff simultaneously

$$1.\ CH(a_{\triangle^1}, \ldots, a_{\triangle^n}) \text{ is a facet of } Y \;=\; CH(0, a_1, \ldots, a_m) \tag{17}$$

$$2.\ CH(a_{\triangle^1}, \ldots, a_{\triangle^n}) \,\cap\, Span(u,v) \neq \emptyset. \tag{18}$$

Now we can fully concentrate on Y and count its basic simplices satisfying both

Figure 2: A dual polyhedron Y and the facets that are intersected by $Span(u,v)$. The cut is illustrated by the white line. The left figure shows those background and the right one those foreground facets. Note that $\mathbb{R}^+ v$ and $\mathbb{R}^+ u$ intersect exactly one facet each.

conditions (17),(18). Hence the number of these simplices is S as well. The evaluation of $E_{m,n}(S)$ is simplified by the linearity of expectation-values and the symmetry of index-choices in RSM. If $\triangle$ is an arbitrary index-set as in (13), then

$$P\left(x_\triangle \text{ is a shadow vertex}\right) \;=\; P(CH(a_{\triangle^1}, \ldots, a_{\triangle^n}) \text{ satisfies (17) and (18)})$$

is identical for all $\binom{m}{n}$ candidate-sets $\triangle$. So it is clear that

$$E_{m,n}(S) \;=\; \binom{m}{n} \cdot P(CH(a_1, \ldots, a_n) \text{ satisfies (17) and (18)}). \tag{19}$$

And,we could derive a (rather complicated) integral-formula for that figure. But a direct evaluation (resp. estimation) worked only for the asymptotic case ($m \to \infty, n$

fixed), since in that configuration the calculation-tools are much higher developed. These methods led to the bound of (9). In order to manage an estimation for general dimensions, we used a somehow tricky idea. We compared $E_{m,n}(S)$ with another expected value $E_{m,n}(Z)$, where Z is a random variable closely related to S. This is effective, because $E_{m,n}(Z)$ can be trivially bounded from above.

Let Z denote the number of facets of Y, which are intersected by $\mathbb{R}^+ v$.

It is clear that under nondegeneracy (4) at most one facet of Y will be intersected by $\mathbb{R}^+ v$, hence $E_{m,n}(Z) \leq 1$. This enables us to conclude that

$$E_{m,n}(S) \leq \frac{E_{m,n}(S)}{E_{m,n}(Z)} \quad (\leq J). \tag{20}$$

(20) means that if we find an upper bound J for the quotient, that J will be an upper bound for $E_{m,n}(S)$ as well. Since $E_{m,n}(S)$ is given in a rather complicated, unevaluable integral form, it is much easier to compare the two expectation values, because their integral formuals differ only slightly. The next section will give some details and insight in that comparison.

3 Comparison of Spherical Measures

In this section we study the two kinds of intersection-probabilities and we exploit the assumption of rotation-symmetry in the distribution of v and u.

Lemma 1: *For a fixed simplex* $CH(a_1, \ldots, a_n)$ *we have*

$$\begin{aligned} P(\mathbb{R}^+ v \ intersects\ CH(a_1, \ldots, a_n)) &= \\ \frac{\lambda_n(CC(a_1, \ldots, a_n) \cap \Omega_n)}{\lambda_n(\Omega_n)} &=: V(a_1, \ldots, a_n). \end{aligned} \tag{21}$$

Here λ_k is the k-dimensional Lebesgue-measure, Ω_k is the unit ball of $\mathbb{R}^k$ and $CC(a_1, \ldots, a_n)$ is the convex cone generated by $a_1, \ldots, a_n$ corresponding to the convex hull $CH(a_1, \ldots, a_n)$. The first equation of Lemma 1 tells us that the intersection probability is identical with the share of Ω_n contained in $CC(a_1, \ldots, a_n)$. $V(a_1, \ldots, a_n)$ is a notation for the spherical measure of this sector.
A bit more complicated is our insight on intersection-probabilities with $Span(u, v)$.

Lemma 2: *1) For a fixed* $(n-2)$*-dimensional simplex* $CH(a_1, \ldots, a_{n-1})$ *we have*

$$\begin{aligned} P(Span(u, v)\ intersects\ CH(a_1, \ldots, a_{n-1})) &= \\ \frac{2\lambda_{n-1}\{CC(a_1, \ldots, a_{n-1}) \cap \Omega_n\}}{\lambda_{n-1}\{\Omega_{n-1}\}} &:= 2W(a_1, \ldots, a_{n-1}). \end{aligned} \tag{22}$$

2) Each intersection of $Span(u, v)$ *with* $CH(a_1, \ldots, a_n)$ *produces exactly two intersected simplex sides of type* $CH(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$ *and*

$$P(Span(u, v)\ intersects\ CH(a_1, \ldots, a_n)) = \sum_{i=1}^{n} W(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n). \tag{23}$$

Figure 3: $CC(a_1, a_2)$ intersects $Span(u_2, v_2)$, but not $Span(u_1, v_1)$. The intersection-probability is proportional to the spherical measure of the sector spanned by the two points.

Figure 4: $CC(a_1, a_2, a_3)$ intersects $\mathbb{R}^+ v_2$, but not $\mathbb{R}^+ v_1, \mathbb{R}^+ u_1, \mathbb{R}^+ u_2$. The intersection-probability is proportional to the spherical measure of the sector spanned by the three points.

Here, we introduce a spherical measure W, for a sector generated by $n-1$ points. In our expectation values, where we average also over $a_1, \ldots, a_n$, this reads

$$E_{m,n}(S) = \binom{m}{n} n \int\limits_{\mathbb{R}^n} . \int\limits_{\mathbb{R}^n} P(a_1, ., a_n \text{ induce facet}) \cdot W(a_1, ., a_{n-1}) d\tilde{F}(a_1).d\tilde{F}(a_n) \quad (24)$$

$$E_{m,n}(Z) = \binom{m}{n} \int\limits_{\mathbb{R}^n} . \int\limits_{\mathbb{R}^n} P(a_1, ., a_n \text{ induce facet}) \cdot V(a_1, ., a_n) d\tilde{F}(a_1).d\tilde{F}(a_n) \quad (25)$$

In (24) and (25) $\tilde{F}$ denotes the distribution function of a_i.
So we can write our relation in the following form:

$$\frac{E_{m,n}(S)}{E_{m,n}(Z)} = n \frac{\int\limits_{\mathbb{R}^n} \cdots \int\limits_{\mathbb{R}^n} \cdots W(a_1, \ldots, a_{n-1}) \cdots}{\int\limits_{\mathbb{R}^n} \cdots \int\limits_{\mathbb{R}^n} \cdots W(a_1, \ldots, a_{n-1}) \cdot \frac{V(a_1, \ldots, a_n)}{W(a_1, \ldots, a_{n-1})} \cdots} . \quad (26)$$

It can be regarded as (the inverse of) an expectation value of $\frac{V}{W}$ under the rather strange density function $P(a_1, \ldots, a_n$ induce facet $) \cdot W(a_1, \ldots, a_{n-1})$. This quotient of spherical measures (of different dimensions) plays the crucial role for our analysis. The comparison becomes more evident, when we exploit

$$\frac{V(a_1, \ldots, a_n)}{W(a_1, \ldots, a_{n-1})} = C(n) \cdot \frac{\lambda_n(CC(a_1, \ldots, a_n) \cap \Omega_n)}{\lambda_{n-1}(CC(a_1, \ldots, a_{n-1}) \cap \Omega_n)}, \quad (27)$$

which means that it is proportional to the quotient of Lebesgue measures of two ball-sectors of different dimension. If we are able to derive a *lower* positive bound for *that* expected quotient, then this yields an upper bound for $E_{m,n}(S)$.

4 Old and New Comparison

Let us think about techniques to compare both sector-volumes. The crucial role plays a_n, which appears only in the numerator, whereas $CC(a_1, \ldots, a_{n-1}) \cap \Omega_n$ is measured in the denominator and appears in the numerator as *one side* of the sector.

This comparison was the challenge within the old approach. We found a *lower* bound for the quotient in (27) by replacing the set in the numerator $CC(a_1,\ldots,a_n)\cap\Omega_n$ by a subset $CH(\frac{a_n}{\|a_n\|}, CC(a_1,\ldots,a_{n-1})\cap\Omega_n)$. By the way, the volume decreases. But for the new geometric figure, the *Cavalieri-Principle* delivers a simple formula:

$$\frac{\lambda_n\{CH(\frac{a_n}{\|a_n\|}, CC(a_1,\ldots,a_{n-1})\cap\Omega_n)\}}{\lambda_{n-1}\{CC(a_1,\ldots,a_{n-1})\cap\Omega_n\}} = \frac{1}{n}\cdot dist(\frac{a_n}{\|a_n\|}, H(a_1,\ldots,a_{n-1})), \quad (28)$$

where $H(a_1,\ldots,a_{n-1})$ stands for the hyperplane containing $0,a_1,\ldots,a_{n-1}$ and $dist(\frac{a_n}{\|a_n\|},H)$ gives the distance of $\frac{a_n}{\|a_n\|}$ to that hyperplane.
It simply remains to evaluate the average value of that distance, as in [2] and [3].

Figure 5: The spherical triangle and the spherical sector generated by a_1,a_2,a_3

Figure 6: Underestimation of the spherical sector with the Cavalieri-Principle

But we should be aware of the fact that this replacement led to a dramatic underestimation of V and of the relation in (27), because we now ignore the curvature of our sector between $\frac{a_n}{\|a_n\|}$ and the ground area completely. This explains the loss of a factor $\sqrt{n}$ in the old result. The underestimation would be harmless if all n points were close together. This is a typical feature of facets in the "asymptotic case" ($m\to\infty$, n fixed), providing a reason for the better result in (8).

Our new approach took the curvature into account. While doing this, we lost a strong advantage of Cavalieri's principle. It is so simple, because in (28) the extension factor $\frac{V}{W}$ exclusively depends on the distance of $\frac{a_n}{\|a_n\|}$ to the hyperplane $H(a_1,\ldots,a_{n-1})$. But for a precise calculation of the curvature, the relative position of $\frac{a_n}{\|a_n\|}$ to $CC(a_1,\ldots,a_{n-1})$ must be fully specified, because $\frac{V}{W}$ strongly depends on the complete location. This drawback forces us to introduce additional levels of integration in our multiple integral formulas.
Besides that we had to use sharper estimates for the marginal density of the uniform distribution on ω_n. Combining this with the evaluation of the extended integration formulas yields a sharper bound on $\frac{E_{m,n}(S)}{E_{m,n}(Z)}$ and saves a (second) factor $\sqrt{n}$.

5 Comparison of Expectations of Sums of Variables

For explaining another saving of $\sqrt{n}$, we enter the proof in [4] at an inequality

$$\frac{E_{m,n}(S)}{E_{m,n}(Z)} \le C(n)\cdot\frac{\int_t^1\int_0^t G(h)^{m-n}\,\frac{R^{n-3}}{r^{n-2}T}\,\left[T+\frac{1}{\sqrt{n}}R\right]\,dh\,dF(r)}{\int_t^1\int_0^t G(h)^{m-n}\,\frac{R^{n-3}}{r^{n-2}T}\,\left[\frac{1}{r}\left(T^2+\frac{1}{\sqrt{n}}R^2\right)\right]\,dh\,dF(r)}. \quad (29)$$

Here we find two integration variables h and r and a fixed threshold t such that

$$0 \leq h \leq t \leq r \leq 1, \;\; T := T(h) = \sqrt{t^2 - h^2}, \;\; R := R(h) = \sqrt{r^2 - h^2}. \tag{30}$$

Note that $R(h) \geq T(h)$ for all $h \in [0, t]$. F denotes the distribution function of r and G represents a marginal distribution function, evaluated at h.
Essentially, (29) can be seen as a comparison between expected values of $\left[T + \frac{1}{\sqrt{n}} R\right]$ and $\left[\frac{1}{r}\left(T^2 + \frac{1}{\sqrt{n}} R^2\right)\right]$, when we regard $G(h)^{m-n}\frac{R^{n-3}}{r^{n-2}T}$ as a "density function". But it is hard to compare precisely, as long as the objectives in numerator and denominator are sums of random variables. Our first idea to avoid the sums is to replace

$$T + \frac{1}{\sqrt{n}} R \;\text{ by }\; \text{Max}\,\{T, \frac{1}{\sqrt{n}} R\} \cdot 2 \quad \text{and} \quad T^2 + \frac{1}{n} R^2 \;\text{ by }\; \text{Max}\,\{T^2, \frac{1}{n} R^2\}. \tag{31}$$

By the way, we have increased the numerator and decreased the denominator, hence we achieve an upper bound for (29). But now we have to determine the maximum of the two terms. Unfortunately, the dominance between $T(h) = \sqrt{t^2 - h^2}$ and $\frac{1}{\sqrt{n}}R(h) = \frac{1}{\sqrt{n}}\sqrt{r^2 - h^2}$ may switch for fixed r at a value h_r (depending on r). So, the dominance principle requires a partition of the inner integrals, where the bounds depend on the outer integration variable r. That overcomplicates an evaluation.
To avoid partitions, we made a crude estimation in [3]. We respectively replaced

$$T + \frac{1}{\sqrt{n}} R \;\text{ by }\; 2\text{Max}\{T, R\} = 2R \;\text{ and }\; T^2 + \frac{1}{n} R^2 \;\text{ by }\; \frac{1}{n}\text{Max}\left\{T^2, R^2\right\} = \frac{1}{n} R^2 \tag{32}$$

Since this time the R-term dominates consistently, in (29) the essential quotient

$$\text{changes from} \quad \frac{T + \frac{1}{\sqrt{n}} R}{\frac{1}{r}\left[T^2 + \frac{1}{\sqrt{n}} R^2\right]} \quad \text{to} \;\; n \cdot 2 \cdot \frac{r}{R}\,. \tag{33}$$

But we observe the appearance of the factor n very sceptically, because the left quotient in (33) seems to justify at most a factor of $\sqrt{n}$ (if $R >> T$).
After (32) and (32) we had in [3] developed an estimation technique for the quotient

$$C(n)2n\frac{\int_t^1\int_0^t G(h)^{m-n}\frac{R^{n-3}}{r^{n-2}T}[R]\,dh\,dF(r)}{\int_t^1\int_0^t G(h)^{m-n}\frac{R^{n-3}}{r^{n-2}T}\left[\frac{R^2}{r}\right]\,dh\,dF(r)} = C(n)2n\frac{\int_t^1\int_0^t G(h)^{m-n}\frac{R^{n-2}}{r^{n-2}T}\,dh\,dF(r)}{\int_t^1\int_0^t G(h)^{m-n}\frac{R^{n-2}}{r^{n-2}T}\left[\frac{R}{r}\right]\,dh\,dF(r)} \tag{34}$$

The main step in that derivation was an application of Jensen's inequality for an integral quotient for fixed values of r $(r > t)$ and an arbitrary value $\xi \in (0, t)$.

$$\frac{\int_\xi^t \frac{1}{T}\frac{R^{n-2}}{r^{n-2}}\,dh}{\int_\xi^t \frac{1}{T}\frac{R^{n-1}}{r^{n-1}}\,dh} \leq \left[\frac{\int_\xi^t \frac{1}{T}\,dh}{\int_\xi^t \frac{1}{T}\frac{R^{n-1}}{r^{n-1}}\,dh}\right]^{\frac{1}{n-1}} \leq [n]^{\frac{1}{n-1}}\;\frac{r}{\sqrt{r^2 - \xi^2}}\,. \tag{35}$$

Now we explain essentials of the **new approach**. Instead of determining dominators as in (31), (32), we bound the denominator from below and insert into (29)

$$\left[T^2 + \frac{1}{n}\,R^2\right] \;\geq \frac{1}{2}\;\frac{1}{\sqrt{n}}\,R\left[T + \frac{1}{\sqrt{n}}\,R\right]. \tag{36}$$

$$\frac{E_{m,n}(S)}{E_{m,n}(Z)} \;\leq\; C(n)\,2\sqrt{n}\;\frac{\int\limits_t^1\int\limits_0^t G(h)^{m-n}\,\frac{R^{n-3}}{r^{n-2}\,T}\;\left[T + \frac{1}{\sqrt{n}}\,R\right]\;dh\,dF(r)}{\int\limits_t^1\int\limits_0^t G(h)^{m-n}\,\frac{R^{n-3}}{r^{n-2}\,T}\;\left[T + \frac{1}{\sqrt{n}}\,R\right]\;\{\frac{R}{r}\}\,dh\,dF(r)} \tag{37}$$

This time the extracted factor is only $\sqrt{n}$, but now we have a joint factor $\left[T + \frac{1}{\sqrt{n}}\,R\right]$ in both integral formulas. This factor can be regarded as part of the "density". The "objective variable" $\frac{R}{r}$ is the same as in (34). The remaining question is whether the "expected value" of $\frac{R}{r}$ will change dramatically under that modification of the "density function". For checking this, we employ Jensen's inequality and show

$$\frac{\int\limits_\xi^t \frac{1}{T}\frac{R^{n-3}}{r^{n-2}}\left[T + \frac{1}{\sqrt{n}}\,R\right]\;dh}{\int\limits_\xi^t \frac{1}{T}\frac{R^{n-3}}{r^{n-2}}\;\left[T + \frac{1}{\sqrt{n}}\,R\right]\;\frac{R}{r}\,dh} \;\leq\; \left[1 + \frac{1}{\sqrt{n}}\right]^{\frac{1}{n-1}}\left[\frac{\int\limits_\xi^t \frac{1}{T}\,dh}{\int\limits_\xi^t \frac{1}{T}\;\frac{R^{n-1}}{r^{n-1}}\;\left[\frac{T}{R} + \frac{1}{\sqrt{n}}\right]\;dh}\right]^{\frac{1}{n-1}} \leq$$

$$\leq\; \left[n(\sqrt{n}+1)\right]^{\frac{1}{n-1}}\cdot\left[\frac{r}{\sqrt{r^2-\xi^2}}\right]^{\frac{1}{n-1}}. \tag{38}$$

With this intermediate result we are allowed to enter the old proof and to proceed in the known manner. Note that (38) resembles (35) very closely. New is only the factor $(\sqrt{n}+1)^{\frac{1}{n-1}}$, which can be bounded from above by a small constant. So the saving of $\sqrt{n}$ between (33) and (36) can be pulled through the complete proof.
The consequence of the two (factor $\sqrt{n}$)-savings (described in sections IV and V) is

Theorem 1:
For all RSM-distributions and all (m,n) pairs with $m \geq n$ we know that the expected number of shadow vertices satisfies

$$E_{m,n}(S) \;\leq\; Const.\cdot m^{\frac{1}{n-1}}\,n^2. \tag{39}$$

6 The Average Number of Steps in the Complete Algorithm

So far, we dealt only with Phase II. Now consider the complete method for solving the LP (1). Let us denote

$$\Pi_k(x) \;:=\; \begin{pmatrix} x^1\\ \vdots\\ x^k\end{pmatrix} \;\text{ for } x = \begin{pmatrix} x^1\\ \vdots\\ x^n\end{pmatrix} \in \mathbb{R}^n. \tag{40}$$

That means that Π_k is the orthogonal projection of $\mathbb{R}^n$ onto $\mathbb{R}^k$.

And let I_k be the following LP in k variables ($k = 1, \dots, n$).

$$\begin{array}{lrcl} \text{maximize} & \Pi_k(v)^T\Pi_k(x) & & \\ \text{subject to} & \Pi_k(a_i)^T\Pi_k(x) & \leq & 1 \text{ for } i = 1, \dots, m \\ \text{where} & x, v, a_1, \dots, a_m \in \mathbb{R}^n & & \text{and } m \geq n. \end{array} \tag{41}$$

X_k will be called the feasible polyhedron of I_k. Note that I_n coincides with the original LP (1). For a complete solution we may apply the following algorithm.

Dimension-By-Dimension-Algorithm

1. Set $k = 1$ and find a vertex of X_1.
2. If existing, find the optimal vertex $(\hat{x}^1)$ of I_1 . Else go to 6).
3. If $k = n$, then go to 7). Set $k = k + 1$.
4. For a $k \in \{2, \dots, n\}$ the solution of I_{k-1} may be available. We call it $\begin{bmatrix} \hat{x}^1 \\ \vdots \\ \hat{x}^{k-1} \end{bmatrix}$. Then $\begin{bmatrix} \hat{x}^1 \\ \vdots \\ \hat{x}^{k-1} \\ 0 \end{bmatrix}$ is feasible for I_k on an edge of X_k. Determine $\begin{bmatrix} \tilde{x}^1 \\ \vdots \\ \tilde{x}^k \end{bmatrix}$, which denotes a vertex on that edge.
5. Use $Span\,(\Pi_k(e_k), \Pi(v))$ as projection plane and start the shadow vertex algorithm in $\begin{bmatrix} \tilde{x}^1 \\ \vdots \\ \tilde{x}^k \end{bmatrix}$ to find an optimal vertex $\begin{bmatrix} \hat{x}^1 \\ \vdots \\ \hat{x}^k \end{bmatrix}$ for I_k. Go to 3) if $\hat{x}$ exists.
6. A solution of the complete problem does not exist. STOP
7. The vector $\hat{x} \in \mathbb{R}^n$ is the solution of I_n (the original problem). STOP

Note that all input vectors of I_k in (41), namely $\Pi_k(a_i), \Pi_k(v)$, are distributed according to RSM. So the theory of sections I - V applies to each use of the shadow-vertex-algorithm in step 5) with u replaced by e_k. For counting the total number of pivot steps s_t, we simply have to cumulate over all $n-1$ applications of the shadow vertex algorithm and to add the $n-1$ vertex exchanges in step 4). A simple upper bound would result from summation over all upper bounds given in Theorem 1. But that would give a false impression, because we would ignore that most of the RSM-distributions over $\mathbb{R}^k$ are no possible projection-distributions from $\mathbb{R}^n$ after applying Π_k. Instead we have to deal here with a subset of the distributions studied for Theorem 1. In [3] it had been shown, that for the set of projection-distributions over $\mathbb{R}^k$ there is a better upper bound on the expected number of shadow vertices of I_k. So, we study $E^k_{m,n}(S)$, the average number of shadow-vertices for those distributions. The case $k = 1$ is trivial, and already in [3] the case $k = 2$ had been clarified with

$$E^2_{m,n}(S) \;\leq\; 4\, m^{\frac{1}{n-1}} \cdot n. \tag{42}$$

Also in [3], it had been shown that for $k = 3, \dots, n$

$$E^k_{m,n}(S) \;\le\; Const.\, m^{\frac{1}{n-1}} \cdot k^{\frac{3}{2}} \cdot n^{\frac{3}{2}}. \tag{43}$$

Hence the old upper bound for $E_{m,n}(s_t)$ had been $E_{m,n}(s_t) \le Const.\, m^{\frac{1}{n-1}} \cdot n^4$.

The saving of a factor n in Theorem 1 (the analysis for stage n) suggests that such a saving may even be possible for each stage. We want to clarify this conjecture and deal with the stages $k = 3, \dots, n$. For that purpose, we simulate the analysis of sections I to V, regarding that our k-distribution has a root in $\mathbb{R}^n$. First, we introduce some notation comparing the configurations in $\mathbb{R}^n$ and $\mathbb{R}^k$. Let f^n be the original density of the a_i-distribution over $\mathbb{R}^n$. Correspondingly, we define a density f^k over $\mathbb{R}^k$, which is the density for the Π_k-projected vectors

$$f^k(x^1,\dots,x^k) \;=\; \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f^n(x^1,\dots,x^k,\eta^{k+1},\dots,\eta^n)\, d\eta^{k+1}\dots d\eta^n. \tag{44}$$

An important insight concerns marginal distributions on the first coordinate. They are not affected by the transfer from $x \in \mathbb{R}^n$ to $\Pi_k(x) \in \mathbb{R}^k$. In particular

$$G_n(h) \;:=\; P_n[x^1 \le h] \;=\; P_k\left[(\Pi_k(x))^1 \le h\right] \;=:\; G_k(h) \text{ for } h \in \mathbb{R}, \tag{45}$$

where P_n is the probability over $\mathbb{R}^n$ and P_k is a probability over $\mathbb{R}^k$ (based on the distribution of the k-projected vectors). $G_n(h)$ and $G_k(h)$ denote the respective marginal distribution functions. The simple reason for (45) is that the first coordinate of x coincides with the first coordinate of $\Pi_k(x)$.
Now we enter the proof of [4], resp. of (39) at an appropriate inequality. Our k-(projection)-distribution satisfies all RSM conditions for general $\mathbb{R}^k$-distributions. In stage k we have $\binom{m}{k}$ candidates for being a shadow vertex. In our inequality n can recklessly be replaced by k.

$$E^k_{m,n}(S) \le Const. \cdot k^{\frac{3}{2}} \cdot \frac{\int_0^t G_k(h)^{m-k}\, T^{-1} \int_{\mathbb{R}^{k-1}} |T - c^{k-1}| f^k(c)\, d\bar{c}\, dh}{\int_0^t G_k(h)^{m-k} T^{-1} \frac{h}{t} \int_{\mathbb{R}^{k-1}} |T - c^{k-1}|^2 \, \frac{\psi(h,t,\|c\|_k)}{\|c\|_k}\, f^k(c)\, d\bar{c}\, dh} \tag{46}$$

We use $\bar{c} = (c^1,.,c^{k-1})^T$, $\|c\|_k := \sqrt{(c^1)^2 + . + (c^k)^2}$, G_k as in (45); h, t, T as in (30),

$$\text{and} \quad \Psi(h,t,\|c\|_k) \;:=\; \frac{1}{\frac{hh}{t\cdot\|c\|_k} + \left(1 - \frac{hh}{t\|c\|_k}\right)\frac{1}{\sqrt{k}}}\,. \tag{47}$$

We exploit that this inequality for a k-dimensional figure is based on a generation of input data in $\mathbb{R}^n$ reflecting an n-dimensional distribution. This is partly done by translating and substituting the k-terms into the n-terms and partly by showing that a replacement of k-terms by corresponding n-terms increases our bound.
As explained in (45), we can replace $G_k(h)$ by $G_n(h)$. Besides, the integration over

$\overline{c} \in \mathbb{R}^{k-1}$ with density $f^k(c)$ may be substituted by an integration over $\overline{\eta} \in \mathbb{R}^{n-1}$ with density $f^n(\eta)$ over $\mathbb{R}^n$, when we define $\overline{\eta} = (\eta^1, \ldots, \eta^{n-1})^T$. This reflects the definition of f^k in (44). And we can replace $|T - c^{k-1}|$ by $|T - \eta^{k-1}|$ and go on to $|T - \eta^{n-1}|$, since $f^n(\eta)$ is invariant under permutations of coordinates.

$$\text{Finally, let us have a look at} \quad \frac{\Psi(h,t,\|c\|_k)}{\|c\|_k}. \tag{48}$$

Here a direct substitution is not possible, but we know that for a vector $c \in \mathbb{R}^n$

$$\|c\|_k \;=\; \sqrt{(c^1)^2 + \ldots + (c^k)^2} \;\leq\; \sqrt{(c^1)^2 + \ldots + (c^n)^2} \;=\; \|c\|_n. \tag{49}$$

This allows the following estimation

$$\frac{\Psi(h,t,\|c\|_k)}{\|c\|_k} \;\geq\; \frac{\sqrt{k}}{\frac{hh}{t}\left(1 - \frac{1}{\sqrt{n}}\right)\sqrt{n} + \|c\|_n} \;=\; \frac{\sqrt{k}}{\sqrt{n}} \cdot \frac{\Psi(h,t,\|c\|_n)}{\|c\|_n}. \tag{50}$$

The replacement, admitted by (50), leads us to

$$E^k_{m,n}(S) \leq Const. \cdot k^{\frac{3}{2}} \cdot \frac{\sqrt{n}}{\sqrt{k}} \cdot \frac{\int\limits_0^t G_n(h)\, T^{-1} \int\limits_{\mathbb{R}^{n-1}} |T - \eta^{n-1}|\, f^n(\eta)\, d\overline{\eta}\, dh}{\int\limits_0^t G_n(h) T^{-1} \int\limits_{\mathbb{R}^{n-1}} |T - \eta^{n-1}|^2\, \frac{\Psi(h,t,\|c\|_n)}{\|c\|_n} f^n(\eta)\, d\overline{\eta}\, dh} \tag{51}$$

This integral quotient had been under investigation in [4], which yields

$$E^k_{m,n}(S) \;\leq\; Const. \cdot k^{\frac{3}{2}} \cdot \frac{\sqrt{n}}{\sqrt{k}} \cdot m^{\frac{1}{n-1}} \cdot \sqrt{n} \;= Const.\, k \cdot n \cdot m^{\frac{1}{n-1}}. \tag{52}$$

Summing up over all stages of k, shows that the saving of two factors $\sqrt{n}$ (as in the derivation of an upper bound for $E_{m,n}(S)$) works for the complete algorithm as well.

Theorem 2:
For all RSM-distributions and for all dimension pairs (m,n) with $m \geq n$ the Dimension-By-Dimension-Algorithm requires on the average for solving the LP (1) completely $E_{m,n}(s_t)$ pivot steps, where

$$E_{m,n}(s_t) \;\leq\; Const.\, m^{\frac{1}{n-1}} \cdot n^3\,.$$

References

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

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

[3] Borgwardt, K.H. (1987): *The Simplex Method – A Probabilistic Analysis*, Springer-Verlag, Heidelberg, 1987.

[4] Borgwardt, K.H. (1997): *A Sharp Upper Bound For The Expected Number Of Shadow-Vertices In The Rotation-Symmetry-Model*, submitted to Mathematics of Operations Research, 1997.