A Lower Bound on the Average Number of Pivot-Steps for Solving Linear Programs – Valid for all Variants of the Simplex-Algorithm

Karl Heinz Borgwardt, Petra Huhn

Universität Augsburg, Institut für Mathematik, Universitätsstr. 14, 86159 Augsburg
email: borgwardt@math.uni-augsburg.de, huhn@math.uni-augsburg.de

Abstract

In this paper we derive a lower bound on the average complexity of the Simplex-Method as a solution-process for linear programs (LP) of the type:

$$\text{maximize } v^T x \text{ subject to } a_1^T x \leq 1, \ldots, a_m^T x \leq 1.$$

We assume these problems to be randomly generated according to the Rotation-Symmetry-Model:

Let $a_1, \ldots, a_m, v$ be distributed independently, identically and symmetrically under rotations on $\mathbb{R}^n \setminus \{0\}$.

We concentrate on distributions over $\mathbb{R}^n$ with bounded support and we do our calculations only for a subfamily of such distributions, which provides computability and is representative for the whole set of these distributions.
The Simplex-Method employs two phases to solve such an LP. In Phase I it determines a vertex x_0 of the feasible region – if there is any. In Phase II it starts at x_0 to generate a sequence of vertices $x_0, \ldots, x_s$ such that successive vertices are adjacent and that the objective $v^T x_i$ increases. The sequence ends at a vertex x_s which is either the optimal vertex or a vertex exhibiting the information that no optimal vertex can exist. The precise rule for choosing the successor-vertex in the sequence determines a variant of the Simplex-Algorithm.
We can show for every variant, that the expected number of steps s^{var} for a variant, when m inequalities and n variables are present, satisfies

$$E_{m,n}[s^{var}] \geq const.\, m^{\frac{1}{n-1}} n^0 \text{ for all pairs } m \geq n \text{ and for all variants.}$$

This result holds, if the selection of x_0 in Phase I has been done independently of the objective v.

Keywords
Linear Programming, Average-Case Complexity of Algorithms, Stochastic Geometry
AMS-Classification 90C05, 68Q25, 60D05

1 The Problem and its Significance

1.1 Notation

We study the average number of pivot steps required by the Simplex-Method in the solution-process of a linear program (LP):

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

We are interested in stochastic results and therefore we employ the following stochastic model:

$$\begin{aligned} &\text{Rotation-Symmetry-Model (RSM):} \\ &a_1, \ldots, a_m \text{ and } v \text{ are distributed independently,} \\ &\text{identically and symmetrically under rotations on } \mathbb{R}^n \setminus \{0\}. \end{aligned} \tag{2}$$

These stochastic assumptions have the side-effect of giving the set of so-called "degenerate" problems only probability zero. A problem of type (1) is called nondegenerate, if:

$$\begin{aligned} &\text{Each subset of } n \text{ elements out of } \{a_1, \ldots, a_m, v\} \text{ is} \\ &\text{linearly independent and each subset of } (n+1) \\ &\text{elements out of } \{a_1, \ldots, a_m\} \text{ is in general position.} \end{aligned} \tag{3}$$

A single distribution over $\mathbb{R}^n \setminus \{0\}$ according to (2) can be characterized by its distribution function $\hat{F}$. Without loss of generality (w. l. o. g.) we can assume that $\hat{F}$ has a density $\hat{f}$. But the assumption of rotation-symmetry even admits a simpler characterization via the radial distribution function

$$F(r) := P(x \mid \|x\| \le r) = \int_{\mathbb{R}^n} I(\|x\| \le r)\, d\hat{F}(x)\,,$$

where $I(\,.\,)$ denotes the indicator function of an event.

Our analysis will rely on the condition that the distributions under investigation have bounded support. W. l. o. g. we can simplify this restriction to the condition that the support is contained in the unit ball $\Omega_n := \{x \in \mathbb{R}^n \mid \|x\| \le 1\}$. The unit sphere on $\mathbb{R}^n$ will be denoted by $\omega_n := \{x \in \mathbb{R}^n \mid \|x\| = 1\}$.

Moreover, we concentrate on a certain family of these distributions with bounded support (with the property that $P(\|x\| \le r)$ decreases algebraically, when r approximates 1, see [5, 10, 11]). The reason for investigating these distributions lies in the good evaluability of many integrals appearing in the probabilistic analysis.

Specifically, we deal with distributions whose radial distribution function is controlled by a parameter $z \in (-1, \infty)$ in the form

$$F_z(r) := \begin{cases} \dfrac{\displaystyle\int_0^r (1-\tau^2)^z \tau^{n-1} d\tau}{\displaystyle\int_0^1 (1-\tau^2)^z \tau^{n-1} d\tau} & \text{for} \quad 0 \le r \le 1\,, \\ 1 & \text{for} \quad r > 1\,. \end{cases} \tag{4}$$

The parameter z can be seen as a centralization parameter, since an increment in z will move more and more weight to the interior of Ω_n.
Prominent special cases for the choice of z are

$z = 0$	uniform distribution on Ω_n,
$z \to -1$	uniform distribution on ω_n,
$z \to \infty$	extremal concentration of weight about the origin,
$z = \frac{n-1}{2}$	symmetric concentration about the central radius $\frac{1}{2}$.

The corresponding radial density function f_z is

$$f_z(r) := \begin{cases} \dfrac{(1-r^2)^z r^{n-1}}{\displaystyle\int_0^1 (1-\tau^2)^z \tau^{n-1} d\tau} & \text{for } 0 \le r \le 1, \\ 0 & \text{elsewhere.} \end{cases} \tag{5}$$

Simplex-Methods solve problems of type (1) in two phases.

Phase I: Determination of a vertex x_0 of the feasible region $X := \{x \in \mathbb{R}^n \,|\, a_1^T x \le 1, \dots, a_m^T x \le 1\}$ *or* confirmation that X has no vertices.

Phase II: Construction of a sequence $x_0, \dots, x_s$ of vertices of X, such that successive vertices x_i, x_{i+1} are adjacent and that $v^T x_i < v^T x_{i+1}$ for $i = 0, \dots, s-1$. The sequence ends at x_s, if (6)

- x_s is the optimal vertex *or*
- at x_s the nonexistence of an optimal vertex becomes evident.

1.2 Complexity of Simplex-Methods

As we are interested in a lower bound on the (average) complexity, it suffices to derive a lower bound on the number s (the number of pivot steps in Phase II). Besides that, Phase I can be done in a similar manner to Phase II. So we are allowed to concentrate on Phase II. We rely on an application of Phase I, which delivers a unique vertex x_0 for a given polyhedron X, independently of the current choice of v. (That means that v is seen as a specific input for Phase II and does not enter the vertex-searching process.)

Since the definition of Phase II in (6) does not specify a formula or rule for choosing the successor vertex to an iteration vertex x_i, we have the freedom to carry out this construction in several ways. Such a unique rule will define a variant of the Simplex-Algorithm.
Prominent variants (in the huge set of possible variants) are e. g.

- the Rule of Greatest Improvement
- the Rule of Steepest Ascent
- the Dantzig Rule
- the Shadow-Vertex Algorithm (or parametric rule)
- the Rule of Random Choice
- Bland's Rule.

If we apply such a rule, then the effort for carrying out one iteration is simple to recognize. Unclear is the number s. So the natural question for the complexity of single variants concerns the value s^{var} or $E_{m,n}[s^{var}]$, where var indicates the specific variant under use.
Variants as the above mentioned examples implement a fixed rule for making a choice between the possible successor vertices or the incident edges, which improve the objective. For this decision different strategies with different effort can be employed. The two first variants study the full shape of the polyhedron and decide after that amongst the available edges (in the style of a local optimization). The two next variants calculate only forecast-values for the progress in the objective locally at the current vertex and decide for the best. And the two last do not optimize (locally) at all, but base their decision on combinatorial methods (handling the original indices) only. Knowing this, one should expect that the first two are superior to the third and fourth and that the fifth and sixth variant

are considerably worse. These judgements are somehow confirmed by numerical tests as in [7].

Only for one of these variants a satisfying upper bound for the average number of s could be derived so far. This is the Shadow-Vertex-Algorithm (SVA). Here it could be proven that the expected number of pivot steps is polynomial in m and n [3, 4] and that the upper bound [6] is as follows

$$E_{m,n}[s^{\text{SVA}}] \leq const.\, m^{\frac{1}{n-1}} n^2$$

for all m, n and for all distributions of the RSM.

This bound is sharp in the sense, that for the uniform distribution on the unit sphere there is an asymptotic (n fixed, $m \to \infty$) lower bound for $E_{m,n}[s^{\text{SVA}}]$ of the type $const.\, m^{\frac{1}{n-1}} n^2$, see [4].
Another result concerning lower bounds on the average number of steps of the Simplex-Method has been derived by Adler and Megiddo [1] for a parametric variant of the Simplex-Algorithm, but their analysis is based on different stochastic assumptions and therefore the results are hardly comparable with the above mentioned analysis.

The reason for the special role of the Shadow-Vertex-Algorithm and its qualification for a probabilistic analysis lies in the fact that it is possible to give a purely geometrical characterization of the event "a certain vertex is on the Simplex-Path" using only the active constraints and two objective directions. It is – for this variant – not necessary to know the preceding part of the Simplex-Path. This would be essential for other variants and makes their evaluation by far overcomplicated.

This situation leads to the question for the representativity of the Shadow-Vertex-Algorithm results for general variants. Our lower bound derivation for the distribution family F_z gives a partial answer.
For simplification and a better comparability we only discuss the results for $z \to -1$. Here, we can show that – under the condition that Phase I worked independently of v and provided us with a vertex x_0 of X – the following lower bound holds for every variant:

$$E_{m,n}[s] \geq const.\, m^{\frac{1}{n-1}(1-\varepsilon(n))} n^{-\delta(n)} \tag{7}$$

where $\varepsilon(n)$ and $\delta(n)$ tend to 0 for $n \to \infty$. Essentially this means that we have a growth as

$$m^{\frac{1}{n-1}} n^0.$$

So it can be stated that our analysis still leaves a gap for better variants (from $m^{\frac{1}{n-1}}n^2$ to $m^{\frac{1}{n-1}}n^0$). But this gap is not too dramatic. Note that the factor n^0 can in the case $m \gg n$ not be regarded as a triviality, since in each vertex up to n incident edges have a chance for providing a significant progress (increment) of the objective. So we would not have been astonished, if we had observed a negative exponent describing the n-dependency as result of the exploitation of an increasing number of options. And, now it is clear that under the RSM no variant can surpass the SVA with respect to the m-dependency.
In addition, note that our result is true for every m and n, but that it yields a nontautological statement only for $m \gg n$, because the constant in (7) may be very small.

The technique used in this paper uses the fact that a certain distance between x_0 and x_s has to be bridged in single applications of the Simplex-variant. Then we can show that – without regard to the variant – there are not enough "long" edges to construct a path with few vertices (traversing the required distance). This geometrical idea leads to a lower bound on the average number of steps – independent of the variant – and will be elaborated in section 2. In detail, our result can technically be described as follows.
We apply the Markov-Inequality (to integrals appearing in the calculation of the lower bound) with varying parameter $\rho \in [2, \infty)$ and we can show that for each such ρ and each such z (the parameter for our distribution family) we have

$$\begin{aligned} E_{m,n}[s+1] \quad \geq \quad & const.\, m^{\frac{1}{n+1+2z} - \frac{n-2}{\rho(n+1+2z)^2}} \\ & \cdot (2z+n+2)^{-\frac{n-1}{2\rho(n+1+2z)}} (n+\rho)^{-\frac{1}{n+1+2z} + \frac{2n+2z-1}{\rho(n+1+2z)^2}}, \end{aligned}$$

where the constant is independent of n, m, z, ρ. This lower bound will be proven in section 3.

2 Geometrical Background

2.1 Observing the Walk on a Simplex-Path

Our nondegeneracy condition ascertains some fundamental properties for LPs, which will be very useful for our investigation. To describe them, we introduce the notation

$$\Delta := \{\Delta^1, \ldots, \Delta^n\} \subset \{1, \ldots, m\}$$

for an (arbitrary) n-element index set of $\{1, \ldots, m\}$. Upper indices denote different elements. In case that different Δ-sets have to be discussed simultaneously, we use lower indices as Δ_1, Δ_2 for distinction.

Remark 1

1. *Each Δ-set induces a basic solution of our LP (1), i. e. the solution of a system of n equations $a_{\Delta^1}^T x = 1, \ldots, a_{\Delta^n}^T x = 1$. We call this (unique) point x_Δ. Because of nondegeneracy, $a_j^T x_\Delta \neq 1 \ \forall j \notin \Delta$.*
   Such a basic solution is a vertex of X if and only if $a_j^T x_\Delta \leq 1 \ \forall j \notin \Delta$ (i. e. all nonbasic constraints are satisfied).

2. *If x_Δ is a vertex of X, then there are different $(n-1)$-element subsets $\{\Delta^1, \ldots, \Delta^{i-1}, \Delta^{i+1}, \ldots, \Delta^n\}$ and each such subset induces a line (affine subspace) $\{x \in \mathbb{R}^n \mid a_{\Delta^1}^T x = 1, \ldots, a_{\Delta^{i-1}}^T x = 1, a_{\Delta^{i+1}}^T x = 1, \ldots, a_{\Delta^n}^T x = 1\}$, containing an edge of X, which is incident to x_Δ.*

3. *Edges incident to a vertex x_Δ are either rays starting at x_Δ or line segments between x_Δ and another vertex $x_{\Delta_\star}$, where $\#(\Delta \cap \Delta_\star) = n-1$. That means that $n-1$ constraints are active in the interval $(x_\Delta, x_{\Delta_\star})$ and in x_Δ resp. $x_{\Delta\star}$ one additional restriction becomes active (not the same). In this case x_Δ and $x_{\Delta_\star}$ are adjacent on X.*

4. *No two adjacent vertices can have equal objective value.*

5. *If and only if an optimal vertex x_Δ exists, then $v \in \operatorname{cone}(a_{\Delta^1}, \ldots, a_{\Delta^n})$, i. e. v is an element of the polar cone of x_Δ.*
   If and only if an optimal vertex does not exist, then $v \notin \operatorname{cone}(a_1, \ldots, a_m)$.

From the elementary theory of the Simplex-Algorithm it is known that the construction in Phase II delivers a connected path consisting of edges

$$\overline{x_0 x_1}, \overline{x_1 x_2}, \ldots, \overline{x_{s-1} x_s}\,. \tag{8}$$

At the stopping vertex x_s, two outcomes are possible.

1. x_s is the optimal vertex with index set Δ_s. This will be confirmed in the current tableau by giving a conical representation for v by $a_{\Delta_s^1}, \ldots, a_{\Delta_s^n}$.

2. x_s is not optimal, but the current tableau shows that v cannot belong to cone$(a_1, \ldots, a_m)$ resp. that no pivot element exists. Geometrically this amounts to the exhibition of an unbounded edge incident to x_s which improves the objective $v^T x$.

It is our aim to quantify (or at least to bound from below) the unknown number s in all potential cases. For this purpose we should handle both stopping cases in a compatible way. And here it is important that in case 1 – bounded objective – every variant working on a given polyhedron X and starting at x_0 leads to the same stopping vertex $x_s = x_{opt}$. But in case 2 – unbounded objective – a start at x_0 and the use of different variants may provide a variety of stopping vertices x_s. So it is desirable to face every variant with the same challenge. And we want to achieve as much conformity as possible in the way how we treat bounded and unbounded problems.

Our expedient for that difficulty is the following.
We augment our Simplex-Path (8) with a final move from x_s to v. This move can be interpreted as an additional "edge" $\overline{x_s v}$. So our complete walk is now

$$\overline{x_0 x_1}, \overline{x_1 x_2}, \ldots, \overline{x_{s-1} x_s}, \overline{x_s v}\,.$$

Hence, bridging the way from x_0 to v is the joint challenge for each variant, no matter whether $v^T x$ is bounded from above or not. And – by the way – we have now achieved a certain symmetry between x_0 and v.
For this interpretation we equate and identify (symbol $\widehat{=}$) all paths from x_0 to v – unregarded the different vertices on those paths and the different lengths. The common characteristic is the starting vertex x_0 and the endpoint v. So, we can describe our walk as the sum of the movements, e. g.

$$\overline{x_0 v} \;\widehat{=}\; \sum_{i=0}^{s} \overline{x_i x_{i+1}} \qquad \text{where we set } x_{s+1} = v\,. \tag{9}$$

2.2 Measuring the length of a Simplex-Path

Since our (augmented) path is connected, we may measure the sum of piecewise moves that we make and compare the result (the total distance) with the direct distance from x_0 to v (when the move is made straightforward). Whenever we use a distance measure satisfying the triangle inequality, then the direct distance will be smaller.

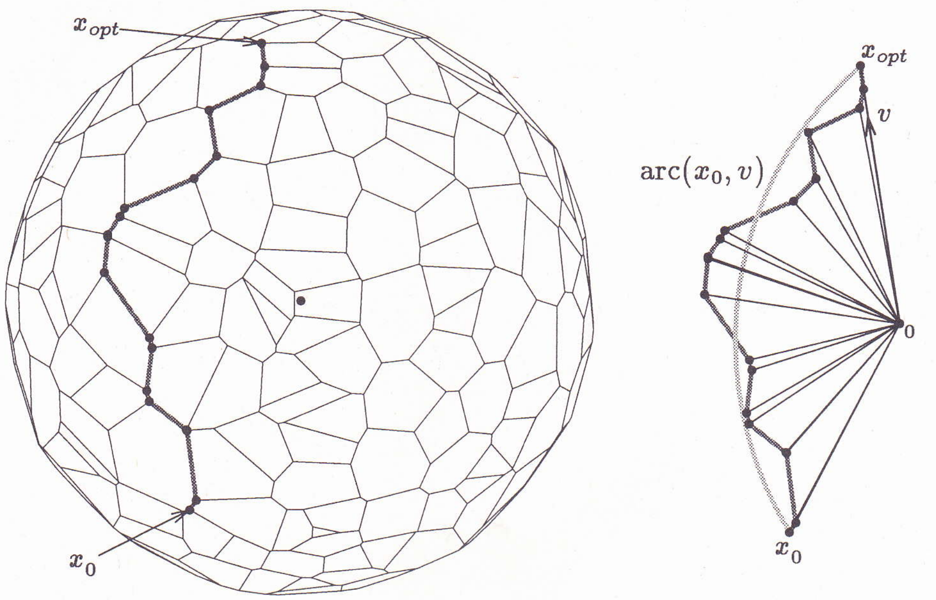


Figure 1: A Simplex-Path and its measurement via angles

The left figure shows the foreground of a three-dimensional polyhedron X and a Simplex-Path with the starting vertex x_0 and the final and optimal vertex $x_s = x_{opt}$. This Simplex-Path is extracted into the right figure to illustrate our idea of measuring the edge-movements and the measuring of the total distance from x_0 to v via angles.

As it is possible to observe our walk and the piecewise linear moves from the origin, we interpret measuring such a move as determining the angle between the directions induced by the starting point and the end point of the move. This model is based on a projection on the sphere and on fictive movements over geodetic curves on the sphere. Therefore, we introduce the notation

$$\begin{aligned}
&\measuredangle(\overline{x_0 v}) := \mathrm{arc}(x_0, v)\,, \\
&\measuredangle(\overline{x_i x_{i+1}}) := \mathrm{arc}(x_i, x_{i+1}) \quad \text{for } i = 0, \ldots, s-1\,, \\
&\measuredangle(\overline{x_s v}) := \mathrm{arc}(x_s, v)\,,
\end{aligned}$$

i. e. we measure edge-movements using the angles of direction-pairs.

Because the triangle inequality holds with those angles as distance measure, we have

$$\measuredangle(\overline{x_0 v}) = \measuredangle\left(\sum_{i=0}^{s} \overline{x_i x_{i+1}}\right) \leq \sum_{i=0}^{s} \measuredangle(\overline{x_i x_{i+1}})\,. \tag{10}$$

Our strategy for obtaining a lower bound on s lies in an attempt to show that the edges of X (resp. their angles) are deterministically resp. stochastically not long enough to bridge the distance from x_0 to v (resp. $\measuredangle(x_0, v)$) in a small number of steps.

We shall try to employ the methodology developed in [4, 8, 10] for a stochastic confirmation of such a statement. But since most of these techniques deal with the role of separate/single vertices and their contributions to polyhedral functionals, we would be better off if we could attribute our movements to single vertices rather than to edges. So we suggest to make the following translation.

Consider an edge $\overline{x_i x_{i+1}}$ $(0 \leq i \leq s-1)$ connecting two vertices $x_{\Delta_i} = x_i$ and $x_{\Delta_{i+1}} = x_{i+1}$. As explained before in Remark 1, part 3, we have $\#(\Delta_i \cap \Delta_{i+1}) = n-1$, and x_{Δ_i} and $x_{\Delta_{i+1}}$ are incident to the same edge $[x_{\Delta_i}, x_{\Delta_{i+1}}]$ and lie on its affine hull. This line is the solution set for the system $a_j^T x = 1 \; \forall j \in \Delta_i \cap \Delta_{i+1}$ and does not contain the origin.

So we have a unique (normalized) perpendicular to this line and a unique foot p_i of this perpendicular on the line. Therefore we know

$$(x_{\Delta_{i+1}} - x_{\Delta_i})^T p_i = 0 \quad \text{and} \quad p_i \in \operatorname{aff}(x_{\Delta_i}, x_{\Delta_{i+1}}).$$

Now three locations of p_i are possible:

1. between the two vertices (on the edge), i. e. $p_i \in [x_{\Delta_i}, x_{\Delta_{i+1}}]$;
2. "left" of x_{Δ_i}, such that $x_{\Delta_i} \in [p_i, x_{\Delta_{i+1}}]$;
3. "right" of $x_{\Delta_{i+1}}$, such that $x_{\Delta_{i+1}} \in [x_{\Delta_i}, p_i]$.

It is possible to dissect the movement from x_i to x_{i+1} into two parts: $\overline{x_i x_{i+1}} = \overline{x_i p_i} + \overline{p_i x_{i+1}}$ and our sum of movements reads as follows (compare (9))

$$\begin{aligned} \overline{x_0 v} \;&\widehat{=}\; \sum_{i=0}^{s} \overline{x_i x_{i+1}} \;\widehat{=}\; \sum_{i=0}^{s-1} (\overline{x_i p_i} + \overline{p_i x_{i+1}}) + \overline{x_s v} \\ &\widehat{=}\; \overline{x_0 p_0} + \sum_{i=1}^{s-1} (\overline{p_{i-1} x_i} + \overline{x_i p_i}) + \overline{p_{s-1} x_s} + \overline{x_s v}\,. \end{aligned} \tag{11}$$

Hereby we have to accept a higher distance for the dissected movement in cases 2) and 3), i. e. $\measuredangle(\overline{x_i x_{i+1}}) \leq \measuredangle(\overline{x_i p_i}) + \measuredangle(\overline{p_i x_{i+1}})$. In case 1) the distance remains the same.

For future purposes we want to remark that overestimations of the traversed angles are acceptable for our ultimate goal, the derivation of a lower bound on

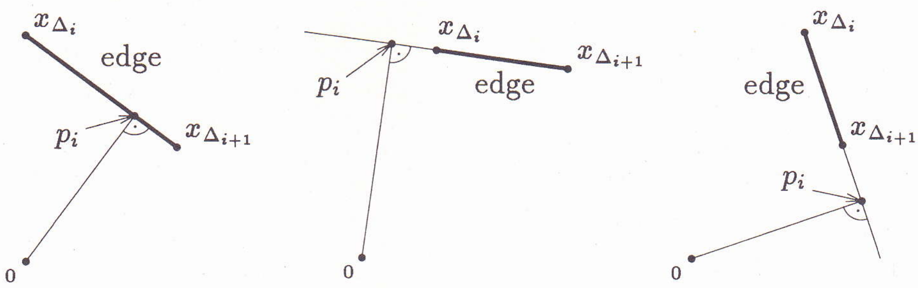


Figure 2: Location of the foot of the perpendicular

These figures illustrate the possible locations of the foot p_i of the perpendicular. In the left figure p_i is lying on the edge, i. e. $p_i \in [x_{\Delta_i}, x_{\Delta_{i+1}}]$. The figure in the middle illustrates the case 2, where p_i is lying "left" of x_{Δ_i}, such that $x_{\Delta_i} \in [p_i, x_{\Delta_{i+1}}]$. Case 3 is shown in the right figure, where p_i is lying "right" of $x_{\Delta_{i+1}}$, such that $x_{\Delta_{i+1}} \in [x_{\Delta_i}, p_i]$.

the number of steps. (This will become clear in the following.)
So, we have the following inequality for the lengths of our movements.

$$\begin{aligned} \measuredangle(\overline{x_0 v}) &\leq \sum_{i=0}^{s} \measuredangle(\overline{x_i x_{i+1}}) \leq \sum_{i=0}^{s-1} \Big(\measuredangle(\overline{x_i p_i}) + \measuredangle(\overline{p_i x_{i+1}}) \Big) + \measuredangle(\overline{x_s v}) \\ &= \measuredangle(\overline{x_0 p_0}) + \sum_{i=1}^{s-1} \Big(\measuredangle(\overline{p_{i-1} x_i}) + \measuredangle(\overline{x_i p_i}) \Big) + \measuredangle(\overline{p_{s-1} x_s}) + \measuredangle(\overline{x_s v}) \,. \end{aligned} \tag{12}$$

In this dissection we have a "halfedge" contributed by the vertex x_0, each time two "halfedges" contributed by $x_1, \ldots, x_{s-1}$, and x_s contributes one "halfedge" and the augmenting move $\overline{x_s v}$. As we want to assign the movements to single vertices in a consistent way, it will be advantageous to dissect the augmenting move in the case of unbounded problems in the following way. If x_s is not optimal, then the Simplex-Algorithm delivers an unbounded edge and we can dissect the movement from x_s to v into a movement from x_s to the foot of the perpendicular p_s of the unbounded edge and a second movement from p_s to v.

$$\begin{aligned} \overline{x_s v} \;\widehat{=}\; & \overline{x_s v} \cdot I(x_s \text{ is the final vertex and optimal}) \\ & + (\overline{x_s p_s} + \overline{p_s v}) \cdot I(x_s \text{ is the final vertex and exhibits} \\ & \qquad \text{the unboundedness of } v^T x \text{ on } X) \,. \end{aligned} \tag{13}$$

As it is clear that x_s denotes the last vertex on the Simplex-Path and that it is either the optimal vertex or exhibits the unboundedness of the problem, we abbreviate the indicator functions above as follows:

$$I(x_s \text{ is optimal}) := I(x_s \text{ is the final vertex and optimal})$$

$$\begin{aligned} &I(x_s \text{ exhibits unboundedness}) := \\ &\quad = I(x_s \text{ is the final vertex and exhibits the unboundedness of } v^T x \text{ on } X) . \end{aligned}$$

We measure these movements by the contributions of single vertices $\mathrm{CB}(x_i)$ to (12) and employ (13)

$$\begin{aligned} \mathrm{CB}(x_0) &:= \measuredangle(\overline{x_0 p_0}) \\ \mathrm{CB}(x_i) &:= \measuredangle(\overline{p_{i-1} x_i}) + \measuredangle(\overline{x_i p_i}) \quad \text{for } i = 1, \ldots, s-1 \\ \mathrm{CB}(x_s) &:= \measuredangle(\overline{p_{s-1} x_s}) + \measuredangle(\overline{x_s v})\, I(x_s \text{ is optimal}) \\ &\quad + \measuredangle(\overline{x_s p_s})\, I(x_s \text{ exhibits unboundedness}) . \end{aligned} \tag{14}$$

So, we know

$$\begin{aligned} \measuredangle(\overline{x_0 v}) &\le \measuredangle(\overline{x_0 p_0}) + \sum_{i=1}^{s-1} (\measuredangle(\overline{p_{i-1} x_i}) + \measuredangle(\overline{x_i p_i})) + \measuredangle(\overline{p_{s-1} x_s}) \\ &\quad + \measuredangle(\overline{x_s v})\, I(x_s \text{ is optimal}) \\ &\quad + \measuredangle(\overline{x_s p_s})\, I(x_s \text{ exhibits unboundedness}) \\ &\quad + \measuredangle(\overline{p_s v})\, I(x_s \text{ exhibits unboundedness}) \\ &= \sum_{i=0}^{s} \mathrm{CB}(x_i) + \measuredangle(\overline{p_s v})\, I(x_s \text{ exhibits unboundedness}) . \end{aligned}$$

Now we could partition the sum on the right side into two categories, namely large contributions and small contributions. Therefore we introduce a threshold $\sigma > 0$ and call a vertex contribution $\mathrm{CB}(x_i)$

$$\begin{aligned} &\sigma - \text{large, if } \mathrm{CB}(x_i) > \sigma , \\ &\sigma - \text{small, if } \mathrm{CB}(x_i) \le \sigma . \end{aligned} \tag{15}$$

And, obviously

$$\begin{aligned} \measuredangle(\overline{x_0 v}) &\le \sum_{i=0}^{s} \mathrm{CB}(x_i) + \measuredangle(\overline{p_s v})\, I(x_s \text{ exhibits unboundedness}) \\ &= \sum_{i=0}^{s} \mathrm{CB}(x_i)\, I(\mathrm{CB}(x_i) > \sigma) + \sum_{i=0}^{s} \mathrm{CB}(x_i)\, I(\mathrm{CB}(x_i) \le \sigma) \\ &\quad + \measuredangle(\overline{p_s v})\, I(x_s \text{ exhibits unboundedness}) . \end{aligned}$$

Next, we move the sum of "large contributions" and $\measuredangle(\overline{p_s v})$ to the left hand side

$$\begin{aligned}
&\measuredangle(\overline{x_0 v}) - \sum_{i=0}^{s} \mathrm{CB}(x_i)\, I(\mathrm{CB}(x_i) > \sigma) - \measuredangle(\overline{p_s v})\, I(x_s \text{ exhibits unboundedness}) \\
&\quad \leq \sum_{i=0}^{s} \mathrm{CB}(x_i)\, I(\mathrm{CB}(x_i) \leq \sigma) \,.
\end{aligned}$$

Imagine that we were able to calculate the left side (no matter how and why, but take this as given). Then we could conclude as follows:

$$\begin{aligned}
&\measuredangle(\overline{x_0 v}) - \sum_{i=0}^{s} \mathrm{CB}(x_i)\, I(\mathrm{CB}(x_i) > \sigma) - \measuredangle(\overline{p_s v})\, I(x_s \text{ exhibits unboundedness}) \\
\leq\ & \sum_{i=0}^{s} \mathrm{CB}(x_i)\, I(\mathrm{CB}(x_i) \leq \sigma) \\
\leq\ & \sum_{i=0}^{s} \sigma\, I(\mathrm{CB}(x_i) \leq \sigma) \\
=\ & \sigma \cdot \#\{x_i \mid \mathrm{CB}(x_i) \leq \sigma\} = \sigma \cdot N_\sigma \,.
\end{aligned}$$

We introduce this number

$$N_\sigma := \#\{x_i \mid \mathrm{CB}(x_i) \leq \sigma\} \tag{16}$$

and define the left side as σ-pathgap:

$$\begin{aligned}
\sigma\text{-pathgap} \quad := \quad & \measuredangle(\overline{x_0 v}) - \sum_{i=0}^{s} \mathrm{CB}(x_i)\, I(\mathrm{CB}(x_i) > \sigma) \\
& - \measuredangle(\overline{p_s v})\, I(x_s \text{ exhibits unboundedness}) \,.
\end{aligned} \tag{17}$$

Note that this σ-pathgap may be positive or negative. If it is negative, it is useless for our purpose, but if it is positive, it yields

$$\frac{\sigma\text{-pathgap}}{\sigma} \leq N_\sigma \leq s + 1 \,. \tag{18}$$

The right inequality is a simple consequence of the fact that N_σ counts a certain subset of the vertices on the path. So our (so far unexplained) knowledge of the σ-pathgap could lead to a lower bound for s.

Note, that for increasing σ the sum $\sum_{i=0}^{s} \mathrm{CB}(x_i)\, I(\mathrm{CB}(x_i) > \sigma)$ will decrease, so the numerator in (18) will increase. But simultaneously the denominator gets larger. In order to get a good lower bound we should choose σ small, but in such a way that the numerator still is considerably positive.

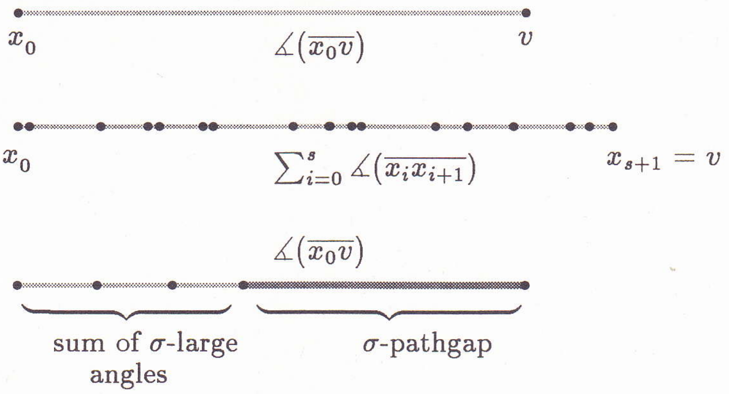


Figure 3: σ-pathgap

For the Simplex-Path in Figure 1 the measured distance of $\measuredangle(\overline{x_0 v})$ is represented by a line segment with length according to $\measuredangle(\overline{x_0 v})$. The same is done for $\sum_{i=0}^{s} \measuredangle(\overline{x_i x_{i+1}})$. Obviously, we have: $\measuredangle(\overline{x_0 v}) \leq \sum_{i=0}^{s} \measuredangle(\overline{x_i x_{i+1}})$.
The third line segment illustrates that the distance from x_0 to v cannot be bridged with this sum of great angles and that there remains a positive σ-pathgap, which has to be bridged with σ-small angles.

Here the practicability of that approach should be questioned. Calculation of the σ-pathgap will only be possible in deterministic special cases. The main reason lies in the fact that the characterization of a path for a general variant is by far too complicated.
This had been just the reason why so far all attempts to derive average upper bounds for general variants had failed. But before we give up, we should think about a relaxation. If we could sum up over all vertices, which are able to make a σ-large contribution, instead of using only those on the path, we could overestimate the sum associated to σ-large contributions and underestimate the σ-pathgap by a "σ-polyhedrongap". If still this σ-polyhedrongap is positive then the according approach could lead to a (worse, but still correct) lower bound on s.

2.3 A Polyhedron in the Dual Space

Before we make an attempt to derive an evaluable lower bound for the σ-pathgap, we should transfer our considerations to the dual space, where the input vectors a_i are generated. There we find a certain counterpart to X, namely the polyhedron resp. polytope

$$Y := \mathrm{conv}(a_1, \ldots, a_m)\,.$$

The reason for moving our concentration from X to Y lies in the very direct impact of the input data on the shape of Y. The input data do form the polyhedron X as well, but this is made in a very indirect and hardly computable way. Another reason for the concentration on Y is the fact that we can obtain all relevant information on X also by studying Y.

Y is – due to nondegeneracy – a simplicial polytope, i. e. every facet is a simplex generated by n points out of $\{a_1, \ldots, a_m\}$. We call such a simplex $\mathrm{conv}(a_{\Delta^1}, \ldots, a_{\Delta^n})$ a basic simplex. There are $\binom{m}{n}$ such candidates and some of them will turn out to be facets of Y.
So we observe as augmentation of the one to one correspondence in Remark 1, part 1,

$$\text{basic solution } x_\Delta \longleftrightarrow \Delta \longleftrightarrow \mathrm{conv}(a_{\Delta^1}, \ldots, a_{\Delta^n}) = \mathrm{conv}(\Delta) \text{ basic simplex.}$$

Each such basic simplex is part of a hyperplane $H(a_{\Delta^1}, \ldots, a_{\Delta^n}) =: H(\Delta)$, which is the affine hull of $\mathrm{conv}(\Delta)$. This hyperplane bounds two open halfspaces, namely $H^+(a_{\Delta^1}, \ldots, a_{\Delta^n}) =: H^+(\Delta)$ and $H^-(a_{\Delta^1}, \ldots, a_{\Delta^n}) =: H^-(\Delta)$.
Due to nondegeneracy, $0 \notin H(\Delta)$. So we define the distinction between $H^+(\Delta)$ and $H^-(\Delta)$ according to the location of 0: $H^-(\Delta)$ is the halfspace, that contains 0, and $H^+(\Delta)$ is the opposite halfspace with $0 \notin H^+(\Delta)$.
By $h(\Delta)$ resp. $h(a_{\Delta^1}, \ldots, a_{\Delta^n})$ we denote the distance of $H(\Delta)$ to the origin. Note, that x_Δ – as defined in the section before – is a normal vector on $H(\Delta)$.

Now let us look at Y and its facet-structure. A basic simplex $\mathrm{conv}(\Delta)$ will be a facet of Y if and only if its bearing hyperplane $H(\Delta)$ is a supporting hyperplane of Y. That means that

1. either $Y \subset H(\Delta) \cup H^-(\Delta)$ (all a_j $(j \notin \Delta)$ are contained in $H^-(\Delta)$)
2. or $Y \subset H(\Delta) \cup H^+(\Delta)$ (all a_j $(j \notin \Delta)$ are contained in $H^+(\Delta)$).

In the first case, we call $\mathrm{conv}(\Delta)$ a facet of 1st kind; and in the second case, $\mathrm{conv}(\Delta)$ is a facet of 2nd kind.

It is easy to see, that exactly in the case that $\mathrm{conv}(\Delta)$ is a facet of 1st kind, then x_Δ (the corresponding basic solution) is a vertex of X.
Investigating the structure of a simplicial polytope further, we see that every facet of Y, e. g. $\mathrm{conv}(a_{\Delta^1}, \ldots, a_{\Delta^n})$ has n side simplices of dimension $n-2$ of the form $\mathrm{conv}(a_{\Delta^1}, \ldots, a_{\Delta^{i-1}}, a_{\Delta^{i+1}}, \ldots, a_{\Delta^n})$. To each such side-simplex it is guaranteed that exactly one a_j $(j \notin \Delta)$ exists, that augments our side-simplex in a

way that another facet is formed, i. e. $\operatorname{conv}(a_{\Delta^1},\ldots,a_{\Delta^{i-1}},a_j,a_{\Delta^{i+1}},\ldots,a_{\Delta^n})$ is another facet. We call a pair of such facets (with $n-1$ joint generators) "fully adjacent".

Now there may be three types of "full adjacency" between two facets, which meet in a common side-simplex $\operatorname{conv}(a_{\Delta^1},\ldots,a_{\Delta^{i-1}},a_{\Delta^{i+1}},\ldots,a_{\Delta^n})$ and are augmentations of this side-simplex with a_{j_1} resp. a_{j_2}:

1. both are facets of 1st kind,
2. both are facets of 2nd kind
3. a facet of 1st kind meets a facet of 2nd kind.

For describing the properties of X, the cases 1) and 3) are important, whereas case 2) is not interesting.

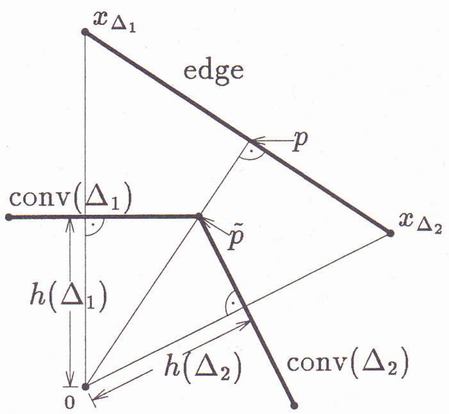
Case 1: Consider the index sets $\Delta_1 = (\Delta^1,\ldots,\Delta^{i-1},j_1,\Delta^{i+1},\ldots,\Delta^n)$ and $\Delta_2 = (\Delta^1,\ldots,\Delta^{i-1},j_2,\Delta^{i+1},\ldots,\Delta^n)$. As $\operatorname{conv}(\Delta_1)$ and $\operatorname{conv}(\Delta_2)$ are facets of 1st kind, x_{Δ_1} and x_{Δ_2} are both vertices of X and $n-1$ common constraints are active.
The solution set of $n-1$ equations $\{x \in \mathbb{R}^n \mid a_j^T x = 1 \; \forall j \in \Delta_1 \cap \Delta_2\}$ is a line, which contains x_{Δ_1} and x_{Δ_2} and can be described as $x_{\Delta_1} + \mathbb{R}d$, where d denotes the direction of the line. The direction d is orthogonal to $\operatorname{lin}(a_{\Delta^1},\ldots,a_{\Delta^{i-1}},a_{\Delta^{i+1}},\ldots,a_{\Delta^n})$ (the linear hull containing the side-simplex $\operatorname{conv}(a_{\Delta^1},\ldots,a_{\Delta^{i-1}},a_{\Delta^{i+1}},\ldots,a_{\Delta^n})$).
So the line intersects the linear subspace in a point p, which is just the foot of the perpendicular on this line. We orient d in such a way that $a_{\Delta_1^i}^T d < 0$. As we have in x_{Δ_1}: $a_{\Delta_1^i}^T x_{\Delta_1} = 1$, $a_{\Delta_2^i}^T x_{\Delta_1} < 1$; and in x_{Δ_2}: $a_{\Delta_2^i}^T x_{\Delta_2} = 1$, $a_{\Delta_1^i}^T x_{\Delta_2} < 1$, it becomes clear that $[x_{\Delta_1}, x_{\Delta_2}]$ is a bounded edge of X with direction d.

Case 3: We set Δ_1 and Δ_2 as before. As two types of facets meet, let $\operatorname{conv}(\Delta_1)$ be the facet of 1st kind and $\operatorname{conv}(\Delta_2)$ be the facet of 2nd kind. Then x_{Δ_1} is a vertex of X, but x_{Δ_2} is none.
Again, we find a line consisting of all solutions of the $n-1$ equations $a_j^T x = 1 \;\; \forall j \in \Delta_1 \cap \Delta_2$, and this line contains x_{Δ_1} and is orthogonal to $\operatorname{lin}(a_{\Delta^1},\ldots,a_{\Delta^{i-1}},a_{\Delta^{i+1}},\ldots,a_{\Delta^n})$. The intersection point of the line and the linear subspace is p (which is again the foot of the perpendicular on our line). Let the direction of the line be d, oriented in the same way as in case 1. Then it becomes clear that $x_{\Delta_1} + \mathbb{R}^+ d$ is an unbounded edge of X, because $\operatorname{lin}(a_{\Delta^1},\ldots,a_{\Delta^{i-1}},a_{\Delta^{i+1}},\ldots,a_{\Delta^n})$ is a supporting hyperplane of Y with outer normal vector d and hence for all a_j we have $a_j^T d \leq 0$.

Our Simplex-Path discussed in the previous section can now be interpreted in this dual space as a walk over facets (of 1st kind) of Y, where successive facets are fully adjacent. The walk ends in a facet of 1st kind that intersects $\mathbb{R}^+ v$ or when our move leads us to a side-simplex, which also belongs to a facet of 2nd kind. This is the translation of the potential case that in x_s an unbounded edge improving v proves the nonexistence of an optimal point.
As derived before, we should measure all the angles which result from a move from a vertex x_Δ (the normal on $H(\Delta)$) to a point p, which is the foot of the perpendicular on an outgoing edge.

For handling the final movement in an appropiate way we consider the following fact, formulated with indicator functions of events:

$$\begin{aligned} I(x_s \text{ exhibits unboundedness}) &\leq \\ \leq\ & I(X \text{ has unbounded edges}) = I(Y \text{ has facets of 2nd kind}) \\ \leq\ & \sum_{\Delta} I(\mathrm{conv}(\Delta) \text{ is facet of 2nd kind}) . \end{aligned} \tag{19}$$



In case 1) we have two adjacent vertices x_{Δ_1} and x_{Δ_2} and the foot of the perpendicular p is lying on the edge. Simultaneously, you can see the basic simplices and facets of the polyhedron Y, namely $\mathrm{conv}(\Delta_1)$ and $\mathrm{conv}(\Delta_2)$. $\tilde{p}$ is the foot of the perpendicular on $\mathrm{aff}(a_j \mid j \in \Delta_1 \cap \Delta_2)$ and a multiple of the vector p.

Figure 4: Illustration of case 1)

Let us repeat and add some facts about the point/direction p.
p has been introduced as foot of the perpendicular on the line containing an edge incident to x_Δ. This line is orthogonal to $\mathrm{lin}(a_{\Delta^1}, \ldots, a_{\Delta^{i-1}}, a_{\Delta^{i+1}}, \ldots, a_{\Delta^n})$ and p satisfies $a_{\Delta^1}^T p = \ldots = a_{\Delta^{i-1}}^T p = a_{\Delta^{i+1}}^T p = \ldots = a_{\Delta^n}^T p = 1$. Hence p is a normal vector on $\mathrm{aff}(a_{\Delta^1}, \ldots, a_{\Delta^{i-1}}, a_{\Delta^{i+1}}, \ldots, a_{\Delta^n})$ and there is a point $\tilde{p} = \lambda p$ ($\lambda \in \mathbb{R}^+$), such that $\tilde{p}$ is the foot of the perpendicular on $\mathrm{aff}(a_{\Delta^1}, \ldots, a_{\Delta^{i-1}}, a_{\Delta^{i+1}}, \ldots, a_{\Delta^n})$. For dealing with angles $\measuredangle(x_\Delta, p)$ it does not matter when we replace p by $\frac{1}{||p||} p$ or by $\tilde{p}$.

To simplify the study of the geometric configuration we perform a transformation of coordinates in the following manner. Given (w. l. o. g.) the index sets $\Delta = \{1, \dots, n\}$ and $\Delta_\star = \{1, \dots, n-1, n+1\}$ and the vectors $a_1, \dots, a_{n+1}$ we rotate $\mathbb{R}^n$ in such a way that $a_1, \dots, a_m$ are mapped on points $b_1, \dots, b_m$, i. e.

$$b_i = \begin{pmatrix} b_i^1 \\ \vdots \\ b_i^{n-1} \\ b_i^n \end{pmatrix}$$

such that $b_1^n, \dots, b_n^n = h$. That means that the last coordinate of the points $b_1, \dots, b_n$ coincides at a joint value $h > 0$.
Of course h is the distance of $H(a_1, \dots, a_n)$, resp. $H(b_1, \dots, b_n)$ to the origin (remember that we have used a rotation), i. e. $h = h(\Delta)$.
For the new vectors we define

$$\bar{b}_i = \begin{pmatrix} b_i^1 \\ \vdots \\ b_i^{n-1} \end{pmatrix} \in \mathbb{R}^{n-1}, \; i = 1, \dots, m\,.$$

So far our rotation helps to relocate a facet or a complete basic simplex on a common level in the n-th coordinate. But the study of the points p makes it also necessary to handle side-simplices of such facets, resp. basic simplices. So we apply another rotation to simplify the description of the side-simplex $\mathrm{conv}(b_1, \dots, b_{n-1})$. We rotate the first $n-1$ coordinates again in such a way that the vectors $\bar{b}_i$ are mapped on vectors c_i, $i = 1, \dots, m$, and that

$$c_1^{n-1} = \dots = c_{n-1}^{n-1} = \theta > 0$$

(remember that still $c_1^n = \dots = c_{n-1}^n = c_n^n = h$).
Due to our general condition that the support is contained in Ω_n, it follows that $a_i \in \Omega_n$ and $b_i, c_i \in \Omega_n$ for all $i = 1, \dots, m$ as well. Hence $0 < h < 1$ and $0 < \theta < \sqrt{1 - h^2}$.
Now it is possible to calculate the position of the two vertices x_Δ and $x_{\Delta_\star}$ via the systems of equations

$$\begin{pmatrix} c_1^1 & \cdots & c_1^{n-2} & \theta & h \\ \vdots & & \vdots & \vdots & \vdots \\ c_{n-1}^1 & \cdots & c_{n-1}^{n-2} & \theta & h \\ c_n^1 & \cdots & c_n^{n-2} & c_n^{n-1} & h \end{pmatrix} x_\Delta = \begin{pmatrix} 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix} \qquad \text{and}$$

$$\begin{pmatrix} c_1^1 & \cdots & c_1^{n-2} & \theta & h \\ \vdots & & \vdots & \vdots & \vdots \\ c_{n-1}^1 & \cdots & c_{n-1}^{n-2} & \theta & h \\ c_{n+1}^1 & \cdots & c_{n+1}^{n-2} & c_{n+1}^{n-1} & c_{n+1}^n \end{pmatrix} x_{\Delta_\star} = \begin{pmatrix} 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix}.$$

We obtain

$$x_\Delta = \frac{1}{h} e_n \quad \text{and} \quad x_{\Delta_\star} = \frac{c_{n+1}^n - h}{c_{n+1}^n \theta - c_{n+1}^{n-1} h} e_{n-1} + \frac{\theta - c_{n+1}^{n-1}}{c_{n+1}^n \theta - c_{n+1}^{n-1} h} e_n .$$

For x_Δ becoming a vertex, it must be ascertained that $c_j \in H^-(c_1, \ldots, c_n)$ $\forall j = n+1, \ldots, m$, resp. that

$$c_j^n \leq h \quad \forall j = n+1, \ldots, m . \tag{20}$$

$x_{\Delta_\star}$ will be an additional vertex, if and only if $c_j \in H^-(c_1, \ldots, c_{n-1}, c_{n+1})$ $\forall j = n, n+2, \ldots, m$, or resp.

$$\frac{c_j^{n-1}(c_{n+1}^n - h) + c_j^n(\theta - c_{n+1}^{n-1})}{c_{n+1}^n \theta - c_{n+1}^{n-1} h} \leq 1 \quad \forall j = n, n+2, \ldots, m . \tag{21}$$

So we learn that if we want x_Δ and $x_{\Delta_\star}$ to be vertices, then the potential location of $c_{n+2}, \ldots, c_m$ and of c_n, c_{n+1} is considerably restricted.

Now let us consider the description of the case where $\operatorname{conv}(\Delta)$ is a facet of 2nd kind. This is true if and only if

$$c_{n+1}, \ldots, c_m \in H^+(\Delta) \quad \text{resp.} \quad c_i^n > h \quad \forall i \notin \Delta .$$

Following this concept, we can also describe the configuration, where a facet of 1st kind $\operatorname{conv}(\Delta)$ is fully adjacent to a facet $\operatorname{conv}(\Delta_\star)$ of 2nd kind. This happens if and only if

$$c_{n+1}, \ldots, c_m \in H^-(\Delta) \quad \text{and} \quad c_n, c_{n+2}, \ldots, c_m \in H^+(\Delta_\star)$$

or – explicitly – if and only if

$$\begin{gathered} c_i^n \leq h \quad \forall i \notin \Delta \quad \text{and} \\ \frac{c_j^{n-1}(c_{n+1}^n - h) + c_j^n(\theta - c_{n+1}^{n-1})}{c_{n+1}^n \theta - c_{n+1}^{n-1} h} \geq 1 \quad \forall j = n, n+2, \ldots, m . \end{gathered} \tag{22}$$

Finally, let us identify the point (or direction) p after our transformation of coordinates. We know that p is the foot of perpendicular on the edge through

x_Δ (from the origin), and a positive multiple of p, namely $\tilde{p} = \lambda p$, is the foot of the perpendicular from the origin on $\mathrm{aff}(c_1, \ldots, c_{n-1})$.
Since $\mathrm{aff}(c_1, \ldots, c_{n-1}) = \{x \mid x^{n-1} = \theta,\ x^n = h,\ x \in \mathbb{R}^n\}$, we have

$$\tilde{p} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \theta \\ h \end{pmatrix} \quad \text{and} \quad p = \frac{1}{\sqrt{\theta^2 + h^2}} \cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \theta \\ h \end{pmatrix} .$$

For the vertex x_Δ we get

$$\measuredangle(\overline{x_\Delta p}) = \measuredangle(\overline{x_\Delta \tilde{p}}) = \arcsin\left(\frac{\theta}{\sqrt{\theta^2 + h^2}}\right) = \arccos\left(\frac{h}{\sqrt{\theta^2 + h^2}}\right) .$$

If we know $h = h(\Delta)$, then we remember that $\theta \le \sqrt{1 - h^2}$ due to $c_i \in \Omega_n$:

$$\Rightarrow \quad \arcsin\left(\frac{\theta}{\sqrt{\theta^2 + h^2}}\right) \le \arcsin(\sqrt{1 - h^2}\,) \text{ and}$$
$$\arccos\left(\frac{h}{\sqrt{\theta^2 + h^2}}\right) \le \arccos(h) .$$

The result is

$$\measuredangle(\overline{x_\Delta p}) \le \arcsin(\sqrt{1 - h^2}\,) = \arccos(h) = \arccos(h(\Delta)) \tag{23}$$

for all edges incident to x_Δ, where p is the corresponding foot of the perpendicular on that edge.
Also we know something about $\measuredangle(\overline{x_\Delta u})$ for any $u \in \mathrm{cone}(c_1, \ldots, c_n)$. Without loss of generality let $u^n = h$ (as for $c_1, \ldots, c_n$). That means that $u \in \mathrm{conv}(c_1, \ldots, c_n)$.
Since $\mathrm{conv}(c_1, \ldots, c_n) \subset \{x \mid x^n = h, \|\bar{x}\| \le \sqrt{1 - h^2}\,, x \in \mathbb{R}^n\}$, it is clear that

$$\measuredangle(\overline{x_\Delta u}) \le \arcsin(\sqrt{1 - h^2}\,) = \arccos(h)$$

for all $u \in \mathrm{conv}(c_1, \ldots, c_n)$ as well.
This estimation can be applied in the case where $\mathrm{conv}(\Delta)$ is the optimal facet, resp. x_Δ is the optimal vertex. The optimality is indicated by the fact that $v \in \mathrm{cone}(c_1, \ldots, c_n)$.
And now we have learned that under the condition that x_Δ is the optimal vertex,

$$\measuredangle(\overline{x_\Delta v}) \le \arcsin(\sqrt{1 - h^2}\,) = \arccos(h) = \arccos(h(\Delta)) . \tag{24}$$

In our algorithm this move from $x_\Delta = x_{opt}$ to v will be the final move. So, in case of existence of an optimal vertex, we actually have an overestimation of the angle which depends only on features of the vertex x_Δ.

But if the last vertex x_s on the Simplex-Path is not optimal (resp. when $v \notin Y$), then we could argue that the angle of the move to p_s (on the unbounded edge) $\measuredangle(\overline{x_s p_s})$ is boundable as above by $\arcsin(\sqrt{1-h^2}) = \arccos(h) = \arccos(h(\Delta))$. And after that we have to reach v from p_s and we do not know a better bound than π for the angle of that move. Employing this very conservative estimation for the last move, we are now able to bound the "vertex-contributions" independently of the edges, which have been selected.

2.4 An upper bound on the pathgap (in dual notation)

Remember that we had derived a so-called σ-pathgap (17) in the form

$$\begin{aligned}\sigma\text{-pathgap} \;\; = \;\; & \measuredangle(\overline{x_0 v}) - \sum_{i=0}^{s} \mathrm{CB}(x_i)\, I(\mathrm{CB}(x_i) > \sigma) \\ & - \measuredangle(\overline{p_s v})\, I(x_s \text{ exhibits unboundedness})\end{aligned}$$

with the vertex contributions $\mathrm{CB}(\,.\,)$ as defined in (14). And we could use the σ-pathgap to bound s from below as in (18):

$$s + 1 \geq N_\sigma \geq \frac{\sigma\text{-pathgap}}{\sigma}\,.$$

The σ-pathgap is a difference, which mainly depends on the value of the sum

$$\sum_{i=0}^{s} \mathrm{CB}(x_i)\, I(\mathrm{CB}(x_i) > \sigma) + \measuredangle(\overline{p_s v})\, I(x_s \text{ exhibits unboundedness})\,. \tag{25}$$

If we overestimate this sum and still keep a positive σ-pathgap, then we obtain a lower bound as desired. So let us try to overestimate (25) in a moderate and simplifying way.

Take an arbitrary internal vertex x_i, $1 \leq i \leq s-1$, and let $h_i = h(\Delta_i)$ denote the height of the corresponding facet of Y. Then – using (23) – we have

$$\begin{aligned}\mathrm{CB}(x_i) = \measuredangle(\overline{p_{i-1} x_i}) + \measuredangle(\overline{x_i p_i}) \leq 2\arccos(h_i) \\ \text{and } \; I(\mathrm{CB}(x_i) > \sigma) \leq I\Big(\arccos(h_i) > \frac{\sigma}{2}\Big)\end{aligned} \tag{26}$$

because the left indicator can only be 1 if the right one is.
For the initial vertex we have a parallel, trivial consideration for the same reasons

$$\begin{aligned}\mathrm{CB}(x_0)\, I(\mathrm{CB}(x_0) > \sigma) \;\; \leq \;\; & \arccos(h_0) \cdot I(\arccos(h_0) > \sigma) \\ \leq \;\; & 2\arccos(h_0) \cdot I\Big(\arccos(h_0) > \frac{\sigma}{2}\Big)\,.\end{aligned} \tag{27}$$

A similar treatment of the last vertex's contribution results from (23) and (24):

$$\begin{aligned}
\mathrm{CB}(x_s) &= \measuredangle(\overline{p_{s-1}x_s}) + \measuredangle(\overline{x_s v})\, I(x_s \text{ is optimal}) \\
&\quad + \measuredangle(\overline{x_s p_s})\, I(x_s \text{ exhibits unboundedness}) \\
&\leq \arccos(h_s) + \arccos(h_s)\, I(x_s \text{ is optimal}) \\
&\quad + \arccos(h_s)\, I(x_s \text{ exhibits unboundedness}) \\
&= 2\arccos(h_s)
\end{aligned}$$

and therefore

$$\mathrm{CB}(x_s)\, I(\mathrm{CB}(x_s) > \sigma) \leq 2\arccos(h_s)\, I\Big(\arccos(h_s) > \frac{\sigma}{2}\Big)\ . \tag{28}$$

As mentioned before, the only bound for $\measuredangle(\overline{p_s v})$ we know is π:

$$\begin{aligned}
&\measuredangle(\overline{p_s v})\, I(x_s \text{ exhibits unboundedness}) \\
&\qquad \leq \pi\, I(x_s \text{ exhibits unboundedness})\ .
\end{aligned} \tag{29}$$

Now we have, as a result of (26)–(29) and (19),

$$\begin{aligned}
\sigma\text{-pathgap} &\geq \measuredangle(\overline{x_0 v}) - \sum_{i=0}^{s} \mathrm{CB}(x_i)\, I(\mathrm{CB}(x_i) > \sigma) \\
&\quad - \measuredangle(\overline{p_s v})\, I(x_s \text{ exhibits unboundedness}) \\
&\geq \measuredangle(\overline{x_0 v}) - \sum_{i=0}^{s} 2\arccos(h_i)\, I\Big(\arccos(h_i) > \frac{\sigma}{2}\Big) \\
&\quad - \pi \cdot I(X \text{ has unbounded edges})\ .
\end{aligned} \tag{30}$$

In the next section, we want to employ this underestimation for an average-case analysis. There it will be absolutely disastrous to have the obligation of a selection of the path vertices $x_0, \ldots, x_s$. But here we can underestimate the pathgap resp. overestimate the sum in the following way.

$$\begin{aligned}
&\sum_{i=0}^{s} 2\arccos(h_i)\, I\Big(\arccos(h_i) > \frac{\sigma}{2}\Big) \leq \\
&\quad \leq \sum_{\Delta} 2\arccos(h(\Delta))\, I(\mathrm{conv}(\Delta) \text{ is facet of 1st kind})\ I\Big(\arccos(h(\Delta)) > \frac{\sigma}{2}\Big)\ ,
\end{aligned}$$

where we have summed up over all index sets Δ of basic solutions. And after our discussion of unboundedness we know that

$$\pi \cdot I(X \text{ has unbounded edges}) \leq \pi \sum_{\Delta} I(\mathrm{conv}(\Delta) \text{ is facet of 2nd kind})\ .$$

So our new estimation of the σ-pathgap is:

$$
\begin{aligned}
&\sigma\text{-pathgap}\\
&\quad \geq \measuredangle(x_0, v)\\
&\qquad - \sum_{\Delta} 2\arccos(h(\Delta))\, I(\mathrm{conv}(\Delta) \text{ is facet of 1st kind})\ I\Big(\arccos(h(\Delta)) > \frac{\sigma}{2}\Big)\\
&\qquad - \pi \cdot \sum_{\Delta} I(\mathrm{conv}(\Delta) \text{ is facet of 2nd kind})\ . \qquad (31)
\end{aligned}
$$

3 A lower bound on the average number of steps

Now we start the averaging process. We would like to sum up or average over all possible problems and to calculate the average value of the available lower bound for s.

If we apply our lower bound on the σ-pathgap (31), then we observe that this often will be negative. Here we are in a conflict, as we know that s never will be negative. So we could work with the maximum of 0 and the term in (31), but this would overcomplicate the calculations. So we conclude

$$
\begin{aligned}
E_{m,n}[s+1] \geq E_{m,n}\left[\frac{\sigma\text{-pathgap}}{\sigma}\right] &= \frac{E_{m,n}[\sigma\text{-pathgap}]}{\sigma} \\
\geq\ & \frac{1}{\sigma}\, E_{m,n}[\measuredangle(x_0, v)] \\
& - \frac{1}{\sigma} E_{m,n}\Big[\sum_{\Delta} 2\arccos(h(\Delta))\, I(\operatorname{conv}(\Delta) \text{ is facet of 1st kind}) \\
& \qquad\qquad \cdot I\Big(\arccos(h(\Delta)) > \frac{\sigma}{2}\Big)\Big] \qquad (32) \\
& - \frac{1}{\sigma}\, E_{m,n}\Big[\pi \cdot \sum_{\Delta} I(\operatorname{conv}(\Delta) \text{ is facet of 2nd kind})\Big].
\end{aligned}
$$

3.1 Calculation of expectation values

Let us first think about the expected value $E_{m,n}[\measuredangle(x_0, v)]$: For each subset of problems, where $a_1, \ldots, a_m$ are fixed, Phase I works in the same way. So it produces the same vertex x_0 in all these cases, which means that the direction of x_0 is fixed. Now (independently of x_0) v is distributed symmetrically under rotations. So it is trivial, that the average angle is $\frac{\pi}{2}$. This holds for all choices of $a_1, \ldots, a_m$ and therefore the total average value is

$$
E_{m,n}[\measuredangle(x_0, v)] = \frac{\pi}{2}. \qquad (33)
$$

The second expectation value in (32) can be written as

$$
\begin{aligned}
& E_{m,n}\Big[\sum_{\Delta} 2\arccos(h(\Delta))\, I(\operatorname{conv}(\Delta) \text{ is facet of 1st kind})\ I\Big(\arccos(h(\Delta)) > \frac{\sigma}{2}\Big)\Big] \\
& \quad = 2\binom{m}{n} E_{m,n}\Big[\arccos(h(\Delta_\star))\ I(\operatorname{conv}(\Delta_\star) \text{ is facet of 1st kind}) \\
& \qquad\qquad \cdot I\Big(\arccos(h(\Delta_\star)) > \frac{\sigma}{2}\Big)\Big] \qquad (34)
\end{aligned}
$$

This equality is a result of the fact, that the input data are distributed identically, and therefore we can choose a typical candidate and a typically basic index set $\Delta_\star := \{1,\dots,n\}$ and count the number of possibilities, which is $\binom{m}{n}$. Using integrals we have

$$(34) \quad = \quad 2\binom{m}{n}\int\limits_{\mathbb{R}^n}^{(m)} \arccos(h(\Delta_\star))\, I(\mathrm{conv}(\Delta_\star) \text{ is facet of 1st kind})$$

$$\cdot\, I\Big(\arccos(h(\Delta_\star)) > \frac{\sigma}{2}\Big)\; d\hat{F}(a_1)\dots d\hat{F}(a_m)\,. \tag{35}$$

Here, $\int\limits_{\mathbb{R}^n}^{(m)}$ abbreviates the notation $\underbrace{\int\limits_{\mathbb{R}^n}\cdots\int\limits_{\mathbb{R}^n}}_{m}$.

For the last expectation value in (32) we can estimate $\pi \leq 2\arccos(h)$ for any $h < 0$ and we obtain

$$\begin{aligned}
&E_{m,n}\Big[\pi\cdot\sum_\Delta I(\mathrm{conv}(\Delta) \text{ is facet of 2nd kind})\Big]\\
&\leq E_{m,n}\Big[\sum_\Delta 2\arccos(-h(\Delta))\, I(\mathrm{conv}(\Delta) \text{ is facet of 2nd kind})\Big]\\
&= 2\binom{m}{n} E_{m,n}\Big[\arccos(-h(\Delta_\star))\, I(\mathrm{conv}(\Delta_\star) \text{ is facet of 2nd kind})\Big]\\
&= 2\binom{m}{n}\int\limits_{\mathbb{R}^n}^{(m)} \arccos(-h(\Delta_\star))\, I(\mathrm{conv}(\Delta_\star) \text{ is facet of 2nd kind})\\
&\qquad\cdot d\hat{F}(a_1)\dots d\hat{F}(a_m)\,.
\end{aligned} \tag{36}$$

Combining both upper bounds (35) and (36) in integral form and using the explicit description of the events "$\mathrm{conv}(\Delta_\star)$ is facet of 1st kind" and "$\mathrm{conv}(\Delta_\star)$ is facet of 2nd kind" and exploiting $\Delta_\star = \{1,\dots,n\}$, we obtain

$$E_{m,n}\Big[\sum_\Delta 2\arccos(h(\Delta))\, I(\mathrm{conv}(\Delta) \text{ is facet of 1st kind})\; I\Big(\arccos(h(\Delta)) > \frac{\sigma}{2}\Big)\Big]$$

$$+\quad E_{m,n}\Big[\pi\cdot\sum_\Delta I(\mathrm{conv}(\Delta) \text{ is facet of 2nd kind})\Big] \tag{37}$$

$$\begin{aligned}
\leq 2\binom{m}{n}\int\limits_{\mathbb{R}^n}^{(m)} \Big\langle\, &\arccos(h(a_1,\dots,a_n))\, I\big(a_{n+1},\dots,a_m \in H^-(a_1,\dots,a_n)\big)\\
&\cdot I\Big(\arccos(h(a_1,\dots,a_n)) > \frac{\sigma}{2}\Big)\\
&+ \arccos(-h(a_1,\dots,a_n))\, I\big(a_{n+1},\dots,a_m \in H^+(a_1,\dots,a_n)\big)\Big\rangle\\
&d\hat{F}(a_1)\dots d\hat{F}(a_m)\,.
\end{aligned} \tag{38}$$

Now we apply our transformation of coordinates with $a_i \longrightarrow b_i$ for $i = 1, \ldots, m$ and $b_1^n = \ldots = b_n^n = h = h(a_1, \ldots, a_n)$ to the integral (38). The result is

$$2\binom{m}{n}\,\lambda_{n-1}(\omega_n)\cdot$$

$$\int_0^1 \Big\langle \arccos(h) G(h)^{m-n}\, I\Big(\arccos(h) > \frac{\sigma}{2}\Big) + \arccos(-h)(1 - G(h))^{m-n}\Big\rangle$$

$$\cdot \int_{\mathbb{R}^{n-1}}^{(n)} |\operatorname{Det}(B)| \hat{f}(b_1) \cdots \hat{f}(b_n) d\bar{b}_1 \ldots d\bar{b}_n \, dh\,, \tag{39}$$

where $\lambda_{n-1}(\omega_n)$ denotes the Lebesgue-measure of the unit sphere. Analogously, $\lambda_n(\Omega_n)$ denotes the Lebesgue-measure of the unit ball. We know (compare [4]):

$$\lambda_{n-1}(\omega_n) = \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} \quad \text{and} \quad \lambda_n(\Omega_n) = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n+2}{2})}\,. \tag{40}$$

$G(h)$ is the marginal distribution function of our special $\mathbb{R}^n$-distribution, given by $G(h) := P(x^n \leq h)$ with $G : [-1, 1] \rightarrow [0, 1]$. $g(h)$ is the corresponding density function with $\int_{-1}^{h} g(\mathfrak{h}) d\mathfrak{h} = G(h)$. For $G(h)$ we know that

$$G(h) = \begin{cases} 1 - \dfrac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \displaystyle\int_h^1 \int_{\frac{h}{r}}^1 (1 - \sigma^2)^{\frac{n-3}{2}} d\sigma dF(r) & \text{for } h \geq 0\,, \\ \dfrac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \displaystyle\int_{|h|}^1 \int_{\frac{|h|}{r}}^1 (1 - \sigma^2)^{\frac{n-3}{2}} d\sigma dF(r)\,, & \text{for } h < 0\,, \end{cases} \tag{41}$$

where $F(r)$ is the radial-distribution function under consideration. In that terminology

$$g(h) = \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \int_{|h|}^1 \frac{(r^2 - h^2)^{\frac{n-3}{2}}}{r^{n-2}} dF(r)\,. \tag{42}$$

The matrix B is

$$B = \begin{pmatrix} b_1^1 & \cdots & b_1^{n-1} & 1 \\ \vdots & & \vdots & \vdots \\ b_n^1 & \cdots & b_n^{n-1} & 1 \end{pmatrix}$$

and enters the integral via the Jacobian of the transformation, which is taken into account by $\lambda_{n-1}(\omega_n) \cdot |\operatorname{Det}(B)|$.

$G(h)^{m-n}$ reflects the requirement that the points $a_{n+1}, \ldots, a_m$ all have to lie below the level $x^n = h$ and $(1-G(h))^{m-n}$ is the probability for the condition that all these points have to be beyond that level, i. e. belong to the opposite halfspace. For future purposes we introduce also the function

$$g_2(h) := \frac{\lambda_{n-2}(\omega_{n-1})}{(n-1)\ \lambda_{n-1}(\omega_n)} \int\limits_{|h|}^{1} \frac{(r^2-h^2)^{\frac{n-1}{2}}}{r^{n-2}} dF(r)\,, \tag{43}$$

and we notice that for all $h \in [-1,1]$

$$G(h) = 1 - G(-h), \quad g(h) = g(-h) \quad \text{and} \quad g_2(h) = g_2(-h)\,. \tag{44}$$

As we had announced before, we concentrate and restrict our calculations to the family (4) of "z-distributions" $F_z(r)$. For these distributions we obtain special versions of the above-mentioned functions, namely

$$G_z(h) = \begin{cases} \text{for } h \geq 0 : \\ 1 - \dfrac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \displaystyle\int\limits_{h}^{1}\int\limits_{\frac{h}{r}}^{1} (1-\sigma^2)^{\frac{n-3}{2}} d\sigma (1-r^2)^z r^{n-1} dr \, \dfrac{2\Gamma(z+\frac{n}{2}+1)}{\Gamma(z+1)\Gamma(\frac{n}{2})} \\ \text{for } h < 0 : \\ \dfrac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \displaystyle\int\limits_{h}^{1}\int\limits_{\frac{h}{r}}^{1} (1-\sigma^2)^{\frac{n-3}{2}} d\sigma (1-r^2)^z r^{n-1} dr \, \dfrac{2\Gamma(z+\frac{n}{2}+1)}{\Gamma(z+1)\Gamma(\frac{n}{2})} \end{cases} \tag{45}$$

$$\begin{aligned} g_z(h) &= \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \int\limits_{|h|}^{1} \frac{(r^2-h^2)^{\frac{n-3}{2}}}{r^{n-2}} (1-r^2)^z r^{n-1} dr \cdot \frac{2\Gamma(z+\frac{n}{2}+1)}{\Gamma(z+1)\Gamma(\frac{n}{2})} \\ &= \frac{\Gamma(z+\frac{n}{2}+1)}{\sqrt{\pi}\Gamma(z+\frac{n}{2}+\frac{1}{2})} \cdot (1-h^2)^{\frac{n-1}{2}+z}\,, \end{aligned} \tag{46}$$

$$\begin{aligned} g_{2,z}(h) &= \frac{\lambda_{n-2}(\omega_{n-1})}{(n-1)\ \lambda_{n-1}(\omega_n)} \int\limits_{|h|}^{1} \frac{(r^2-h^2)^{\frac{n-1}{2}}}{r^{n-2}} (1-r^2)^z r^{n-1} dr \cdot \frac{2\Gamma(z+\frac{n}{2}+1)}{\Gamma(z+1)\Gamma(\frac{n}{2})} \\ &= \frac{\Gamma(z+\frac{n}{2}+1)}{2\sqrt{\pi}\Gamma(z+\frac{n}{2}+\frac{3}{2})} \cdot (1-h^2)^{\frac{n+1}{2}+z}\,. \end{aligned} \tag{47}$$

An evaluation of the inner integral in (39) based on our special distribution (compare [10](p. 40) and [8](p. 48/49)) delivers:

$$\int\limits_{\mathbb{R}^{n-1}}^{(n)} |\operatorname{Det}(B)| \hat{f}_z(b_1) \cdots \hat{f}_z(b_n) d\bar{b}_1 \ldots d\bar{b}_n =$$

$$= \frac{(n!)\ \lambda_{n-1}(\Omega_{n-1})}{(2\pi)^{n-1}} \frac{\Gamma(z+\frac{n}{2}+\frac{1}{2})\Gamma(n(z+\frac{n}{2}))}{\Gamma(z+\frac{n}{2})\Gamma(n(z+\frac{n}{2})+\frac{1}{2})} \cdot g_z(h)\cdot(1-h^2)^{(n-1)(z+\frac{n}{2})}\,.$$

We conclude that for an arbitrary z-distribution

$$(37) \le 2\binom{m}{n}\ \lambda_{n-1}(\omega_n)\cdot\frac{(n!)\ \lambda_{n-1}(\Omega_{n-1})}{(2\pi)^{n-1}} \frac{\Gamma(z+\frac{n}{2}+\frac{1}{2})\Gamma(n(z+\frac{n}{2}))}{\Gamma(z+\frac{n}{2})\Gamma(n(z+\frac{n}{2})+\frac{1}{2})}\cdot$$

$$\int_0^1 \left(\arccos(h)G_z(h)^{m-n}\, I\left(\arccos(h) > \frac{\sigma}{2}\right) + \arccos(-h)(1-G_z(h))^{m-n}\right)$$

$$\cdot g_z(h)\cdot(1-h^2)^{(n-1)(z+\frac{n}{2})}dh\,.$$

If we take (44) into regard, then it is also possible to write the upper bound in the form

$$\begin{aligned}
(37) \quad &\le\quad 2\binom{m}{n}\ \lambda_{n-1}(\omega_n)\cdot\frac{(n!)\ \lambda_{n-1}(\Omega_{n-1})}{(2\pi)^{n-1}} \frac{\Gamma(z+\frac{n}{2}+\frac{1}{2})\Gamma(n(z+\frac{n}{2}))}{\Gamma(z+\frac{n}{2})\Gamma(n(z+\frac{n}{2})+\frac{1}{2})} \\
&\qquad \cdot \int_{-1}^{\cos(\frac{\sigma}{2})} \arccos(h)G_z(h)^{m-n}g_z(h)\cdot(1-h^2)^{(n-1)(z+\frac{n}{2})}dh \\
&=\quad \binom{m}{n}\cdot C_1(n)\cdot C_2(z,n) \\
&\qquad \cdot \int_{-1}^{\cos(\frac{\sigma}{2})} \arccos(h)G_z(h)^{m-n}g_z(h)\cdot(1-h^2)^{(n-1)(z+\frac{n}{2})}dh\,.
\end{aligned} \tag{48}$$

with

$$C_1(n) := \frac{2(n!)\ \lambda_{n-1}(\Omega_{n-1})\ \lambda_{n-1}(\omega_n)}{(2\pi)^{n-1}} \quad \text{and} \tag{49}$$

$$C_2(z,n) := \frac{\Gamma(z+\frac{n}{2}+\frac{1}{2})\Gamma(n(z+\frac{n}{2}))}{\Gamma(z+\frac{n}{2})\Gamma(n(z+\frac{n}{2})+\frac{1}{2})}\,. \tag{50}$$

After abbreviating the description of the upper bound, we want to simplify the integral in (48) as much as possible. This can be achieved by an attempt to estimate $\arccos(h)$ by a function of $G(h)$.
For $\eta\in[0,1]$ we have

$$\int_\eta^1 \sqrt{1-\sigma^2}^{\,n-3}d\sigma \ge \int_\eta^1 \sigma\sqrt{1-\sigma^2}^{\,n-3}d\sigma = \frac{1}{n-1}(1-\eta^2)^{(n-1)/2}\,, \tag{51}$$

and therefore we get for $h \in [0,1]$:

$$\begin{aligned}
1 - G_z(h) &= \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \int_h^1 \int_{\frac{h}{r}}^1 (1-\sigma^2)^{\frac{n-3}{2}} d\sigma dF_z(r) \\
&\geq \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \frac{1}{n-1} \int_h^1 \left(1 - \frac{h^2}{r^2}\right)^{\frac{n-1}{2}} dF_z(r) \\
&\geq \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \frac{1}{n-1} \int_h^1 \frac{(r^2-h^2)^{\frac{n-1}{2}}}{r^{n-2}} dF_z(r) \\
&= g_{2,z}(h) = \frac{\Gamma(z+\frac{n}{2}+1)}{2\sqrt{\pi}\Gamma(z+\frac{n}{2}+\frac{3}{2})} \cdot (1-h^2)^{\frac{n+1}{2}+z} \qquad (52) \\
&\geq \frac{\Gamma(z+\frac{n}{2}+1)}{2\sqrt{\pi}\Gamma(z+\frac{n}{2}+\frac{3}{2})} \cdot \left(\frac{2}{\pi}\arccos(h)\right)^{n+1+2z} .
\end{aligned}$$

In this chain of approximations we have first used (51), then $\frac{1}{r} \geq 1$, and the equalities to $g_{2,z}$ are clear, if we remember (43) and (47). The last estimation follows from

$$\sqrt{1-h^2} \geq \frac{2}{\pi}\arcsin(\sqrt{1-h^2}) = \frac{2}{\pi}\arccos(h) \quad \text{for } h \in [0,1].$$

Now, we can conclude from the result of (52) that for $h \in [0,1]$:

$$\arccos(h) \leq \frac{\pi}{2}\left(\frac{2\sqrt{\pi}\Gamma(z+\frac{n}{2}+\frac{3}{2})}{\Gamma(z+\frac{n}{2}+1)}\right)^{\frac{1}{n+1+2z}} \Big(1 - G_z(h)\Big)^{\frac{1}{n+1+2z}} . \qquad (53)$$

On $-1 \leq h \leq 0$ we have

$$1 - G_z(h) \geq 1 - G_z(0) = \frac{1}{2} \geq \frac{1}{2}\left(\frac{\arccos(h)}{\pi}\right)^{n+1+2z}$$

because $\frac{\pi}{2} \leq \arccos(h) \leq \pi$ for $h \in [-1,0]$.

$$\Rightarrow \quad \arccos(h) \leq 2^{\frac{1}{n+1+2z}}\pi\Big(1 - G_z(h)\Big)^{\frac{1}{n+1+2z}} . \qquad (54)$$

A uniform bound of this kind for the total interval $h \in [-1,1]$ comes by combining (53) and (54) in the form

$$\begin{aligned}
\arccos(h) &\leq \operatorname{Max}\left\{\frac{\pi}{2}\left(\frac{2\sqrt{\pi}\Gamma(z+\frac{n}{2}+\frac{3}{2})}{\Gamma(z+\frac{n}{2}+1)}\right)^{\frac{1}{n+1+2z}}, 2^{\frac{1}{n+1+2z}}\pi\right\} \\
&\quad \cdot \Big(1 - G_z(h)\Big)^{\frac{1}{n+1+2z}} \\
&\leq 4.5 \cdot \Big(1 - G_z(h)\Big)^{\frac{1}{n+1+2z}} . \qquad (55)
\end{aligned}$$

This is true for all $n \geq 3$ and $z \geq -1$ because of

$$\frac{\pi}{2}\left(\frac{2\sqrt{\pi}\Gamma(z+\frac{n}{2}+\frac{3}{2})}{\Gamma(z+\frac{n}{2}+1)}\right)^{\frac{1}{n+1+2z}} = 2^{\frac{1}{n+1+2z}}\pi\cdot\frac{1}{2}\left(\frac{\sqrt{\pi}\Gamma(z+\frac{n}{2}+\frac{3}{2})}{\Gamma(z+\frac{n}{2}+1)}\right)^{\frac{1}{n+1+2z}}$$
$$\leq 2^{\frac{1}{n+1+2z}}\pi\cdot\frac{1}{2}\left(\sqrt{\pi}\sqrt{z+\frac{n}{2}+1}\right)^{\frac{1}{n+1+2z}} \leq 2^{\frac{1}{n+1+2z}}\pi\cdot\frac{1}{2}\left(\sqrt{\pi}\sqrt{\frac{3}{2}}\right)^{\frac{1}{2}}$$
$$\leq 2^{\frac{1}{n+1+2z}}\pi \leq \sqrt{2}\pi \leq 4.5\,.$$

For the first estimation we exploit, that for $x \geq 1$ and $\alpha \in (0,1)$ we have $\frac{\Gamma(x+\alpha)}{\Gamma(x)} \leq x^\alpha$ (compare [4], Appendix), and for the second inequality we used the monotonicity of the function $\sqrt{x+1}^{\frac{1}{x}}$ for $x > 0$.
The bound (55) for $\arccos(h)$ leads to

$$(48) \;\leq\; \binom{m}{n} C_1(n)\, C_2(z,n)\cdot 4.5 \cdot \int_{-1}^{\cos(\frac{\sigma}{2})} G_z(h)^{m-n}\Big(1-G_z(h)\Big)^{\frac{1}{n+1+2z}} g_z(h)\cdot(1-h^2)^{(n-1)(z+\frac{n}{2})}dh\,. \tag{56}$$

Next, we want to estimate $(1-h^2)$ by a function of $G(h)$. Remembering the definition (47) of $g_{2,z}(h)$ we get

$$(1-h^2)^{\frac{1}{2}} = \left(\frac{2\sqrt{\pi}\Gamma(z+\frac{n}{2}+\frac{3}{2})}{\Gamma(z+\frac{n}{2}+1)}\right)^{\frac{1}{n+1+2z}} g_{2,z}(h)^{\frac{1}{n+1+2z}}$$

and further, we know for $h \in [0,1]$ (compare the estimations leading to (52)):

$$g_{2,z}(h) \leq 1 - G_z(h)$$

and also for $h \in [-1,0]$:

$$g_{2,z}(h) = g_{2,z}(-h) \leq 1 - G_z(-h) \leq 1 - G_z(h)\,.$$

So, we have

$$(56) \;\leq\; \binom{m}{n} C_1(n)\, C_2(z,n)\, 4.5\cdot C_3(z,n) \cdot \int_{-1}^{\cos(\frac{\sigma}{2})} G_z(h)^{m-n}\Big(1-G_z(h)\Big)^{\frac{1}{n+1+2z}} g_z(h)\cdot g_{2,z}(h)^{\frac{2}{n+1+2z}(n-1)(z+\frac{n}{2})}dh$$
$$\leq \binom{m}{n} C_1(n)\, C_2(z,n)\, 4.5\cdot C_3(z,n) \cdot \int_{-1}^{\cos(\frac{\sigma}{2})} G_z(h)^{m-n} g_z(h)\cdot\Big(1-G_z(h)\Big)^{\frac{(n-1)(2z+n)+1}{n+1+2z}} dh \tag{57}$$

with

$$C_3(z,n) := \left(\frac{2\sqrt{\pi}\Gamma(z+\frac{n}{2}+\frac{3}{2})}{\Gamma(z+\frac{n}{2}+1)}\right)^{\frac{(n-1)(2z+n)}{n+1+2z}} . \tag{58}$$

Now we make use of the transformation/substitution $x = 1 - G_z(h)$ and $\frac{dx}{dh} = -g_z(h)$. This yields

$$(57) \;=\; \binom{m}{n} C_1(n)\, C_2(z,n)\, 4.5 \cdot C_3(z,n) \cdot \int\limits_{x(\cos(\frac{\sigma}{2}))}^{1} (1-x)^{m-n} x^{\frac{(n-1)(2z+n)+1}{n+1+2z}}\, dx$$

with $x(\cos(\frac{\sigma}{2})) := 1 - G_z(\cos(\frac{\sigma}{2}))$.

A chance for information on such truncated Beta-integrals lies in an application of the Markov-inequality, which states that for any $\varepsilon > 0$ and every positive exponent $\rho > 0$ for a random variable X:

$$P(|X| \geq \epsilon) \leq \frac{E[|X|^\rho]}{\epsilon^\rho} .$$

From that we can conclude:

$$\begin{aligned}
&\int\limits_{x(\cos(\frac{\sigma}{2}))}^{1} (1-x)^{m-n} x^{\frac{(n-1)(2z+n)+1}{n+1+2z}}\, dx \\
&\quad \leq \frac{1}{x(\cos(\frac{\sigma}{2}))^\rho} \int\limits_{0}^{1} (1-x)^{m-n} x^{\frac{(n-1)(2z+n)+1}{n+1+2z}+\rho}\, dx \\
&\quad = \frac{1}{x(\cos(\frac{\sigma}{2}))^\rho} \cdot \frac{\Gamma(m-n+1)\Gamma(n-1+\rho+\frac{2z+3}{n+1+2z})}{\Gamma(m+\rho+\frac{2z+3}{n+1+2z}))} .
\end{aligned}$$

So we arrive at the upper bound

$$\begin{aligned}
(57) \;\leq\; &\binom{m}{n} C_1(n)\, C_2(z,n)\, 4.5 \cdot C_3(z,n) \\
&\cdot \frac{1}{x(\cos(\frac{\sigma}{2}))^\rho} \cdot \frac{\Gamma(m-n+1)\Gamma(n-1+\rho+\frac{2z+3}{n+1+2z})}{\Gamma(m+\rho+\frac{2z+3}{n+1+2z}))} .
\end{aligned} \tag{59}$$

We want to estimate the single factors. If we take into regard that $\Gamma(\frac{n+1}{2})\Gamma(\frac{n}{2}) = \frac{(n-1)!\sqrt{\pi}}{2^{n-1}}$ and $(x+\alpha-1)^\alpha \leq \frac{\Gamma(x+\alpha)}{\Gamma(x)} \leq x^\alpha$ for $x \geq 1$ and $\alpha \in (0,1)$, then we get from (49), (50), (58) and (40)

$$C_1(n) \;=\; \frac{2(n!)\ \lambda_{n-1}(\Omega_{n-1})\ \lambda_{n-1}(\omega_n)}{(2\pi)^{n-1}} = \frac{2(n!)\pi^{\frac{n-1}{2}} 2\pi^{\frac{n}{2}}}{(2\pi)^{n-1}\Gamma(\frac{n+1}{2})\Gamma(\frac{n}{2})}$$

$$
\begin{aligned}
&= \frac{4(n!)\pi^{\frac{1}{2}}}{2^{n-1}\Gamma(\frac{n+1}{2})\Gamma(\frac{n}{2})} = \frac{4(n!)\pi^{\frac{1}{2}}2^{n-1}}{2^{n-1}(n-1)!\sqrt{\pi}} = 4n\,, \\
C_2(z,n) &= \frac{\Gamma(z+\frac{n}{2}+\frac{1}{2})\Gamma(n(z+\frac{n}{2}))}{\Gamma(z+\frac{n}{2})\Gamma(n(z+\frac{n}{2})+\frac{1}{2})} \leq \frac{(z+\frac{n}{2})^{\frac{1}{2}}}{(n(z+\frac{n}{2})-\frac{1}{2})^{\frac{1}{2}}} \\
&\leq \frac{1}{\sqrt{n-1}}\,, \\
C_3(z,n) &= \left(\frac{2\sqrt{\pi}\Gamma(z+\frac{n}{2}+\frac{3}{2})}{\Gamma(z+\frac{n}{2}+1)}\right)^{\frac{(n-1)(2z+n)}{n+1+2z}} \leq \left(2\sqrt{\pi}(z+\frac{n}{2}+1)^{\frac{1}{2}}\right)^{\frac{(n-1)(2z+n)}{n+1+2z}}.
\end{aligned}
$$

The quotient of Γ-functions in (59) can be estimated as follows:

$$
\begin{aligned}
&\binom{m}{n}\frac{\Gamma(m-n+1)\Gamma(n-1+\rho+\frac{2z+3}{n+1+2z})}{\Gamma(m+\rho+\frac{2z+3}{n+1+2z}))} = \frac{\Gamma(m+1)\Gamma(n-1+\rho+\frac{2z+3}{n+1+2z})}{\Gamma(n+1)\Gamma(m+\rho+\frac{2z+3}{n+1+2z}))} \\
&= \frac{\Gamma(m+1)}{\Gamma(m+\lfloor\rho+\frac{2z+3}{n+1+2z}\rfloor)}\frac{\Gamma(m+\lfloor\rho+\frac{2z+3}{n+1+2z}\rfloor)}{\Gamma(m+\rho+\frac{2z+3}{n+1+2z})} \\
&\quad\cdot\frac{\Gamma(n-1+\lfloor\rho+\frac{2z+3}{n+1+2z}\rfloor)}{\Gamma(n+1)}\frac{\Gamma(n-1+\rho+\frac{2z+3}{n+1+2z})}{\Gamma(n-1+\lfloor\rho+\frac{2z+3}{n+1+2z}\rfloor)} \\
&\leq \frac{1}{m^{\lfloor\rho+\frac{2z+3}{n+1+2z}\rfloor-1}}\frac{1}{(m+\rho+\frac{2z+3}{n+1+2z}-1)^{\rho+\frac{2z+3}{n+1+2z}-\lfloor\rho+\frac{2z+3}{n+1+2z}\rfloor}} \\
&\quad\cdot\left(n-2+\lfloor\rho+\frac{2z+3}{n+1+2z}\rfloor\right)^{\lfloor\rho+\frac{2z+3}{n+1+2z}\rfloor-2} \\
&\quad\cdot\left(n-1+\rho+\frac{2z+3}{n+1+2z}\right)^{\rho+\frac{2z+3}{n+1+2z}-\lfloor\rho+\frac{2z+3}{n+1+2z}\rfloor} \\
&\leq (n+\rho)^{\rho-2+\frac{2z+3}{n+1+2z}}\cdot m^{-\rho+1-\frac{2z+3}{n+1+2z}} \qquad \text{for } \rho+\frac{2z+3}{n+1+2z}\geq 1\,.
\end{aligned}
$$

Now, we have

$$
(57) \;\leq\; \frac{1}{x(\cos(\frac{\sigma}{2}))^{\rho}}\frac{4n\cdot 4.5}{\sqrt{n-1}}\left(2\sqrt{\pi}(z+\tfrac{n}{2}+1)^{\frac{1}{2}}\right)^{\frac{(n-1)(2z+n)}{n+1+2z}}\frac{(n+\rho)^{\rho-2+\frac{2z+3}{n+1+2z}}}{m^{\rho-1+\frac{2z+3}{n+1+2z}}}
$$

and this yields (compare (32) and (33)):

$E_{m,n}[\sigma\text{-pathgap}]$

$$\begin{aligned}
&\geq E_{m,n}[\measuredangle(x_0, v)] \\
&\quad - E_{m,n}\Big[\sum_{\Delta} 2\arccos(h_\Delta)\, I(\text{conv}(\Delta) \text{ is facet of 1st kind}) \\
&\qquad\qquad I\Big(\arccos(h_\Delta) > \frac{\sigma}{2}\Big)\Big] \\
&\quad - E_{m,n}\Big[\pi \cdot \sum_{\Delta} I(\text{conv}(\Delta) \text{ is facet of 2nd kind})\Big] \\
&\geq \frac{\pi}{2} - \frac{18n \cdot \left(2\sqrt{\pi}(z+\frac{n}{2}+1)^{\frac{1}{2}}\right)^{\frac{(n-1)(2z+n)}{n+1+2z}}}{x(\cos(\frac{\sigma}{2}))^{\rho} \cdot \sqrt{n-1}} \cdot \frac{(n+\rho)^{\rho-2+\frac{2z+3}{n+1+2z}}}{m^{\rho-1+\frac{2z+3}{n+1+2z}}}. \qquad (60)
\end{aligned}$$

3.2 Discussion of the parameter σ

We have to look for a value σ such that the difference in (60) remains positive and that we get a good lower bound on $\dfrac{E_{m,n}[\sigma\text{-pathgap}]}{\sigma}$.
If we determine a value $\bar{\sigma}$ in way that the term in (60) is at least $\frac{\pi}{4}$ then we get a lower bound of the form

$$\begin{aligned}
E_{m,n}[s+1] &\geq \frac{E_{m,n}[\bar{\sigma}\text{-pathgap}]}{\bar{\sigma}} \\
&\geq \frac{1}{\bar{\sigma}}\left(\frac{\pi}{2} - \frac{18n \cdot \left(2\sqrt{\pi(z+\frac{n}{2}+1)}\right)^{\frac{(n-1)(2z+n)}{n+1+2z}}}{x(\cos(\frac{\bar{\sigma}}{2}))^{\rho} \cdot \sqrt{n-1}} \cdot \frac{(n+\rho)^{\rho-2+\frac{2z+3}{n+1+2z}}}{m^{\rho-1+\frac{2z+3}{n+1+2z}}}\right) \\
&\geq \frac{1}{\bar{\sigma}} \cdot \frac{\pi}{4}. \qquad (61)
\end{aligned}$$

Since we want the term in (60) to be at least $\frac{\pi}{4}$, this requires:

$$\frac{\pi}{2} - \frac{1}{x(\cos(\frac{\sigma}{2}))^{\rho}} \cdot \frac{18n \cdot \left(2\sqrt{\pi(z+\frac{n}{2}+1)\,s}\right)^{\frac{(n-1)(2z+n)}{n+1+2z}}}{\sqrt{n-1}} \cdot \frac{(n+\rho)^{\rho-2+\frac{2z+3}{n+1+2z}}}{m^{\rho-1+\frac{2z+3}{n+1+2z}}} \geq \frac{\pi}{4}$$

$$\Leftrightarrow \quad x(\cos(\tfrac{\sigma}{2})) \geq \frac{(4\cdot 18n)^{\frac{1}{\rho}}\left(2\sqrt{\pi(z+\frac{n}{2}+1)}\right)^{\frac{(n-1)(2z+n)}{\rho(n+1+2z)}}}{(\pi\sqrt{n-1})^{\frac{1}{\rho}}} \cdot \frac{(n+\rho)^{1-\frac{2}{\rho}+\frac{2z+3}{\rho(n+1+2z)}}}{m^{1-\frac{1}{\rho}+\frac{2z+3}{\rho(n+1+2z)}}}.$$

We remember from (52) that

$$\begin{aligned}
x(\cos(\tfrac{\sigma}{2})) &= 1 - G_z(\cos(\tfrac{\sigma}{2})) \\
&\geq g_{2,z}(\cos(\tfrac{\sigma}{2})) = \frac{\Gamma(z+\frac{n}{2}+1)}{2\sqrt{\pi}\Gamma(z+\frac{n}{2}+\frac{3}{2})} \cdot (1-\cos(\tfrac{\sigma}{2})^2)^{\frac{n+1}{2}+z} \\
&\geq \frac{\Gamma(z+\frac{n}{2}+1)}{2\sqrt{\pi}\Gamma(z+\frac{n}{2}+\frac{3}{2})} \cdot \left(\frac{\sigma}{\pi}\right)^{n+1+2z}
\end{aligned}$$

because of $1-\cos(\frac{\sigma}{2})^2 = \sin(\frac{\sigma}{2})^2 \geq \left(\frac{2}{\pi}\frac{\sigma}{2}\right)^2$. So it suffices to choose σ in such a way that

$$\frac{\Gamma(z+\frac{n}{2}+1)}{2\sqrt{\pi}\Gamma(z+\frac{n}{2}+\frac{3}{2})}\cdot\left(\frac{\sigma}{\pi}\right)^{n+1+2z}$$
$$\geq \frac{(72n)^{\frac{1}{\rho}}\left(2\sqrt{\pi(z+\frac{n}{2}+1)}\,\right)^{\frac{(n-1)(2z+n)}{\rho(n+1+2z)}}}{(\pi\sqrt{n-1})^{\frac{1}{\rho}}}\cdot\frac{(n+\rho)^{1-\frac{2}{\rho}+\frac{2z+3}{\rho(n+1+2z)}}}{m^{1-\frac{1}{\rho}+\frac{2z+3}{\rho(n+1+2z)}}},$$

which is equivalent to

$$\sigma \;\geq\; \pi\left(\frac{2\sqrt{\pi}\Gamma(z+\frac{n}{2}+\frac{3}{2})}{\Gamma(z+\frac{n}{2}+1)}\right)^{\frac{1}{n+1+2z}}\cdot$$
$$\frac{(72n)^{\frac{1}{\rho(n+1+2z)}}\left(2\sqrt{\pi(z+\frac{n}{2}+1)}\,\right)^{\frac{(n-1)(2z+n)}{\rho(n+1+2z)^2}}}{(\pi\sqrt{n-1})^{\frac{1}{\rho(n+1+2z)}}}\left(\frac{(n+\rho)^{1-\frac{2}{\rho}+\frac{2z+3}{\rho(n+1+2z)}}}{m^{1-\frac{1}{\rho}+\frac{2z+3}{\rho(n+1+2z)}}}\right)^{\frac{1}{n+1+2z}}$$
$$=:\; \bar{\sigma}\,. \tag{62}$$

Finally, we remember from (61) that $\dfrac{1}{\bar{\sigma}}\cdot\dfrac{\pi}{4}$ yields a lower bound on $E_{m,n}[s+1]$ in the form

$$E_{m,n}[s+1] \geq$$
$$\geq \frac{1}{4}\left(\frac{\Gamma(z+\frac{n}{2}+1)}{2\sqrt{\pi}\Gamma(z+\frac{n}{2}+\frac{3}{2})}\right)^{\frac{1}{n+1+2z}}\cdot \tag{63}$$
$$\frac{(\pi\sqrt{n-1})^{\frac{1}{\rho(n+1+2z)}}}{(72n)^{\frac{1}{\rho(n+1+2z)}}\left(2\sqrt{\pi(z+\frac{n}{2}+1)}\,\right)^{\frac{(n-1)(2z+n)}{\rho(n+1+2z)^2}}}\cdot\left(\frac{m^{1-\frac{1}{\rho}+\frac{2z+3}{\rho(n+1+2z)}}}{(n+\rho)^{1-\frac{2}{\rho}+\frac{2z+3}{\rho(n+1+2z)}}}\right)^{\frac{1}{n+1+2z}}.$$

This is the lower bound on $E_{m,n}[s+1]$ that we have aimed for. But since the formula in (63) looks very complicated, we try to simplify some of its parts, in order to get an estimation of its size.
We remark that

$$\left(\frac{\Gamma(z+\frac{n}{2}+1)}{2\sqrt{\pi}\Gamma(z+\frac{n}{2}+\frac{3}{2})}\right)^{\frac{1}{n+1+2z}} \geq \left(\frac{1}{2\sqrt{\pi}(z+\frac{n}{2}+1)^{\frac{1}{2}}}\right)^{\frac{1}{n+1+2z}} \geq \frac{1}{2.1}$$

and

$$\frac{(\pi\sqrt{n-1})^{\frac{1}{\rho(n+1+2z)}}}{(72n)^{\frac{1}{\rho(n+1+2z)}}\left(2\sqrt{\pi(z+\frac{n}{2}+1)}\,\right)^{\frac{(n-1)(2z+n)}{\rho(n+1+2z)^2}}}$$

$$
\begin{aligned}
&= \left(\sqrt{(2z+n+2)}\right)^{-\frac{(n-1)(2z+n)}{\rho(n+1+2z)^2}} \cdot \frac{\left(\pi\sqrt{n-1}\right)^{\frac{1}{\rho(n+1+2z)}}}{(72n)^{\frac{1}{\rho(n+1+2z)}}\left(\sqrt{2\pi}\right)^{\frac{(n-1)(2z+n)}{\rho(n+1+2z)^2}}} \\
&\geq \left(\sqrt{(2z+n+2)}\right)^{-\frac{(n-1)}{\rho(n+1+2z)}} \cdot const.\,,
\end{aligned}
$$

as we have $\left(\frac{\pi\sqrt{n-1}}{(72n)}\right)^{\frac{1}{\rho(n+1+2z)}} \geq \left(\frac{\pi\sqrt{2}}{72\cdot 3}\right)^{\frac{1}{2\rho}} \geq 0.37$ for $\rho \geq 2$ and $\left(\sqrt{2\pi}\right)^{-\frac{(n-1)(2z+n)}{\rho(n+1+2z)^2}} \geq \sqrt{2\pi}^{-\frac{1}{\rho}} \geq 0.63$ for $\rho \geq 2$.
So, this proves the main theorem:

Theorem 1
For $n \geq 3$, $m \geq n$, $\rho \in [2,\infty)$ and $z \in (-1,\infty)$ (from our distribution family) we have

$$
\begin{aligned}
E_{m,n}[s+1] \;&\geq\; const.\, m^{\frac{1}{n+1+2z}-\frac{n-2}{\rho(n+1+2z)^2}} \\
&\quad \cdot (2z+n+2)^{-\frac{n-1}{2\rho(n+1+2z)}}(n+\rho)^{-\frac{1}{n+1+2z}+\frac{2n+2z-1}{\rho(n+1+2z)^2}}\,.
\end{aligned}
$$

The constant is independent of n, m, z, ρ.

3.3 Concluding Remarks

Now we have a lower bound for the expected value of pivot steps, which holds for every variant.

To make an illustrative comparison with the behaviour of the Shadow-Vertex-Algorithm possible, we concentrate on a single distribution ($z \to -1$, i. e. uniform distribution on the unit sphere) and we choose a special value of ρ, namely $\rho = n$. Then we get from Theorem 1

$$
\begin{aligned}
E_{m,n}[s+1] \;&\geq\; const.\, m^{\frac{1}{n-1}-\frac{n-2}{n(n-1)^2}} \cdot (n)^{-\frac{n-1}{2n(n-1)}} \cdot (n+n)^{-\frac{1}{n-1}+\frac{2n-3}{n(n-1)^2}} \\
&= const.\, m^{\frac{1}{n-1}\left(1-\frac{n-2}{n(n-1)}\right)} \cdot (n)^{-\frac{1}{2n}} \cdot (2n)^{-\frac{1}{n-1}+\frac{2n-3}{n(n-1)^2}} \\
&= const_0.\, m^{\frac{1}{n-1}(1-\varepsilon(n))} \cdot n^{-\delta(n)}
\end{aligned}
$$

where $\varepsilon(n) := \frac{n-2}{n(n-1)}$, $\delta(n) := -\frac{1}{2n} - \frac{1}{n-1} + \frac{2n-3}{n(n-1)^2}$ and all constants are taken into regard by $const_0$.. So we see – as announced in the introduction – that the order of growth is essentially $m^{\frac{1}{n-1}} \cdot n^0$.
Remember that we had an upper bound [6] for that configuration with the Shadow-Vertex-Algorithm of $const_1. \cdot m^{\frac{1}{n-1}} n^2$, and an (asymptotic) lower bound [4] of $const_2. \cdot m^{\frac{1}{n-1}} n^2$ for $m \to \infty$ and fixed n.
So the result presented here shows the same behaviour in m, but leaves a gap of

a factor n^2 for possible superior speed of better variants. We remark that even with the Shadow-Vertex-Algorithm the behaviour in n gets better for distributions with increasing parameter z. This shows that the growth in m is an intrinsic feature of our stochastic model and cannot be avoided even when we employ the most sophisticated variant.

Another question arises when we ask for any variant which actually realizes that behaviour of $m^{\frac{1}{n-1}}n^0$, i. e. whether our lower bound is sharp. Here we must understand clearly the concept of our proof. This result is mainly a consequence of our relaxation method described at the end of chapter 2, where we forgot about the specific path and summed up over all vertices of the polyhedron. Our stochastic analysis shows that $m^{\frac{1}{n-1}}n^0$ cannot be increased significantly by doing a better analysis or numerical evaluation on our relaxation. This is the price we have to pay for having a variant–independent lower bound.
We suspect that the sharpest lower bound on Simplex-Paths might be higher (perhaps $m^{\frac{1}{n-1}}n^{\frac{1}{2}}$ or $m^{\frac{1}{n-1}}n^1$), but for proving such a result we are forced to employ the path-generating-rule (pivot rule) of specific variants. That means that such an analysis would have to concentrate on each specific variant. (And with the exception of the Shadow-Vertex-Algorithm, this seems to be impossible from the current point of view.)

A last aspect, that we should mention, is that our result implies a certain information on the average diameter of a polyhedron (the maximum over all shortest paths between two vertices).
In the same manner as we have derived our result on all variants, we could show that the average diameter is bounded from below by

$$const. \cdot m^{\frac{1}{n-1}}n^0 .$$

This is also an evident consequence of the chosen stochastic model.
To improve such bounds remains a challenge for the future.

References

[1] I. Adler and N. Megiddo. A simplex algorithm whose average number of steps is bounded between two quadratic functions of the smaller dimension. *Journal of the Association for Computing Machinery*, 32:871–895, 1985.

[2] R. G. Bland. New finite pivoting rules for the simplex method. *Mathematics of Operations Research*, 2:103–107, 1977.

[3] K. H. Borgwardt. The average number of pivot steps required by the simplex-method is polynomial. *Zeitschrift für Operations Research*, 26:157–177, 1982.

[4] K. H. Borgwardt. *The Simplex Method — A Probabilistic Analysis*. Springer-Verlag, Heidelberg, 1987.

[5] K. H. Borgwardt. Average complexity of a gift-wrapping algorithm for determining the convex hull of randomly given points. *Discrete & Computational Geometry*, 17:79–109, 1997.

[6] K. H. Borgwardt. A sharp upper bound for the expected number of shadow-vertices in the rotation-symmetry-model — quadratic in the smaller and sublinear in the larger dimension of the LP. *Mathematics of Operations Research*, 1998, to appear.

[7] K. H. Borgwardt, R. Damm, R. Donig, and G. Joas. Empirical studies on the average efficiency of simplex-variants under rotation-symmetry. *ORSA Journal on Computing*, 5:249–260, 1993.

[8] P. Huhn. *Schranken für die durchschnittliche Laufzeit des Simplexverfahrens und von Innere-Punkte-Verfahren*. Ph.D. dissertation, Universität Augsburg, 1997. Augsburger Mathematisch–naturwissenschaftliche Schriften, Bd. 16.

[9] V. Klee and G. Minty. How good is the simplex-algorithm? In *Inequalities III*, pages 159–175. Academic Press, 1972.

[10] K.-H. Küfer. *Asymptotische Varianzanalysen in der stochastischen Polyedertheorie*. Ph.D. dissertation, Universität Kaiserslautern, 1992.

[11] K.-H. Küfer. On the variance of the number of pivot steps required by the simplex algorithm. *Zeitschrift für Operations Research*, 42:1–24, 1995.

Summary

We derive a lower bound on the average complexity of the Simplex-Method solving linear programs (LP) of the form: max $v^T x$ s.t. $a_1^T x \leq 1, \ldots, a_m^T x \leq 1$.
We assume these problems to be randomly generated according to the Rotation-Symmetry-Model (RSM), i. e. let $a_1, \ldots, a_m, v$ be distributed independently, identically and symmetrically under rotations on $\mathbb{R}^n \setminus \{0\}$.
Concentrating on a representative family of RSM-distributions with bounded support, we show for every variant *var* of the Simplex-Method, that on the average the number of pivot-steps s^{var} satisfies

$$E_{m,n}[s^{var}] \geq const.\, m^{\frac{1}{n-1}} n^0 \text{ for all pairs } m \geq n \text{ and for all variants.}$$

This result holds, if the selection of the starting vertex for the Simplex-Method has been done independently of the objective v.

Keywords
Linear Programming, Average-Case Complexity of Algorithms, Stochastic Geometry