Interior-Point-Methods:

Worst-Case and Average-Case Analysis of a Phase-I-Algorithm and a Termination Procedure

Petra Huhn and Karl Heinz Borgwardt

University of Augsburg
Institute for Mathematics
86135 Augsburg
Germany
E-mail: huhn@math.uni-augsburg.de, borgwardt@math.uni-augsburg.de

We are interested in the average behaviour of Interior-Point-Methods (IPMs) for Linear Programming problems (LPs). We use the Rotation-Symmetry-Model as the probabilistic model for the average case analysis. This model had been used by Borgwardt in his average case analysis of the Simplex-Method. IPMs solve LPs in three phases. First, one has to find an appropriate starting point, then a sequence of interior points is generated, which converges to the optimal face. Finally, the optimum has to be calculated, as it is not an interior point. We present upper bounds on the average number of iterations in the first and the third phase by looking at random figures of the underlying polyhedron. These bounds show, that IPMs solve LPs in strongly polynomial time in the average case, so only the dimension parameters and not the encoding length of the problem determine the average behaviour of IPMs.

Key Words: Linear Programming, Interior Point Methods, Probabilistic Analysis

1. PREFACE

In 1984 Karmarkar [13] introduced a projective interior-point-method (IPM) with polynomial worst case complexity for linear programming and a new field of research was born — the field of interior-point-methods. Moreover, Karmarkar announced a superior practical performance on large problems compared to the simplex-method, but it took a few years until efficient implementations for IPMs could at least partially confirm Karmarkar's announcement. In practice it has been observed that some variants of IPMs need a number of iterations that seems to be almost independent of the problem dimensions. So, we are looking for a theoretical explanation for this phenomenon. A chance to explain these observations is a probabilistic analysis as it was done for the simplex-method, where different authors proved a strongly polynomial average case complexity (compare Adler et al. [1], [2], Borgwardt [6, 8], Smale [17] and Todd [18]). This would enable us to compare the average behaviour of IPMs and of the simplex-method.

This paper is concerned with the worst case and average case behaviour of IPMs. The algorithms presented in the paper are typical barrier function methods adapted to a particular type of problems. So, the convergency proofs and the worst case analysis are omitted and the reader can refer to den Hertog [9] or other literature on IPMs ([22] a.o.). In fact, the parts of the paper concerned with the average case analysis contain the major contribution and results. As the main focus will be on the average case behaviour, we briefly describe our approach to the probabilistic analysis of algorithms.

We will consider the average running time of a deterministic algorithm when it is applied to problem instances generated according to some probability distribution. The average case analysis of IPMs that we are going to present here is a part of a more extensive project concerned with the probabilistic analysis of algorithms for solving linear programming problems under the same probabilistic model. So, we will use the so-called rotation-symmetry-model from Borgwardt's average case analysis of the simplex-method in [6] and [8] as the stochastic model for our average case analysis of IPMs. In the long term this should lead to a fair comparison between simplex-methods and interior-point-methods.

In this paper we analyze a phase-I algorithm and in phase II a barrier method with a special termination procedure for linear programming problems (LPs).

A phase-I algorithm in the context of IPMs has a threefold purpose. In the case that the LP is feasible and bounded, the phase-I algorithm should provide a suitable starting point for an IPM and otherwise — if the LP is infeasible or unbounded — it should indicate the infeasibility resp. the unboundedness because a further solution process is not needed in these latter cases. Another possibility to provide starting points for an IPM is to transform the problem such that a starting point is known after the transformation. The major drawback of such an approach is that a probabilistic analysis with transformed problems is overcomplicated and hardly to do, as the data of the transformed problems are not distributed according to the original stochastic model and most transformations cause dependencies between different parts of the new data. To avoid the "starting point problem" one can proceed as in Todd [19], where a stochastic model is introduced, that provides starting points. But in Todd's model the righthand side of the problem data is determined by the matrix of the problem and the desired starting point and so the independency of the data is lost again.

As we want to use the rotation-symmetry-model and as we want to work with the original problem data we apply a phase-I algorithm to our LP. This algorithm (for phase-I) is a barrier function method (in the notation of the IPM literature), adapted for approximating the "analytic center" of a polyhedron. After we have ensured the LP to be bounded and after we have found a starting point, we start phase-II. There, we employ a typical barrier-function method to generate interior points that systematically reduce the duality gap. We combine this method with a termination procedure which projects the current iterate (interior point) onto the boundary and checks whether we have found an optimal point or not.

The average case analysis of the IPM (phase-II) is based on the guaranteed deterministic behaviour of the reduction process and on this stopping criterion, which depends on the difference in the objective function values at the best and the second best vertices. This approach is somehow similar to the average case analysis in Anstreicher et al. [4] and [3], Todd et al. [20] and Ye [21], where IPMs are combined with other termination procedures and bounds are given on the average running time under a probabilistic model from Todd [19].

The paper is organized as follows. In Section 2.1 we give a brief introduction to LPs, analytic centers, barrier functions and complexity theory. The phase-I algorithm and its worst case analysis is presented in Section 2.2. Section 2.3 is concerned with the average case analysis of the phase-I algorithm. The next chapter discusses phase II: we analyse a typical barrier function method and a termination criterion and prove worst case complexity results in Section 3.1. A detailed investigation of a special distribution function is given in Section 3.2. This special distribution function is used in Section 3.3 for the average case analysis of the phase II algorithm and the termination criterion.

We show that our phase-I algorithm has a polynomial worst case complexity of $O(mL)$, where m is the number of constraints in a canonical form problem (problem with inequality constraints and no sign-constraints) and L is the encoding length of the problem data. The barrier function method we use in phase II has a worst case complexity of $O(\sqrt{m}L)$, which is polynomial, too.
In the average case analysis we prove under some weak asymptotic conditions ("asymptotic" usually means: n fixed and $m \to \infty$; and here "weak asymptotic" stands for: $m \geq c \cdot n$, where n is the number of variables in the canonical form problem and $c = O(1)$ is a specified constant) that both algorithms have even a better average case complexity — they are strongly polynomial in the average case. In detail, the average case complexity of the phase-I algorithm is $O(m\sqrt{n})$ and the average number of steps of the barrier function method is at most $O(\sqrt{m}\ln m)$.

2. A PHASE-I-ALGORITHM TO START INTERIOR-POINT-METHODS

2.1. Mathematical Introduction

2.1.1. Basic Notations for LPs

We look at linear programming problems of the following type:

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

where $v, x, a_1, \ldots, a_m \in \mathbb{R}^n$ and $m \geq n$, $m, n \in \mathbb{N}$.

The matrix A collects all the restriction vectors a_i and e denotes the vector of all ones in the appropriate dimension, i. e. $A^T = (a_1, \ldots, a_m)$ and $e^T = (1, \ldots, 1)$. We will use this vector of all ones in different dimensions. To clarify the dimension of the respective vector $e \in \mathbb{R}^m$, we will sometimes write $e_{(m)}$. Note that we should distinguish this from unit vectors in $\mathbb{R}^n$, which will be denoted by e_i, $i \in \{1, \ldots, n\}$. In the same way as the vector of all ones we will handle the notation of the null-vector. So, we will denote the null-vector by 0 and if it is necessary to emphasize the dimension the n-dimensional null-vector will be denoted by $0_{(n)}$.

The program (P) will be seen as our *primal* problem — according to a dual forthcoming problem (D). $v^T x$ is called its *objective function*, and $X_P = \{x \in$

$\mathbb{R}^n \,|\, Ax \le e\}$ its *feasible region.* The *interior of* X_P will be denoted by $Int\, X_P$ and the boundary of X_P by ∂X_P.

Often it is convenient to embed (P) into a problem in $\mathbb{R}^{n+m}$ and to reformulate (P) using slack variables:

$$\text{maximize } v^T x \text{ subject to } Ax + s = e,\ s \ge 0\,, \text{ where } s \in \mathbb{R}^m\,. \tag{PS}$$

Note that X_P is the projection of the feasible region of (PS) on $\mathbb{R}^n$, resp. the set of all x, such that a feasible slack s exists. Obviously, such a slack s is a function of the corresponding vector $x \in X_P$, namely $s = s(x) := e - Ax$. Often we shall omit this special emphasis on this dependency. Instead of writing $s(x)$, we will only write s, as long as conflicts or misinterpretations can be excluded.

Remark 2.1. The origin 0 is — in any case — a feasible, interior point of X_P and the point $\binom{0_{(n)}}{e_{(m)}}$ belongs to $X_{PS} := \{\binom{x}{s} \mid Ax + s = e, s \ge 0\}$.

To avoid numerous case-studies and to make a probabilistic evaluation possible, we agree on the following *assumption of nondegeneracy*[1]:

> *Each n-elementic subset of $\{a_1, \ldots, a_m, v\}$ is linearly independent and each $(n+1)$-elementic subset of $\{a_1, \ldots, a_m, -v\}$ is in general position.* (1)

At a later point, it will become obvious, that this assumption does in no way influence the results of our study, because in the rotation-symmetry-model — which is the basis of our evaluation — the set of degenerate problems (those not satisfying (1)) forms a set of probability null.

Remark 2.2. Since $m \ge n$, the assumption of nondegeneracy provides X_P and X_{PS} to be pointed. If X_P is unbounded, then there exists a direction $d \in \mathbb{R}^n \setminus \{0\}$ such that $Ad \le 0$, and because of nondegeneracy there is even a direction d with $Ad < 0$.

Every LP is accompanied by another linear program, the so-called dual problem. The dual problem of (P) is:

$$\text{minimize } e^T y \text{ subject to } A^T y = v,\ y \ge 0. \tag{D}$$

$X_D = \{y \in \mathbb{R}^m \mid A^T y = v, y \ge 0\}$ denotes the *feasible region* of (D).

Remark 2.3. From duality theory we know, that $X_D \ne \emptyset$ if an optimal solution for (P) exists, i. e. the problem (D) cannot be unbounded (assumed to be feasible) because there exists a primal feasible point.

2.1.2. Barrier Function and Analytic Center

[1] This is a slightly different condition of nondegeneracy compared to that in [6] and [10], because we want to guarantee a nondegenerate phase-I polyhedron (as it will be defined in section 2.1.3).

For a point $x \in Int(X_P)$ denote the *barrier function* of X_P by

$$\phi(x) := -\sum_{i=1}^{m} \ln(1 - a_i^T x) = -\sum_{i=1}^{m} \ln s_i(x) = -\sum_{i=1}^{m} \ln s_i \,.$$

We want to mention some properties of this barrier function.

LEMMA 2.1. *1. ϕ is a strictly convex function of x in Int X_P.*
2. $\lim_{x \to \bar{x} \in \partial X_P} \phi(x) = \infty$.
3. If X_P is unbounded, then $\inf_{x \in Int\, X_P} \phi(x) = -\infty$.
4. If X_P is bounded, then $\phi(x)$ is bounded from below and attains its minimum in a unique point $x_{ac} \in Int\, X_P$. This point is called the analytic center *of X_P.*

If X_P is bounded, we want to approximate the analytic center x_{ac} and we want to measure the "distance" of an arbitrary point $x \in X_P$ to x_{ac}.

For this purpose, we introduce some additional terms. In general, we use

$$S := S(x) = diag(s) = diag(e - Ax) \in \mathbb{R}^{m \times m}$$

for the diagonal matrix bearing the components of s in its diagonal entries. We will deal with some vectors and its "diagonal" counterparts in the same way by switching from small letters to capitals and vice versa.

Now, we can define the gradient, the Hessian matrix of $\phi(x)$ and the Newton direction $p(x)$ at x as

$$g(x) \;:=\; \nabla\phi(x) = A^T[S(x)]^{-1}e = A^T S^{-1} e \,, \tag{2}$$

$$H(x) \;:=\; \nabla^2\phi(x) = A^T[S(x)]^{-2}A = A^T S^{-2} A \,, \tag{3}$$

$$p(x) \;:=\; -[H(x)]^{-1}g(x) = -(A^T S^{-2} A)^{-1} A^T S^{-1} e \,. \tag{4}$$

We define a measure $\delta(x)$ for the distance of an interior point x to the analytic center x_{ac} of X_P as the Hessian norm of the Newton direction:

$$\delta(x) \;:=\; ||p(x)||_{H(x)} = ||S^{-1}Ap(x)|| \,. \tag{5}$$

$||\,.\,||$ denotes the Euclidean norm and $||\,.\,||_H$ the Hessian norm.
Note, that $\delta(x) = 0$ implies $g(x) = 0$ and thus $x = x_{ac}$.

Remark 2.4. Another, equivalent definition is

$$\delta(x) = \min_{y \in \mathbb{R}^m} \{||S(x)y - e|| \mid A^T y = 0\} = ||Sy(s) - e||, \tag{6}$$

where the minimizing y is given by

$$y(s) = S^{-1}(I - S^{-1}A(A^T S^{-2} A)^{-1} A^T S^{-1})e = S^{-1}e + S^{-2}Ap(x)\,. \tag{7}$$

To verify the formulas in (7), note that this point $y(s)$ satisfies the Karush-Kuhn-Tucker conditions for the problem "min $(Sy-e)^T(Sy-e)$ subject to $A^T y = 0$" which has a strictly convex objective function.

DEFINITION 2.1. A point $x \in Int\, X_P$ with $\delta(x) < 1$ is called an *approximate center*; for $\tau \in (0,1)$ we call a point $x \in Int\, X_P$ with $\delta(x) \leq \tau < 1$ a *τ-approximate center.*

Again, for a more detailed discussion see [9] and [22].

2.1.3. Phase-I Problem and Phase-I Barrier Function

One purpose of our phase-I algorithm is to decide whether $v^T x$ is unbounded on X_P or not. In the negative case we are supposed to provide an initial point for a certain reduction process of the distance to the $v^T x$-optimal point on X_P. Since we want to make use of barrier functions and of analytic centers, it would be helpful to have a tool which distinguishes between unboundedness of X_P and unboundedness of $v^T x$ on X_P. This is necessary, because no analytic center of X_P exists if X_P is unbounded. Therefore any method determining an analytic center will fail anyway, whether a $v^T x$-optimal point exists or not.

For having such a tool, we define a polyhedron

$$\underline{X}_P := \{x \in \mathbb{R}^n \,|\, Ax \leq e,\ -v^T x \leq 1\}, \tag{8}$$

which differs from X_P by having one additional restriction, namely $v^T x \geq -1$. Based on this new polyhedron we introduce a linear programming problem corresponding to (P) as

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

Together with our assumption of nondegeneracy (1) the introduction of $\underline{X}_P$ will give us the chance to treat the different cases in the appropriate way.

Remark 2.5. If $\underline{X}_P$ is unbounded, then (see Remark 2.2) there exists a direction d with $v^T d > 0$ and $Ad \leq 0$ and thus we know, that the objective function $v^T x$ of (P) resp. ($\underline{\text{P}}$) is unbounded from above on X_P and on $\underline{X}_P$.

We will use the corresponding notation as in the previous section. The slack variables are $s = e - Ax$ and $s_{m+1} = 1 + v^T x$, where this additional slack variable s_{m+1} corresponds to the additional constraint $-v^T x \leq 1$ in $\underline{X}_P$. For the augmented polyhedron the barrier function is

$$\underline{\phi}(x) := -\sum_{i=1}^{m} \ln(1 - a_i^T x) - \ln(1 + v^T x) = -\sum_{i=1}^{m+1} \ln s_i\,. \tag{9}$$

As in section 2.1.2 we need the gradient, Hessian matrix and the Newton direction of $\underline{\phi}$ at x:

$$\underline{g}(x) \;=\; \nabla\underline{\phi}(x) = A^T S^{-1} e - s_{m+1}^{-1} v\,, \tag{10}$$

$$\underline{H}(x) \;=\; \nabla^2\underline{\phi}(x) = A^T S^{-2} A + s_{m+1}^{-2} v v^T \quad \text{and} \tag{11}$$

$$\begin{aligned}\underline{p}(x) &= -\underline{H}(x)^{-1}\underline{g}(x)\\ &= -\left(A^T S^{-2} A + s_{m+1}^{-2} v v^T\right)^{-1}\left(A^T S^{-1} e - s_{m+1}^{-1} v\right).\end{aligned} \tag{12}$$

Remark 2.6. We can make full use of Lemma 2.1 applied to $\underline{\phi}$ and $\underline{X}_P$. If $\underline{X}_P$ is bounded, we will denote the *analytic center* of $\underline{X}_P$ by $\underline{x}_{ac}$.

The measure $\underline{\delta}(x)$ for the distance of a point $x \in Int\, \underline{X}_P$ to the analytic center $\underline{x}_{ac}$ of $\underline{X}_P$ is defined by

$$\underline{\delta}(x) := ||\underline{p}(x)||_{\underline{H}(x)} = ||\underline{S}^{-1}\underline{A}\underline{p}(x)|| \tag{13}$$

$$\text{using} \quad \underline{s} := \begin{pmatrix} s \\ s_{m+1} \end{pmatrix}, \; \underline{y} := \begin{pmatrix} y \\ y_{m+1} \end{pmatrix} \text{ and } \underline{A} := \begin{pmatrix} A \\ -v^T \end{pmatrix}. \tag{14}$$

Then we can formulate $\underline{\delta}(x)$ corresponding to Remark 2.4 as

$$\underline{\delta}(x) = \min_{\underline{y} \in \mathbb{R}^{m+1}} \{||\underline{S}\underline{y} - e_{(m+1)}|| \,|\, (\underline{A})^T\underline{y} = 0\} = ||\underline{S}\underline{y}(\underline{s}) - e_{(m+1)}|| \tag{15}$$

$$\text{with} \quad \underline{y}(\underline{s}) := \underline{S}^{-1}e_{(m+1)} + \underline{S}^{-2}\underline{A}\underline{p}(x). \tag{16}$$

$\underline{g}$, $\underline{H}$ and $\underline{p}$ can be formulated analogously to (2)–(4) as

$$\underline{g} = \underline{A}^T\underline{S}^{-1}e_{(m+1)}, \quad \underline{H} = \underline{A}^T\underline{S}^{-2}\underline{A} \tag{17}$$

$$\text{and} \quad \underline{p} = -\underline{H}^{-1}\underline{g} = -(\underline{A}^T\underline{S}^{-2}\underline{A})^{-1}\underline{A}^T\underline{S}^{-1}e_{(m+1)}. \tag{18}$$

Finally, we repeat some results of deterministic complexity theory.

2.1.4. Complexity Theory

For deriving worst case complexity results we use a discrete complexity model, which admits only rational data. In this context an algorithm is said to be polynomial if the computational effort can be bounded from above polynomially in the encoding length L of the specific problem instance.[2]

The worst-case complexity analysis of our algorithm makes use of the following facts (compare [16]):

LEMMA 2.2. *1. Each basic solution of the system $a_1^Tx \le 1, \ldots a_m^Tx \le 1$ has a Euclidean norm less than 2^L, resp. each vertex of X_P is contained in a ball of radius 2^L (centered at the origin).*

2. If there exists an optimal solution to (P)*, then there exists an optimal solution x_{opt} with $||x_{opt}|| \le 2^L$.*

For the situation, where ($\underline{\text{P}}$) replaces (P) and the polyhedron $\underline{X}_P$ replaces X_P, we can deduce that:

COROLLARY 2.1. *1. The norm (length) $||x||$ of each vertex x of $\underline{X}_P$ is at most 2^L.*

[2] Here L gives the number of bits (resp. digits) which are necessary to store $v, a_1, \ldots, a_m$ correctly.

2. If there exists a point $x \in Int\,\underline{X}_P$ with $||x|| \geq 2^L$, then the objective function $v^T x$ of (P) resp. ($\underline{P}$) is unbounded (from above) on X_P and $\underline{X}_P$.

Remark 2.7. The bound 2^L for the norm of vertices is a worst-case bound. If we look at a specific problem instance of ($\underline{P}$) resp. at $\underline{X}_P$, the maximal norm of a vertex may be extremely smaller than 2^L.

2.2. A Phase-I-Algorithm

2.2.1. Properties of the Barrier Function

We will start by presenting some fundamental lemmata, which prepare for the complexity proof of the algorithm. These lemmata and their proofs are the result of adapting barrier function methods as described in [9] to the phase-I problem ($\underline{P}$) for approximating the analytic center.

Throughout this section we will write $\underline{p}$, $\underline{g}$, $\underline{H}$ and $\underline{\delta}$ instead of $\underline{p}(x)$, $\underline{g}(x)$, $\underline{H}(x)$ and $\underline{\delta}(x)$ for the sake of simplicity of the formulas. This is done whenever there is no doubt about the reference point x.

LEMMA 2.3. *For $x \in Int\,\underline{X}_P$ and $d \in \mathbb{R}^n$ with $||d||_{\underline{H}(x)} < 1$ we have $x + d \in Int\,\underline{X}_P$.*

LEMMA 2.4. *1. For $\underline{\delta} = \underline{\delta}(x) \geq 0$ define $\bar{\alpha} = (1+\underline{\delta})^{-1}$, then we have*

$$\underline{\phi}(x) - \underline{\phi}(x + \bar{\alpha}\underline{p}) \geq \underline{\delta} - \ln(1+\underline{\delta})\,.$$

This means that a Newton step with step length $\bar{\alpha}$ decreases the barrier function by an amount of $\underline{\delta} - \ln(1+\underline{\delta})$ at least.

2. If $\underline{\delta} = \underline{\delta}(x) < 1$, then $\underline{\phi}(x) > -\infty$ for all $x \in \underline{X}_P$, i. e. $\underline{X}_P$ is bounded.

LEMMA 2.5. *Let $T \in \mathbb{R}$, $T > 0$ be fixed. The barrier function $\underline{\phi}(x)$ is bounded from below on the region $\underline{X}_P \cap \{x \in \mathbb{R}^n \mid ||x|| \leq T\}$, i. e. for all $x \in \underline{X}_P \cap \{x \in \mathbb{R}^n \mid ||x|| \leq T\}$ we have*

$$\underline{\phi}(x) \geq -\sum_{i=1}^{m} \ln(1 + ||a_i||T) - \ln(1 + ||v||T)\,.$$

Proof. We get $|a_i^T x| \leq ||a_i||||x|| \leq ||a_i||T$ and $|v^T x| \leq ||v||||x|| \leq ||v||T$ for $x \in \{x \in \mathbb{R}^n \mid ||x|| \leq T\}$ by using the Cauchy-Schwarz inequality and therefore

$$\begin{aligned}\underline{\phi}(x) &= -\sum_{i=1}^{m} \ln(1 - a_i^T x) - \ln(1 + v^T x) \\ &\geq -\sum_{i=1}^{m} \ln(1 + ||a_i||T) - \ln(1 + ||v||T).\end{aligned}$$

∎

Now, we define the lower bound for the barrier function value in Lemma 2.5 and a special T, called $T(\underline{X}_P)$, as follows

$$\underline{\Lambda}(T) := -\sum_{i=1}^{m} \ln(1+||a_i||T) - \ln(1+||v||T) \tag{19}$$

$$\text{and} \quad T(\underline{X}_P) := \mathrm{Max}\{||x|| \mid x \text{ is a vertex of } \underline{X}_P\}. \tag{20}$$

The calculation of $T(\underline{X}_P)$ is as difficult as solving the LP (P), so we will not try to calculate $T(\underline{X}_P)$. But we will use this figure and the insight of Lemma 2.5 for some theoretical considerations implicitly in our algorithm and for the probabilistic analysis. So, the theorem of Krein-Milman and the convexity of the Euclidean unit ball provide the following result.

Remark 2.8. If $\underline{X}_P$ is bounded, then $\underline{X}_P \subseteq \{x \mid ||x|| \le T(\underline{X}_P)\}$. Otherwise, if $\underline{X}_P$ is unbounded, then there exists a point $x \in \underline{X}_P$ with $||x|| > T(\underline{X}_P)$.

2.2.2. *Algorithm and Complexity Analysis*

As our phase-I algorithm shall determine whether the phase-I polyhedron is unbounded or not, we can explore the iterates and the norm of the iterates on being greater or smaller than $T(\underline{X}_P)$. Unfortunately, we do not know the explicit value of $T(\underline{X}_P)$ and so we cannot compare the values $||x||$ for an iterate x with $T(\underline{X}_P)$. But — given a point $x \in \underline{X}_P$ — we can try to find a vertex $\bar{x}$ of $\underline{X}_P$ with $||\bar{x}|| \ge ||x||$. Our test-procedure (for finding such a vertex) is working in a special way, so that we can guarantee the unboundedness of $\underline{X}_P$ in those cases, where the procedure does not find such a vertex.

This kind of "test" can be done by Procedure Test, where we use the following notation: Let ${\underline{a}_i}^T$ denote the i-th row of $\underline{A}$, then

$$I(x) := \{i \mid {\underline{a}_i}^T x = 1\} \subset \{1,\dots,m+1\}$$

is the index set of those constraints, which are active at x. For an index set $I = \{i_1,\dots,i_j\} \subseteq \{1,\dots,m+1\}$ $(j \ge 1)$ we define

$${\underline{A}_I}^T := (a_{i_1},\dots,a_{i_j}) \quad \text{and} \quad P_I := E - {\underline{A}_I}^T(\underline{A}_I {\underline{A}_I}^T)^{-1}\underline{A}_I\,.$$

P_I is a projection matrix on the null space of $\underline{A}_I$.

Algorithm 1 (Procedure Test).

Input: $x \in \mathit{Int}\, \underline{X}_P$

```
l := 0; ξ_l := x; d_l := ξ_l;                         {intialization}
I(ξ_l) := ∅;                              {index set of active constraints}
repeat
   if d_l = 0 then
      return false (exit);           {A vertex ξ of X_P exists with ||ξ|| ≥ ||ξ_0||}
   else
      if A d_l ≤ 0 then
         return true (exit);                   {X_P is unbounded in direction d_l}
```

9. **else**
10. COMPUTE $\alpha_l = \max\{\alpha \mid \alpha \geq 0, \xi_l + \alpha d_l \in \underline{X}_P\}$;
11. $\xi_{l+1} := \xi_l + \alpha_l d_l$;
12. CHOOSE $i_{l+1} \in \{i \mid i \notin I_l, \underline{a}_i{}^T \xi_{l+1} = 1\}$;
13. $I_{l+1} := I_l \cup \{i_{l+1}\}$;
14. COMPUTE $P_{I_{l+1}}$; {*projection matrix*}
15. $d_{l+1} := P_{I_{l+1}} d_l$;
16. $l := l + 1$;
17. **endif**
18. **endif**
19. **until** $l = n$;
20. return *false* (exit); $\{\xi_n$ *is a vertex of* $\underline{X}_P$ *with* $||\xi_n|| \geq ||\xi_0||\}$

Output: *true:* $\underline{X}_P$ is unbounded *or* *false:* there exists a vertex $\bar{x}$ of $\underline{X}_P$ with norm greater than $||x|| = ||\xi_0||$.

THEOREM 2.1 (Complexity and Correctness of ALGORITHM 1).
1. PROCEDURE TEST *terminates after at most n iterations.*
2. The effort of each iteration is at most $O(mn + n^2)$.[3]
3. For all $n > l \geq 0$ *we have* $\xi_{l+1} \in \underline{X}_P$ *and* $||\xi_{l+1}|| \geq ||\xi_l||$.
4. If PROCEDURE TEST *returns* true, *then* $\underline{X}_P$ *is unbounded.*
5. If PROCEDURE TEST *returns* false, *then there exists a vertex* $\bar{x}$ *of* $\underline{X}_P$ *with* $||\bar{x}|| \geq ||\xi_0||$.

Proof. *Part 1:* The variable l is initialized (in line 1) with the value 0, incremented by 1 (in line 16) and we stop if $l = n$ (line 19) at the latest.

Part 2: Testing $\underline{A} d_l \leq 0$ and computing $\bar{\alpha}_l$ raises $O(mn)$ arithmetic operations because we have to take into account $m + 1$ constraints. The computation of the projection matrix $P_{I_{l+1}}$ can be done in $O(n^2)$ arithmetic operations using update formulas. All other statements are done in at most $O(n)$ arithmetic operations.

Part 3: Because of the input and line 10/11 it is clear, that $\xi_{l+1} \in \underline{X}_P$ for all l. Furthermore, we have

$$\begin{aligned} \xi_{l+1} &= \xi_l + \alpha_l d_l\,, \qquad d_l = P_{I_{l-1}} d_{l-1} = P_{I_{l-1}} \xi_0 \\ \xi_l &= \xi_{l-1} + \alpha_{l-1} d_{l-1} = \xi_0 + \sum_{i=0}^{l-1} \alpha_i d_i = \xi_0 + \sum_{i=0}^{l-1} \alpha_i P_{I_i} \xi_0, \end{aligned}$$

and can conclude that

$$\begin{aligned} \xi_{l+1}{}^T \xi_{l+1} &= (\xi_l + \alpha_l d_l)^T (\xi_l + \alpha_l d_l) = \xi_l{}^T \xi_l + 2\alpha_l \xi_l{}^T d_l + \alpha_l^2 d_l{}^T d_l \\ &= \xi_l{}^T \xi_l + 2\alpha_l (\xi_0 + \sum_{i=0}^{l-1} \alpha_i P_{I_i} \xi_0)^T P_{I_{l-1}} \xi_0 + \alpha_l^2 d_l{}^T d_l \end{aligned}$$

[3] We use the Landau-symbol $O(\,.\,)$ in the following way:
$\eta(m,n) = O(\bar{\eta}(m,n))$ means that there exists a constant $\kappa \in \mathbb{R}^+$ with $|\eta(m,n)| \leq \kappa |\bar{\eta}(m,n)|$ and this constant κ depends neither on n nor on m.

$$= \xi_l{}^T \xi_l + 2\alpha_l \xi_0{}^T P_{I_{l-1}} \xi_0 + 2\alpha_l \sum_{i=0}^{l-1} \alpha_i \xi_0{}^T P_{I_i}^T P_{I_{l-1}} \xi_0 + \alpha_l^2 d_l{}^T d_l$$
$$\geq \xi_l{}^T \xi_l ,$$

because $P_{I_{l-1}}$ and $P_{I_i}^T P_{I_{l-1}} = P_{I_{l-1}}$ (as $i \leq l-1$ and therefore $I_i \subseteq I_{l-1}$) are projection matrices and positive semidefinite and hence, $\alpha_l \xi_0{}^T P_{I_{l-1}} \xi_0 \geq 0$ and $\alpha_l \sum_{i=0}^{l-1} \alpha_i \xi_0{}^T P_{I_i}^T P_{I_{l-1}} \xi_0 \geq 0$ as long as $a_i \geq 0$, $i = 0, \ldots, l-1$.

Part 4: If PROCEDURE TEST returns *true*, then the direction d_l satisfies $\underline{A} d_l \leq 0$. So, the ray $\{x \mid x = \xi_l + \lambda d_l, \lambda \geq 0\}$ is contained in $\underline{X}_P$.

Part 5: PROCEDURE TEST returns *false*, if the repeat-loop is completed with $l = n$ or if a direction d_l $(0 \leq l \leq n)$ happens to be 0.
In the first case, we have $\#I(\xi_n) = n$ (n active constraints) and because of nondegeneracy ξ_n has to be a vertex. Using part 3 we get $||\xi_n|| \geq ||\xi_0||$ and the existence of such a vertex is shown.
For the second case, where $d_l = 0$ for some index $l \in \{0, \ldots, n\}$ we distinguish between $l = 0$ and $l > 0$.
If $l = 0$ then we must have started at the point $\xi_0 = 0$. As the polyhedron $\underline{X}_P$ is pointed (because of nondegeneracy and $m \geq n$), obviously there exists a vertex with norm greater than $||\xi_0|| = 0$.
Now, if $l > 0$, we have reached a point ξ_l on a face, where l constraints (those of I_l) are active. As we have $d_l = 0$, we know that $0 = d_l = P_{I_l} d_{l-1} = P_{I_l} d_{l-2} = \ldots = P_{I_l} d_0 = P_{I_l} \xi_0$ and that in the movement from ξ_0 to $\xi_l = \xi_0 + \sum_{i=0}^{l-1} \alpha_i d_i$ every step uses a direction orthogonal to the face reached at ξ_l. Since $\underline{X}_P$ is pointed, every face of $\underline{X}_P$ is pointed too, in particular the face under consideration. Let $\tilde{\xi}$ be a vertex on the face where the constraints in I_l are active and that contains ξ_l. We can move from ξ_l to the vertex $\tilde{\xi}$ straightforward without leaving that face. So, we have $\tilde{\xi} = \xi_l + \tilde{d}$ with a vector $\tilde{d}$ which is orthogonal to $x_0, d_0, \ldots, d_{l-1}$. The Euclidean norm of $\tilde{\xi}$ is $||\tilde{\xi}||^2 = (\xi_l + \tilde{d})^T (\xi_l + \tilde{d}) = \xi_l^T \xi_l + \tilde{d}^T \tilde{d} \geq \xi_l^T \xi_l$, because $\tilde{d}^T \xi_l = \tilde{d}^T (\xi_0 + \sum_{i=0}^{l-1} \alpha_i d_i) = 0$. This proves that there exists a vertex with norm larger than $||\xi_l||$, and $||\xi_l|| \geq ||\xi_0||$ according to part 3. ▌

Remark 2.9. If we invoke PROCEDURE TEST at a point $x \in \underline{X}_P$ with $||x|| > T(\underline{X}_P)$, then PROCEDURE TEST$(x)$ returns *true*.

Invoking this "test"-procedure we can formulate the phase-I algorithm as a method for minimizing the barrier function $\underline{\phi}$ and to check for each iterate whether there exists a vertex with greater norm or not.

ALGORITHM 2 (PHASE-I ALGORITHM).

Input: $x \in Int\,\underline{X}_P$
1. $k := 0$; $x_k := x$; $\{$*initialization*$\}$
2. COMPUTE $\underline{\delta}(x_k)$;
3. COMPUTE TEST(x_k);
4. **while** $\underline{\delta}(x_k) \geq 1$ **and** TEST$(x_k) =$ *false*

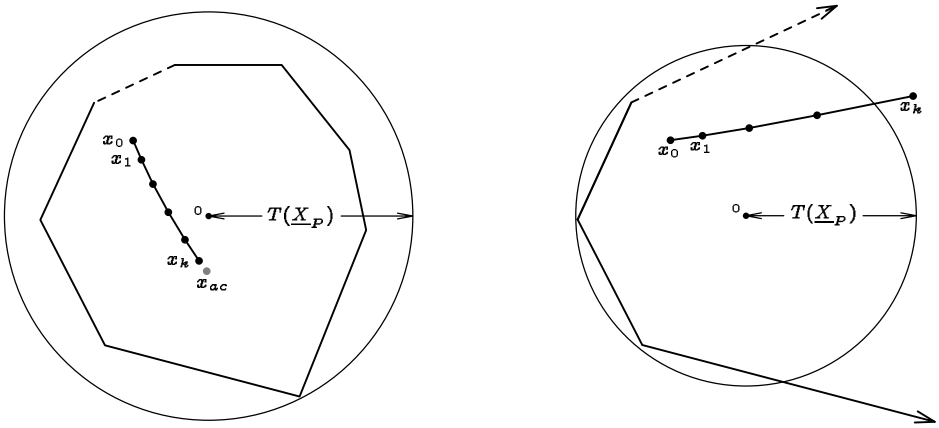

FIG. 1. Illustration of how the PHASE-I ALGORITHM works. In the case of a bounded polyhedron (left hand side) the sequence of iterates x_0, x_1, ..., x_k converges to the analytic center x_{ac} of $\underline{X}_P$ and the algorithm stops at x_k since this point is approximately centered. In the case of an unbounded polyhedron (right hand side) the sequence of iterates does not converge. But finally (at the latest when the iterates leave the region $\{x \mid ||x|| \leq T(\underline{X}_P)\} \cap \underline{X}_P$) the algorithm stops with the detection of unboundedness.

COMPUTE $\underline{p}(x_k) = -\underline{H}(x_k)^{-1}\underline{g}(x_k)$;
$\bar{\alpha}_k := (1 + \underline{\delta}(x_k))^{-1}$;
$x_{k+1} := x_k + \bar{\alpha}_k \underline{p}(x_k)$;
COMPUTE $\underline{\delta}(x_{k+1})$;
COMPUTE TEST(x_{k+1});
$k := k + 1$;
endwhile
Output: x_k with $\underline{\delta}(x_k) < 1$ *or* TEST$(x_k) =$ *true.*

THEOREM 2.2 (Complexity and Correctness of the ALGORITHM 2).
1. After k iterations the value of the barrier function is less than $-0.3k$.
2. After $K_T := 0.3^{-1}|\underline{\Delta}(T)|$ iterations we have $||x_{K_T}|| > T$.
3. The PHASE-I ALGORITHM *terminates after at most $O(|\underline{\Delta}(T(\underline{X}_P))|)$ iterations, i. e. after at most $O(mL)$ iterations in the worst case (using the complexity model of paragraph 2.1.4).*
4. The effort of each iteration is at most $O(mn^2)$.
5. If the PHASE-I ALGORITHM *terminates with* PROCEDURE TEST$(x_k) =$ *true, then the phase-I polyhedron $\underline{X}_P$ is unbounded and the objective function $v^T x$ of problem (P) is unbounded on X_P.*
6. If the PHASE-I ALGORITHM *terminates with $\underline{\delta}(x_k) < 1$, then x_k is an approximate center of $\underline{X}_P$.*

Proof. Part 1: In each of the k iterations $\underline{\delta}(x_k)$ has been greater than 1 (and TEST has been false, otherwise we would have stopped before). So, using Lemma

2.4 (part 1) and the monotonicity of the function $\chi - \ln(1+\chi)$ we get

$$\begin{aligned} \underline{\phi}(x_{k-1}) &\geq \underline{\phi}(x_k) + \underline{\delta}(x_{k-1}) - \ln(1+\underline{\delta}(x_{k-1})) \\ &\geq \underline{\phi}(x_k) + 1 - \ln 2 \geq \underline{\phi}(x_k) + 0.3 \\ \Rightarrow \quad \underline{\phi}(x_k) &\leq \underline{\phi}(x_{k-1}) - 0.3 \\ &\leq \underline{\phi}(x_0) - 0.3k = \underline{\phi}(0) - 0.3k = -0.3k \end{aligned} \tag{21}$$

Part 2: Assume, that $\|x_{K_T}\| \leq T$. Substitution of k in (21) by $K_T = 0.3^{-1}|\underline{\Delta}(T)|$ delivers

$$\underline{\phi}(x_{K_T}) \leq -0.3 \cdot 0.3^{-1}|\underline{\Delta}(T)| = -\sum_{i=1}^{m} \ln(1+\|a_i\|T) - \ln(1+\|v\|T).$$

But this is a contradiction to Lemma 2.5, therefore $\|x_{K_T}\| > T$.

Part 3: Similar to part 2 we have $\phi(x_{K_{T(\underline{X}_P)}}) \leq \underline{\Delta}(T(\underline{X}_P))$ after $K_{T(\underline{X}_P)} = O(|\underline{\Delta}(T(\underline{X}_P))|)$ iterations (assuming that the algorithm has not terminated in one of the prior iterations).
So, the iterate $x_{K_{T(\underline{X}_P)}}$ satisfies $\|x_{K_{T(\underline{X}_P)}}\| > T(\underline{X}_P)$. From the definition of $T(\underline{X}_P)$ in (20) we conclude that there exists no vertex of $\underline{X}_P$ with norm greater than $T(\underline{X}_P)$. Hence, PROCEDURE TEST can not deliver such a vertex. It is forced to stop and to return "true" and so the PHASE-I ALGORITHM has to stop, too.
From complexity theory we know that in the worst case

$$\begin{aligned} |\underline{\Delta}(T(\underline{X}_P))| &= \sum_{i=1}^{m} \ln(1+\|a_i\|T(\underline{X}_P)) + \ln(1+\|v\|T(\underline{X}_P)) \\ &\leq \sum_{i=1}^{m} \ln(1+2^L \cdot 2^L) + \ln(1+2^L \cdot 2^L) = O(mL)\,. \end{aligned}$$

Part 4: The computation of $\underline{H}(x_k)$ can be done in $O(mn^2)$ arithmetic operations and the effort of inverting this matrix is at most $O(n^3)$. All other calculations can be done with less effort.

Part 5: Part 5 follows from Theorem 2.1 and Remark 2.5.

Part 6: If $\underline{\delta}(x_k) < 1$, then x_k is an approximate center of $\underline{X}_P$ by definition. ▮

To obtain a $\underline{\tau}$-approximate center with the PHASE-I ALGORITHM we can modify the algorithm in a way described in Remark 2.10. The approach is based on the subsequent lemma, which shows locally quadratic convergence.

LEMMA 2.6. *If* $\underline{\delta}(x) < 1$, *then* $x_+ := x + \underline{p}(x) \in \mathit{Int}\,\underline{X}_P$ *and* $\underline{\delta}(x_+) \leq \underline{\delta}(x)^2$.

Remark 2.10. To determine a $\underline{\tau}$-approximate center we can change line 4 of the PHASE-I ALGORITHM to:

while $\underline{\delta}(x_k) \geq 0.95$ **and** TEST$(x_k) =$ **false do**.

Then the algorithm will output a point x_k with $\underline{\delta}(x_k) < 0.95$ or TEST$(x_k) =$ true. In the case of $\underline{\delta}(x_k) < 0.95$, we proceed by improving the proximity of the

approximate center according to Lemma 2.6.
This combination of the (modified) algorithm and of improving the proximity of an approximate center terminates after at most $6.65|\underline{\Delta}(T(\underline{X}_P))| + 1.5\ln(20|\ln\underline{\tau}|)$ iterations with a $\underline{\tau}$-approximate center.[4]

Proof. Analogously to the proof of Theorem 2.2 we have in each of the k iterations $\underline{\delta}(x_k) \geq \underline{\tau}$ (and TEST has been false, otherwise we would have stopped before). Using Lemma 2.4 (part 1) we get

$$\underline{\phi}(x_{k-1}) \geq \underline{\phi}(x_k) + \underline{\delta}(x_{k-1}) - \ln(1+\underline{\delta}(x_{k-1})) \geq \underline{\phi}(x_k) + \underline{\tau} - \ln(1+\underline{\tau})\,.$$

Using the Taylor-series representation for $\ln(1+\chi) = \sum_{i=1}^{\infty}(-1)^{i+1}\frac{\chi^i}{i}$ for $\chi \in (-1,1)$ we obtain $\ln(1+\underline{\tau}) \leq \underline{\tau} - \frac{\underline{\tau}^2}{2} + \frac{\underline{\tau}^3}{3}$ and as $\underline{\tau} \in (0,1)$ we conclude

$$\begin{aligned} \underline{\phi}(x_{k-1}) &\geq \underline{\phi}(x_k) + \underline{\tau}^2\left(\tfrac{1}{2} - \tfrac{1}{3}\underline{\tau}\right) \geq \underline{\phi}(x_k) + \tfrac{1}{6}\underline{\tau}^2 \\ \Rightarrow \quad \underline{\phi}(x_k) &\leq \underline{\phi}(x_{k-1}) - \tfrac{1}{6}\underline{\tau}^2 \leq \underline{\phi}(x_0) - \tfrac{1}{6}\underline{\tau}^2 k = -\tfrac{1}{6}\underline{\tau}^2 k\,. \end{aligned}$$

Proceeding as in part 2 and 3 of the proof of Theorem 2.2 we see that after $K_{T(\underline{X}_P)} = \frac{6}{0.95^2}|\underline{\Delta}(T(\underline{X}_P))|$ iterations the value of the barrier function is reduced to a value less than $-|\underline{\Delta}(T(\underline{X}_P))|$. So the modified PHASE-I ALGORITHM has stopped after at most $\frac{6}{0.95^2}|\underline{\Delta}(T(\underline{X}_P))| \leq 6.65|\underline{\Delta}(T(\underline{X}_P))|$ iterations with a 0.95-approximate center or with $\text{TEST}(x_k) = \text{true}$.

At this point we still have to find a $\underline{\tau}$-approximate center. This is done as described in Lemma 2.6, where we use the quadratic convergence. So, starting at x_k with $\underline{\delta}(x_k) \leq 0.95$ we obtain after l applications of Lemma 2.6 a point x_{k+l} with $\underline{\delta}(x_{k+l}) \leq \underline{\delta}(x_k)^{2^l} = 0.95^{2^l}$. To obtain a $\underline{\tau}$-approximate center it is sufficient to find a number l with

$$0.95^{2^l} \leq \underline{\tau} \quad \Leftrightarrow \quad 2^l \geq \frac{\ln\underline{\tau}}{\ln 0.95} \quad \Leftrightarrow \quad l \geq \frac{1}{\ln 2}\ln\left(\frac{\ln\underline{\tau}}{\ln 0.95}\right)$$

As $\frac{1}{\ln 2} \leq 1.5$ and $\frac{\ln\underline{\tau}}{\ln 0.95} = \frac{|\ln\underline{\tau}|}{|\ln 0.95|} \leq 20|\ln\underline{\tau}|$ we find that $l = 1.5\ln(20|\ln\underline{\tau}|)$ will work. So the overall effort of the modified algorithm is at most $6.65|\underline{\Delta}(T(\underline{X}_P))| + 1.5\ln(20|\ln\underline{\tau}|)$. ▮

2.2.3. Construction of a Starting Point for an IPM

Our original goal was to find a starting point for an interior point method for solving (P). But, up to now we only got an approximate center of $\underline{X}_P$.

If we once have an approximate center, i. e. a point $x \in \mathit{Int}\,\underline{X}_P$ with $\underline{\delta}(x) < 1$, we can use the Newton direction $\underline{p}(x)$ to calculate a "better" approximation of the analytic center. Lemma 2.6 implies, that these iterates will converge quadratically to the analytic center.

[4]Note, that this bound on the number of iterations is polynomial in $\langle\underline{\tau}\rangle$, the encoding length of $\underline{\tau}$. This is important if we want to find a $\underline{\tau}$-approximate center of $\underline{X}_P$ and consider $\underline{\tau}$ as a part of the input data. So, we are interested in a worst case bound, that is polynomial in $\langle\underline{\tau}\rangle$.

So, we can assume that we have a point

$$x \in Int\,\underline{X}_P \quad \text{with} \quad \underline{\delta}(x) \le \underline{\tau} < 1 \tag{22}$$

for some fixed $\underline{\tau} \in (0,1)$. The parameter $\underline{\tau}$ has to be chosen according to the conditions, which have to be satisfied by the starting point for the IPM. Here, we want to discuss the typical starting condition when we use a barrier function method (now for solving (P)) as described in [9].

We only employ the IPM for solving (P), if the phase-I algorithm has stopped with $\underline{\delta}(x_k) < 1$ and we therefore know, that $v^T x$ is bounded from above on X_P. Moreover, in this case the barrier function $\phi_P(x,\mu)$ (defined below) achieves its minimum over X_P at a unique point $x(\mu)$. This barrier function is defined by

$$\phi_P(x,\mu) := -\frac{v^T x}{\mu} - \sum_{i=1}^{m} \ln(1 - a_i^T x), \tag{23}$$

where μ is a positive parameter. The minimal point $x(\mu)$ of $\phi_P(x,\mu)$ is uniquely characterized by the KKT conditions:

$$\begin{aligned} &Ax + s = e, \quad s \ge 0 \\ &A^T y = v, \quad y \ge 0 \\ &Sy = \mu e_{(m)} \end{aligned} \tag{24}$$

and we will call the minimal point of $\phi_P(x,\mu)$ the *analytic μ-center.* A measure for the distance of an interior feasible point x to $x(\mu)$ is

$$\delta(x,\mu) := ||p(x,\mu)||_{H(x,\mu)}. \tag{25}$$

Again, $g(x,\mu)$ denotes the gradient, $H(x,\mu)$ the Hessian matrix and $p(x,\mu)$ the Newton direction of $\phi_P(x,\mu)$ at x, i. e.

$$\begin{aligned} &g(x,\mu) = \nabla_x \phi_P(x,\mu) = -\frac{1}{\mu} v + A^T S^{-1} e, \\ &H(x,\mu) = \nabla_x^2 \phi_P(x,\mu) = A^T S^{-2} A \\ &\text{and} \quad p(x,\mu) = -H(x,\mu)^{-1} g(x,\mu). \end{aligned} \tag{26}$$

Again (compare (7)), we can formulate $\delta(x,\mu)$ in different ways:

$$\begin{aligned} \delta(x,\mu) &= ||S^{-1} A p(x,\mu)|| = \min_{y \in \mathbb{R}^m} \left\{ \left\| \frac{Sy}{\mu} - e \right\| \,\middle|\, A^T y = v \right\} \\ &= \left\| \frac{Sy(s,\mu)}{\mu} - e \right\| \end{aligned} \tag{27}$$

$$\text{where} \qquad y(s,\mu) = \mu(S^{-1} e + S^{-2} A p(x,\mu)). \tag{28}$$

And we call a point $x \in Int\, X_P$ an *approximate μ-center*, if we have $\delta(x,\mu) < 1$. Moreover, if we know that $\delta(x,\mu) \le \tau < 1$ we call $x \in Int\, X_P$ a *τ-approximate μ-center.*

Now for starting phase II, we need a point $\bar{x}$ and a suitable parameter $\bar{\mu}$ such that $\delta(\bar{x},\bar{\mu}) < 1$ at least. And — depending on the specific variant of a phase-II algorithm — we are asked to provide a pair $(\bar{x},\bar{\mu})$ such that $\delta(\bar{x},\bar{\mu}) < \tau_{var} < 1$, where τ_{var} is a specific closeness constant required for starting that variant.

The question is whether we can manage to transform our point $x \in Int\, \underline{X}_P$ with $\underline{\delta}(x) < \underline{\tau}$ into such a pair or derive such a pair from x.
Our strategy for constructing such a pair is as follows:
For a given $x \in Int\, \underline{X}_P$ with $\underline{\delta}(x) < \underline{\tau} < 1$ and corresponding slack variables $\underline{s} = \begin{pmatrix} s \\ s_{m+1} \end{pmatrix} = e_{(m+1)} - \underline{A}x$ and the dual point $\underline{y}(\underline{s}) = \begin{pmatrix} y \\ y_{m+1} \end{pmatrix} = \underline{S}^{-1}e_{(m+1)} + \underline{S}^{-2}\underline{A}\underline{p}(x)$ we define

$$\bar{x} := x, \quad \bar{s} := s, \quad \bar{y} := y_{m+1}^{-1}y \quad \text{and} \quad \bar{\mu} := \frac{\bar{s}^T\bar{y}}{m}. \tag{29}$$

LEMMA 2.7. *The construction of $\bar{x}, \bar{s}$ and $\bar{y}$ from (29) out of a point $x \in Int\, \underline{X}_P$ with $\underline{\delta}(x) < 1$ assures that $\bar{x} \in Int\, X_P$, $\bar{y} \in Int\, X_D$.*

The next question is which quality of the distance measure $\delta(x,\mu)$ can be achieved by this construction.

THEOREM 2.3. *Application of the transformation described in (29) to $x \in Int\, \underline{X}_P$ with $\underline{\delta}(x) \le \underline{\tau} < 1$ provides that*

$$\delta(\bar{x},\bar{\mu}) \le \frac{\underline{\tau}(1+\underline{\tau})m}{m - \underline{\tau}(1+\underline{\tau})}.$$

Proof. First, we introduce $\underline{\eta} = \underline{\delta}(x) \le \underline{\tau}$. Then notice, that

$$\|\underline{S}\underline{y}(\underline{s}) - e_{(m+1)}\| = \underline{\eta} < 1 \;\Rightarrow\; \|Sy - e_{(m)}\|^2 + (s_{m+1}y_{m+1} - 1)^2 = \underline{\eta}^2 \tag{30}$$

$$\Rightarrow\; 1 - \underline{\eta} \le s_{m+1}y_{m+1} \le 1 + \underline{\eta} \tag{31}$$

and (compare (13) – (18) for relations between $\underline{y}(\underline{s})$, $\underline{p}(x)$ and $\underline{\delta}(x) = \underline{\eta}$)

$$\begin{aligned} e_{(m+1)}{}^T\underline{S}\underline{y}(\underline{s}) &= e_{(m+1)}{}^T\underline{S}\left(\underline{S}^{-1}e_{(m+1)} + \underline{S}^{-2}\underline{A}\underline{p}(x)\right) \\ &= m + 1 - \underline{p}(x)^T\underline{H}(x)\underline{p}(x) = m + 1 - \underline{\eta}^2 \\ \Rightarrow \qquad & s^Ty + s_{m+1}y_{m+1} = m + 1 - \underline{\eta}^2. \end{aligned} \tag{32}$$

Defining

$$\mu := \frac{s^Ty}{m} \tag{33}$$

and using (31) and (32) we conclude

$$\mu \;=\; \frac{m+1-\underline{\eta}^2-s_{m+1}y_{m+1}}{m}\begin{cases}\leq & \frac{m+1-\underline{\eta}^2-1+\underline{\eta}}{m}=1+\frac{\underline{\eta}(1-\underline{\eta})}{m}\,,\\ \geq & \frac{m+1-\underline{\eta}^2-1-\underline{\eta}}{m}=1-\frac{\underline{\eta}(1+\underline{\eta})}{m}\,,\end{cases} \tag{34}$$

$$\Rightarrow\;\frac{\underline{\eta}(1-\underline{\eta})}{m}\leq 1-\mu\leq\frac{\underline{\eta}(1+\underline{\eta})}{m}\,. \tag{35}$$

Returning to the figures $\bar{s},\bar{y}$ and $\delta(\bar{x},\bar{\mu})$ we obtain with (27):

$$\bar{S}\bar{y}=\frac{1}{y_{m+1}}Sy \quad\text{and}\quad \bar{\mu}=\frac{s^Ty}{y_{m+1}m} \tag{36}$$

$$\Rightarrow\;\delta(\bar{x},\bar{\mu})=\left\|\frac{\bar{S}\bar{y}(\bar{s},\bar{\mu})}{\bar{\mu}}-e_{(m)}\right\|$$

$$\leq\left\|\frac{\bar{S}\bar{y}}{\bar{\mu}}-e_{(m)}\right\|=\left\|\frac{Sy}{\mu}-e_{(m)}\right\|=\frac{1}{\mu}\|Sy-\mu e_{(m)}\|\,.$$

Using (30), (32), (33) and (35) we get

$$\begin{aligned}\|Sy-\mu e_{(m)}\|^2 &= \|Sy-e_{(m)}+(1-\mu)e_{(m)}\|^2\\ &= \underline{\eta}^2-(s_{m+1}y_{m+1}-1)^2+2(1-\mu)(\mu m-m)+(1-\mu)^2m\\ &= \underline{\eta}^2-\underline{\eta}^4+2(1-\mu)m\underline{\eta}^2-(1-\mu)^2m(m+1)\\ &\leq \underline{\eta}^2-\underline{\eta}^4+2m\underline{\eta}^2\frac{\underline{\eta}(1+\underline{\eta})}{m}-(1-\mu)^2m(m+1)\\ &\leq \underline{\eta}^2(1+\underline{\eta})^2\,.\end{aligned}$$

On the other hand it is clear from (34) that $\mu\geq 1-\frac{\underline{\eta}(1+\underline{\eta})}{m}$, so

$$\delta(\bar{x},\bar{\mu})\leq\frac{1}{\mu}\|Sy-\mu e_{(m)}\|\leq\frac{\underline{\eta}(1+\underline{\eta})}{1-\frac{\underline{\eta}(1+\underline{\eta})}{m}}=\frac{\underline{\eta}(1+\underline{\eta})m}{m-\underline{\eta}(1+\underline{\eta})}$$

This upper bound for $\delta(\bar{x},\bar{\mu})$ is strictly increasing in $\underline{\eta}$ and therefore we have $\delta(\bar{x},\bar{\mu})\leq\dfrac{\underline{\tau}(1+\underline{\tau})m}{m-\underline{\tau}(1+\underline{\tau})}$ for $\underline{\tau}\geq\underline{\eta}$. ▮

Remark 2.11. The upper bound in Theorem 2.3 is strictly increasing in $\underline{\tau}$. It starts at 0 for $\underline{\tau}=0$ and grows to $\frac{2m}{m-2}$ for $\underline{\tau}=1$.

To obtain a point $\bar{x}$ with $\delta(\bar{x},\bar{\mu})<1$ choose $\underline{\tau}$ such that $\frac{\underline{\tau}(1+\underline{\tau})m}{m-\underline{\tau}(1+\underline{\tau})}<1$ resp. $\underline{\tau}(1+\underline{\tau})<\frac{m}{m+1}$. This is satisfied for $\underline{\tau}\leq 0.36$ for all $m\geq 1$.
And if we want to achieve a point $\bar{x}$ with $\delta(\bar{x},\bar{\mu})\leq\tau_{var}$ then it is sufficient to satisfy the condition $\underline{\tau}(1+\underline{\tau})<\frac{m\tau_{var}}{m+\tau_{var}}$, which is possible if $\underline{\tau}\leq\frac{1}{2}\left(-1+\sqrt{1+4\frac{m\tau_{var}}{m+\tau_{var}}}\right)$. So for every variant specific requirement τ_{var} we have a positive region for $\underline{\tau}$. (Another way to achieve a point with $\delta(x,\mu)\leq\tau_{var}$ is to obtain first a point $\bar{x}$ with $\delta(\bar{x},\bar{\mu})\leq 0.95$, and then use the Newton direction $p(x,\mu)$ to calculate a

better approximation of the analytic μ-center. This can be done analogously to Lemma 2.6.)

For the complexity of a forthcoming phase-II algorithm it may be helpful to know something about the initial barrier parameter $\bar{\mu}$.

Remark 2.12. For $\bar{x}$, $\bar{s}$ and $\bar{\mu}$ defined as in (29) we know that

$$\bar{\mu} \leq \left(\frac{1}{1-\underline{\eta}} + \frac{\underline{\eta}}{m}\right)(1 + v^T\bar{x}) \leq O(2^L)\,,$$

where $\underline{\eta} = \underline{\delta}(\bar{x})$.

If we assume $||x_{opt}|| \leq \frac{1}{q}$ then we can give the following upper bound

$$\bar{\mu} \leq \left(\frac{1}{1-\underline{\eta}} + \frac{\underline{\eta}}{m}\right)\left(1 + \frac{1}{q}\right).$$

Proof. Using (36) and (33) we obtain $\bar{\mu} = \frac{s^T y}{y_{m+1}m} = \frac{\mu}{y_{m+1}}$. From the inequalities in (31) and (35) and because of $s_{m+1} = 1 + v^T x = 1 + v^T\bar{x}$ it follows that $\bar{\mu} \leq \frac{1 + \frac{\underline{\eta}(1-\underline{\eta})}{m}}{y_{m+1}} \leq \left(1 + \frac{\underline{\eta}(1-\underline{\eta})}{m}\right)\frac{1 + v^T\bar{x}}{1-\underline{\eta}}$.

Since $\underline{\eta} \in (0,1)$ is fixed and $1 + v^T\bar{x} \leq O(2^L)$ the barrier function $\bar{\mu}$ is at most $O(2^L)$.

Assuming $||x_{opt}|| \leq \frac{1}{q}$ we argue analogously, but we use $s_{m+1} = 1 + v^T x \leq 1 + v^T x_{opt} \leq 1 + \frac{1}{q}$. ▮

2.3. Probabilistic Analysis

2.3.1. The Rotation-Symmetry-Model

We have seen that the upper bound on the number of iterations of the phase-I algorithm depends on the encoding length L. But practical experience suggests that this upper bound overestimates the usual number of iterations. To explain this effect in a convincing theoretical way, we will carry out a probabilistic analysis.

For that purpose one has to make assumptions on the distribution of the data of (P); and one has to evaluate an expected value of iterations (or corresponding moments) for the solution process under our distribution. As mentioned in the introduction, we will base our probabilistic analysis on the Rotation-Symmetry-Model (RSM). It demands that

$$\begin{array}{l} a_1, \ldots, a_m, v \textit{ are distributed on } \mathbb{R}^n \setminus \{0\} \\ \textit{identically, independently and symmetrically under rotations.} \end{array} \tag{37}$$

We specialize this model here to the uniform distribution on the ($(n-1)$-dimensional) unit sphere ω_n in $\mathbb{R}^n$ and refer to it as *uni*-RSM.

$$\begin{array}{l} a_1, \ldots, a_m, v \textit{ are distributed} \\ \textit{identically, independently and uniformly on } \omega_n. \end{array} \tag{38}$$

Note that, in fact, this model gives to nondegeneracy (as in (1)) the probability 1 and makes degeneracy ignorable in our average-case analysis.

2.3.2. The Average Number of Iterations (Part 1)

With these probabilistic settings we can consider $T(\underline{X}_P)$ and $\underline{\Lambda}(T(\underline{X}_P))$ as random variables depending on $a_1, \ldots, a_m, v$. From Theorem 2.2 (part 3) we know that $|\underline{\Lambda}(T(\underline{X}_P))|$ is essentially the problem-specific number of iterations of the PHASE-I ALGORITHM and we can get the average number of iterations of this algorithm by calculating $E[|\underline{\Lambda}(T(\underline{X}_P))|]$.
Recall the definition of $\underline{\Lambda}(T(\underline{X}_P)) = -\sum_{i=1}^{m} \ln(1 + ||a_i||T) - \ln(1 + ||v||T)$ as in (19). Since $||v|| = ||a_i|| = 1$ for $i = 1, \ldots, m$ in the *uni*-RSM, it follows that

$$|\underline{\Lambda}(T(\underline{X}_P))| = (m+1)\ln(1 + T(\underline{X}_P)) =: \widetilde{\Lambda}(T(\underline{X}_P))\,, \tag{39}$$

and for the expectation values we have[5]

$$E[|\underline{\Lambda}(T(\underline{X}_P))|] = E[\widetilde{\Lambda}(T(\underline{X}_P))]\,. \tag{40}$$

The expectation value is evaluated with respect to the distribution of $T(\underline{X}_P)$. So, we are interested in the distribution function of $T(\underline{X}_P)$, which will be denoted by F_T. But, an exact derivation of F_T is overcomplicated and therefore we will derive and use an approximate distribution function $\widehat{F}$ with the following property:

$$\widehat{F}(t) \le F_T(t) \quad \forall\, t \in (0, \infty). \tag{41}$$

A distribution $\widehat{F}$ with this feature will put more weight on the large values of t than the exact distribution F_T and this variation of weights leads to

$$E_{F_T}[\widetilde{\Lambda}(T(\underline{X}_P))] \le E_{\widehat{F}}[\widetilde{\Lambda}(T(\underline{X}_P))], \tag{42}$$

where $E_{F_T}[\,.\,]$ is the expectation value with respect to the exact distribution function F_T and $E_{\widehat{F}}[\,.\,]$ is the expectation value with respect to the approximate distribution function $\widehat{F}$.
The next section deals with these distribution functions.

2.3.3. An Approximate Distribution Function

First, we look at the exact distribution function $F_T(t)$, i. e. it is the probability that $T(\underline{X}_P)$ is less than t.

$$\begin{aligned} F_T(t) &= P(T(\underline{X}_P) \le t) = P(||x|| \le t\,\forall \text{ vertices } x \text{ of } \underline{X}_P) \\ &= 1 - P(\text{there exists a vertex } x \text{ of } \underline{X}_P \text{ with } ||x|| > t). \end{aligned} \tag{43}$$

[5]The figure $\widetilde{\Lambda}(T(\underline{X}_P) := (m+1)\ln(1 + T(\underline{X}_P))$ is an upper bound on $\underline{\Lambda}(T(\underline{X}_P))$ for all distributions (of v, a_i, $i = 1, \ldots, m$) with Ω_n (unit ball in $\mathbb{R}^n$) as bounded support, for the uniform distribution both figures are equal.

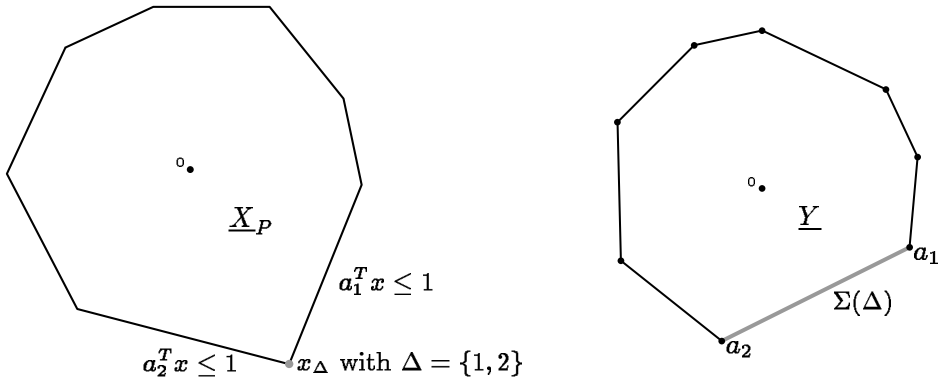

FIG. 2. Primal polyhedron $\underline{X}_P$ and corresponding dual polyhedron $\underline{Y}$. The vertex x_Δ of $\underline{X}_P$ has a dual counterpart, the simplex $\Sigma(\Delta)$.

If we find an upper bound on P(there exists a vertex x of $\underline{X}_P$ with $||x|| > t$), we get a lower bound on $F_T(t)$ and can use this lower bound to define the approximate distribution function $\widehat{F}$.

For the evaluation of the probability we use the polar interpretation in the dual space as in [6]. The advantage is that the random events can be explained directly by use of the random input vectors $a_1, \dots, a_m$ and $a_{m+1} := -v$. (The use of a_{m+1} instead of $-v$ will simplify the forthcoming notation and formulas.)

We start by introducing the polar polyhedron. Let *conv*$(\dots)$ denote the convex hull. Then the polar polyhedron corresponding to $\underline{X}_P$ is defined by

$$\underline{Y} := conv(0, a_1, \dots, a_m, a_{m+1})\,. \tag{44}$$

Each vertex x of $\underline{X}_P$ is — because of nondegeneracy — the unique solution of a system

$$\begin{array}{l} a_{\Delta^1}^T x = 1, \dots, a_{\Delta^n}^T x = 1 \text{ with } \Delta = \{\Delta^1, \dots, \Delta^n\} \subseteq \{1, \dots, m+1\} \\ \text{and it satisfies } a_i^T x \leq 1 \text{ for all } i \in \{1, \dots, m+1\}\,. \end{array} \tag{45}$$

Δ is the index set of those constraints, which are active at the vertex x. To distinguish between different vertices, we will label a vertex with its index set and denote it by x_Δ.
For the polar polyhedron $\underline{Y}$ we define simplices $\Sigma(\Delta)$ as

$$\Sigma(\Delta) := conv(a_i, i \in \Delta) \tag{46}$$

and we call $\Sigma(\Delta)$ a *boundary simplex* if $\Sigma(\Delta) \subseteq \partial\underline{Y}$, i. e. if $\Sigma(\Delta)$ is a facet of $\underline{Y}$. From [6] we know

LEMMA 2.8 ([6], Lemma 1.7). *x_Δ is a vertex of $\underline{X}_P$ if and only if $\Sigma(\Delta)$ is a boundary simplex of $\underline{Y}$.*

For further considerations we will use the following notations.
Given n linearly independent points $x_1, \dots, x_n \in \mathbb{R}^n$ we denote the hyperplane

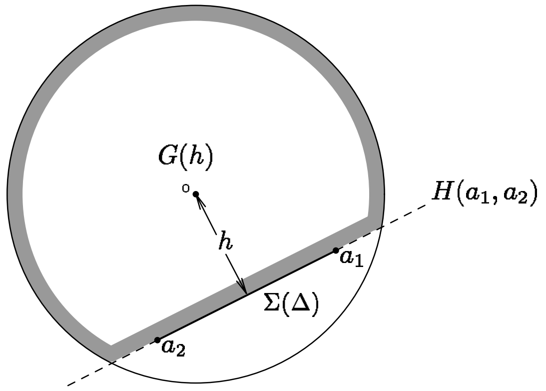

FIG. 3. The sphere illustrates the support of the distribution and $G(h)$ is the probability that all points are lying in the halfspace $H^-(a_1, a_2)$ which is bounded by the hyperplane $H(a_1, a_2)$.

through $x_1, \ldots, x_n$ by $H(x_1, \ldots, x_n)$ and the distance between $H(x_1, \ldots, x_n)$ and the origin by $h(x_1, \ldots, x_n)$. $H^-(x_1, \ldots, x_n)$ denotes the halfspace which is bounded by $H(x_1, \ldots, x_n)$ and contains the origin. Using these notations we have the following relation:

$$\Sigma(\Delta) \text{ is a boundary simplex of } \underline{Y} \quad \Leftrightarrow \quad a_j \in H^-(a_i, i \in \Delta) \quad \forall j \in \{1, \ldots, m+1\}. \tag{47}$$

Moreover, we know that x_Δ can be seen as the normal vector to the hyperplane $H(a_i, i \in \Delta)$ and $\|x_\Delta\| = \frac{1}{h(a_i, i \in \Delta)}$.

Under our probabilistic model, the Rotation-Symmetry-Model, the probability that a point x is lying in the halfspace $H^-(a_i, i \in \Delta)$ can be described by a marginal distribution function $G : [-1, 1] \to [0, 1]$ which depends on the height $h(a_i, i \in \Delta)$ (compare [6]). $G(h)$ is defined as $P(x^n \leq h)$, resp. it is the marginal distribution function along one (the last) coordinate. In our special case, the *uni*-RSM, this reads

$$P(x \in H^-(a_i, i \in \Delta)) =: G(h(a_i, i \in \Delta)) = 1 - \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \int_{h(a_i,\, i \in \Delta)}^{1} (1 - \sigma^2)^{\frac{n-3}{2}} d\sigma\,, \tag{48}$$

where ω_i denotes the ($(i-1)$-dim.) unit sphere in $\mathbb{R}^i$ and λ_i is the Lebesgue-measure of dimension i.

In order to develop an integral formula for the probability

$$P(\text{there exists a vertex } x \text{ of } \underline{X}_P \text{ with } \|x\| > t) \tag{49}$$

we will write $F(a_i)$ for the distribution function of the vectors $a_i, i = 1, \ldots, m+1$ (these vectors are distributed identically) and we will use indicator functions $I_{\{\,.\,\}}$

to formulate the fact that an event $\{\,.\,\}$ becomes true. For further considerations we remark, that all vertices x satisfy $||x|| \geq 1$ and therefore we can rely on the fact that t is greater than 1 and that

$$F_T(t) = P(T(\underline{X}_P) \leq t) = 0 \quad \forall\, t \leq 1\,. \tag{50}$$

So, we can formulate resp. approximate the probability in (49) as follows:

$$\begin{aligned}
&P(\text{there exists a vertex } x \text{ with } ||x|| > t) = \\
&= \int\limits_{\mathbb{R}^n}^{(m+1)} I_{\{\text{there exists a vertex } x \text{ with } ||x|| > t\}}(t, a_1, \dots, a_{m+1}) \\
&\qquad dF(a_1) \cdots dF(a_{m+1}) \\
&\leq \sum_{\Delta \subseteq \{1,\dots,m+1\}} \int\limits_{\mathbb{R}^n}^{(m+1)} I_{\{x_\Delta \text{ is a vertex with } ||x_\Delta|| > t\}}(t, a_1, \dots, a_{m+1}) \\
&\qquad dF(a_1) \cdots dF(a_{m+1}) \\
&\qquad \text{w.l.o.g. we can assume } \Delta = \{1, \dots, n\} \\
&= \binom{m+1}{n} \int\limits_{\mathbb{R}^n}^{(m+1)} I_{\{x_\Delta \text{ is a vertex}\}}(a_1, \dots, a_{m+1}) \cdot \\
&\qquad I_{\{||x_\Delta|| > t\}}(t, a_1, \dots, a_n) dF(a_1) \cdots dF(a_{m+1}) \\
&= \binom{m+1}{n} \int\limits_{\mathbb{R}^n}^{(n)} I_{\{||x_\Delta|| > t\}}(t, a_1, \dots, a_n) \cdot \\
&\qquad \int\limits_{\mathbb{R}^n}^{(m-n+1)} I_{\{x_\Delta \text{ is a vertex}\}}(a_1, \dots, a_{m+1}) dF(a_{n+1}) \cdots dF(a_{m+1}) \\
&\qquad dF(a_1) \cdots dF(a_n)\,.
\end{aligned} \tag{51}$$

We have (compare (48))

$$\begin{aligned}
&\int\limits_{\mathbb{R}^n}^{(m-n+1)} I_{\{x_\Delta \text{ is a vertex}\}}(a_1, \dots, a_n) dF(a_{n+1}) \cdots dF(a_{m+1}) \\
&= G(h(a_1, \dots, a_n))^{m-n+1}\,.
\end{aligned}$$

Hence,

$$\begin{aligned}
&P(\text{there exists a vertex } x \text{ with } ||x|| > t) \leq \\
&\leq \binom{m+1}{n} \int\limits_{\mathbb{R}^n}^{(n)} G(h(a_1, \dots, a_n))^{m-n+1} I_{\{||x_\Delta|| > t\}}(t, a_1, \dots, a_n) \\
&\qquad dF(a_1) \cdots dF(a_n)\,.
\end{aligned} \tag{52}$$

Next, we apply a special transformation of coordinates to make (52) evaluable. W.l.o.g. we may assume that F has a density function f.[6] We are going to replace the vectors $a_1,\dots,a_n$ by vectors $b_1,\dots,b_n$ with $b_i^n = h$ for some h with $0<h<1$ and for all $i=1,\dots,n$. This transformation is described in [6], chapter 2.3. The determinant of the Jacobian of this transformation can be taken into account by $\lambda_{n-1}(\omega_n)|\operatorname{Det}(B)|$, where $\lambda_{n-1}(\omega_n)$ denotes the Lebesgue measure of ω_n and

$$B = \begin{pmatrix} b_1^1 & \cdots & b_1^{n-1} & 1 \\ \vdots & & \vdots & \vdots \\ b_n^1 & \cdots & b_n^{n-1} & 1 \end{pmatrix}. \tag{53}$$

Furthermore, let $\bar{b}_i := (b_i^1,\dots,b_i^{n-1})$ for $i=1,\dots,n$ and let $f(b_i)$ denote the density of b_i.
Now, $\|x_\Delta\| = 1/h(a_1,\dots,a_n)$ can be formulated as $\|x_\Delta\| = 1/h > 1$, because by the transformation h is the distance between the origin and the hyperplane $H(b_1,\dots,b_n)$. Hence

$$\begin{aligned} &P(\text{there exists a vertex } x \text{ with } \|x\|>t) \\ &\le \binom{m+1}{n}\lambda_{n-1}(\omega_n)\int_0^1 \int_{\mathbb{R}^{n-1}}^{(n)} G(h)^{m-n+1} I_{\{\frac{1}{h}>t\}}(t,h) \\ &\qquad |\operatorname{Det}(B)| f(b_1)\cdots f(b_n)d\bar{b}_1\dots d\bar{b}_n dh \\ &= \binom{m+1}{n}\lambda_{n-1}(\omega_n)\cdot \\ &\qquad \int_0^{1/t} G(h)^{m-n+1}\int_{\mathbb{R}^{n-1}}^{(n)} |\operatorname{Det}(B)| f(b_1)\cdots f(b_n)d\bar{b}_1\dots d\bar{b}_n dh. \end{aligned} \tag{54}$$

For the uniform distribution on ω_n the inner integral can be evaluated by using the function $g(h)$, which is the density corresponding to $G(h)$, i. e.

$$g(h) = \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)}(1-h^2)^{\frac{n-3}{2}} \tag{55}$$

and therefore (compare [10])

$$\begin{aligned} &\int_{\mathbb{R}^{n-1}}^{(n)} |\operatorname{Det}(B)| f(b_1)\cdots f(b_n)d\bar{b}_1\dots d\bar{b}_n = \\ &= \frac{(n!)\lambda_{n-1}(\Omega_{n-1})\Gamma(\frac{n}{2}-\frac{1}{2})\Gamma(n(\frac{n}{2}-1))}{(2\pi)^{n-1}\Gamma(\frac{n}{2}-1)\Gamma(n(\frac{n}{2}-1)+\frac{1}{2})}\, g(h)(1-h^2)^{(n-1)(\frac{n}{2}-1)}, \end{aligned} \tag{56}$$

[6] It is not necessary at this point to assume the existence of a density function. All calculations can explicitly be done for the uniform distribution. But the arguments given above easily permit to generalize the results to other (families of) distributions.

where $\lambda_{n-1}(\Omega_{n-1})$ denotes the Lebesgue measure of the unit ball in $\mathbb{R}^{n-1}$. Insertion of (56) into (54) delivers

$$\begin{aligned}
&P(\text{there exists a vertex } x \text{ with } \|x\| > t)\\
\leq\ & \binom{m+1}{n} \lambda_{n-1}(\omega_n) \frac{(n!)\lambda_{n-1}(\Omega_{n-1})}{(2\pi)^{n-1}} \frac{\Gamma(\frac{n}{2}-\frac{1}{2})\Gamma(n(\frac{n}{2}-1))}{\Gamma(\frac{n}{2}-1)\Gamma(n(\frac{n}{2}-1)+\frac{1}{2})}\\
& \int_0^{1/t} G(h)^{m-n+1}\, g(h)(1-h^2)^{(n-1)(\frac{n}{2}-1)} dh\\
\leq\ & \binom{m+1}{n} \frac{2n}{\sqrt{n-1}} \int_0^{1/t} G(h)^{m-n+1}\, g(h)(1-h^2)^{(n-1)(\frac{n}{2}-1)} dh \qquad (57)\\
\leq\ & 0.8n \binom{m+1}{n} \int_0^{1/t} G(h)^{m-n+1}\, (1-h^2)^{n(\frac{n}{2}-1)-2} dh \qquad (58)\\
\leq\ & 0.8n \binom{m+1}{n} \cdot \frac{1}{t} \cdot \max_{h\in[0,\frac{1}{t}]} \{G(h)^{m-n+1}(1-h^2)^{n(\frac{n}{2}-1)-2}\}\\
\leq\ & 0.8n \binom{m+1}{n} \cdot \frac{1}{t} \cdot G\left(\frac{1}{t}\right)^{m-n+1}. \qquad (59)
\end{aligned}$$

The inequality in (57) is valid because

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

and $\quad \dfrac{\Gamma(\frac{n}{2}-\frac{1}{2})\Gamma(n(\frac{n}{2}-1))}{\Gamma(\frac{n}{2}-1)\Gamma(n(\frac{n}{2}-1)+\frac{1}{2})} \leq \dfrac{(\frac{n}{2}-1)^{\frac{1}{2}}}{(n(\frac{n}{2}-1)-\frac{1}{2})^{\frac{1}{2}}} \leq \dfrac{1}{\sqrt{n-1}}.$

The inequality (58) follows from

$$g(h) = \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)}(1-h^2)^{\frac{n-3}{2}} \leq \sqrt{\tfrac{n-1}{2\pi}}\,(1-h^2)^{\frac{n-3}{2}}.$$

Now, we have approximated $P(\text{there exists a vertex } x \text{ with } \|x\| > t)$ as follows:

$$P(\text{there exists a vertex } x \text{ with } \|x\| > t) \leq 0.8n \tbinom{m+1}{n} \tfrac{1}{t} G(\tfrac{1}{t})^{m-n+1}.$$

If we look at the term $0.8n\binom{m+1}{n}\frac{1}{t}\cdot G(\frac{1}{t})^{m-n+1}$ we see that for $t\to\infty$ this term tends to 0 and for $t \to 1$ it tends to $0.8n\binom{m+1}{n} > 1$. Since the term is monotonously decreasing with t, it is guaranteed that there exists a $\tilde{t} > 1$ satisfying

$$0.8n \binom{m+1}{n} \cdot \frac{1}{\tilde{t}} \cdot G\left(\frac{1}{\tilde{t}}\right)^{m-n+1} = 1. \qquad (60)$$

Using $\tilde{t}$ we can specialize the upper bound on the probability to:

$$\begin{aligned}
&P(\text{there exists a vertex } x \text{ with } \|x\| > t)\\
\leq\ & \begin{cases} 1 & \text{for all } 1 \leq t \leq \tilde{t}, \\ 0.8n \binom{m+1}{n} \frac{1}{t} G(1/\tilde{t})^{m-n+1} & \text{for all } t \geq \tilde{t}, \end{cases}
\end{aligned}$$

where we have replaced $G(\frac{1}{t})$ by its upper bound $G(\frac{1}{\tilde{t}})$ for simplification.

2.3.4. The Average Number of Iterations (Part 2)

The upper bound just developed can be used to define a lower bound on the distribution function $F_T(t)$ via (43) and to define an approximate distribution function $\widehat{F}$ in the same way

$$\begin{aligned} F_T(t) &= 1 - P(\text{there exists a vertex } x \text{ of } \underline{X}_P \text{ with } \|x\| > t) \qquad (61)\\ &\geq \begin{cases} 0 & \text{for all } 1 \leq t \leq \tilde{t}, \\ 1 - 0.8n\binom{m+1}{n}\frac{1}{t}G(1/\tilde{t})^{m-n+1} & \text{for all } t \geq \tilde{t}. \end{cases} \\ &=: \widehat{F}(t). \qquad (62) \end{aligned}$$

$\widehat{F}$ satisfies $\widehat{F}(t) = 0$ for $t \leq \tilde{t}$ and $\widehat{F}(t) \geq 0$ for all t, $\widehat{F}$ is monotonically increasing and $\lim_{t\to\infty} \widehat{F}(t) = 1$.
Now, as we have an explicit formula for $\widehat{F}$, it becomes easy to calculate $E_{\widehat{F}}[\widetilde{\Lambda}(T(\underline{X}_P))]$ and this expectation value provides an upper bound on the average number of iterations of the PHASE-I ALGORITHM (by using (40) and (42).

$$\begin{aligned} &E[|\underline{\Lambda}(T(\underline{X}_P))|] = E[\widetilde{\Lambda}(T(\underline{X}_P))] \leq E_{\widehat{F}}[\widetilde{\Lambda}(T(\underline{X}_P))] \\ &= (m+1)E_{\widehat{F}}[\ln(1+T(\underline{X}_P))] = (m+1)\int_1^\infty \ln(1+t)d\widehat{F}(t) \\ &= 0.8n(m+1)\binom{m+1}{n}G\left(\frac{1}{\tilde{t}}\right)^{m-n+1}\int_{\tilde{t}}^\infty \ln(1+t)\frac{1}{t^2}dt \\ &= 0.8n(m+1)\binom{m+1}{n}G\left(\frac{1}{\tilde{t}}\right)^{m-n+1}\left(\frac{\ln(1+\tilde{t})}{\tilde{t}} + \ln\left(\frac{1+\tilde{t}}{\tilde{t}}\right)\right) \\ &= (m+1)\left(\ln(1+\tilde{t}) + \tilde{t}\ln\left(\frac{1+\tilde{t}}{\tilde{t}}\right)\right) \qquad \text{(compare (60))} \\ &\leq (m+1)\left(\ln(1+\tilde{t}) + \tilde{t}\ln 2\right). \qquad (63) \end{aligned}$$

In order to get an upper bound on the average number of iterations, which depends on the parameters m, n only, we should find the explicit value of $\tilde{t}$, resp. an upper bound for it.
Remembering the implicit definition of $\tilde{t}$ in (60) we find that any t with

$$0.8n\binom{m+1}{n}\cdot\frac{1}{t}\cdot G\left(\frac{1}{t}\right)^{m-n+1} \leq 1 \qquad (64)$$

delivers an upper bound on $\tilde{t}$. So, we have to consider the left-hand side in (64) and we start by having a look at the function $G\left(\frac{1}{t}\right)$. According to the Mean Value Theorem there exists a $\tilde{h} \in [0, h]$ such that

$$G(h) = G(0) + h\,g(\tilde{h}) \leq \tfrac{1}{2} + h\tfrac{\sqrt{n-1}}{\sqrt{2\pi}}(1-\tilde{h}^2)^{\frac{n-3}{2}} \leq \tfrac{1}{2} + h\tfrac{\sqrt{n-1}}{\sqrt{2\pi}}$$

$$\Rightarrow \quad G\left(\tfrac{1}{t}\right) \leq \tfrac{1}{2} + \tfrac{1}{t}\tfrac{\sqrt{n-1}}{\sqrt{2\pi}}\,. \tag{65}$$

A simple upper bound on $0.8n\binom{m+1}{n}$ can be derived with Stirling's formula as follows

$$0.8n\binom{m+1}{n} \leq 0.8\,\frac{(m+1)^n}{(n-1)!} \leq 0.2\,(m+1)^n\left(\tfrac{e}{n-1}\right)^{n-\frac{1}{2}}. \tag{66}$$

Now, we use these approximations to bound the left-hand side of (64) and we should seek for a value of t which is great enough to assure that

$$0.2\,(m+1)^n\left(\tfrac{e}{n-1}\right)^{n-\frac{1}{2}}\tfrac{1}{t}\left(\tfrac{1}{2}+\tfrac{1}{t}\tfrac{\sqrt{n-1}}{\sqrt{2\pi}}\right)^{m+1-n} \leq 1 \tag{67}$$

$$\Leftrightarrow \quad \left(0.2\,(m+1)^n\left(\tfrac{e}{n-1}\right)^{n-\frac{1}{2}}\tfrac{1}{t}\right)^{\frac{1}{m+1-n}}\left(\tfrac{1}{2}+\tfrac{1}{t}\tfrac{\sqrt{n-1}}{\sqrt{2\pi}}\right) \leq 1\,. \tag{68}$$

It is clear that $\tilde{t} \leq t$, that means the value of t would bound $\tilde{t}$ from above.

Let us, for simplicity of the representation, substitute $t := \vartheta\frac{\sqrt{n-1}}{\sqrt{2\pi}}$. Now our inequality reads as follows:

$$\left(0.2\,(m+1)^n\left(\tfrac{e}{n-1}\right)^{n-\frac{1}{2}}\tfrac{\sqrt{2\pi}}{\vartheta\sqrt{n-1}}\right)^{\frac{1}{m+1-n}}\left(\tfrac{1}{2}+\tfrac{1}{\vartheta}\right) \leq 1\,.$$

Since there is no chance to satisfy the above inequality for all m, n with values $\vartheta \leq 2$, we restrict our considerations to values $\vartheta > 2$. But still we want to find a ϑ as small as possible.
The left side has as an upper bound

$$\left(0.51\,(m+1)^n\frac{e^{n-\frac{1}{2}}}{(n-1)^n}\right)^{\frac{1}{m+1-n}}\frac{1}{\vartheta^{\frac{1}{m+1-n}}}\left(\frac{\vartheta+2}{2\vartheta}\right)$$

and still it would be sufficient to find a ϑ making this expression less than 1. But this would be satisfied if and only if the following inequality holds.

$$\left(0.51\,(m+1)^n\frac{e^{n-\frac{1}{2}}}{(n-1)^n}\right)^{\frac{1}{m+1-n}} \leq \frac{2\vartheta}{\vartheta+2}\cdot\vartheta^{\frac{1}{m+1-n}} \tag{69}$$

$$\Leftrightarrow \quad \left(0.51\,\frac{e^{n-\frac{1}{2}}}{(n-1)^n}\right)^{\frac{1}{m+1-n}}\cdot(m+1)^{\frac{n}{m+1-n}} \leq \frac{2\vartheta}{\vartheta+2}\cdot\vartheta^{\frac{1}{m+1-n}}$$

If n is fixed, and m gets large, (i. e. $m \to \infty$) we find that

$$\left(0.51\,\frac{e^{n-\frac{1}{2}}}{(n-1)^n}\right)^{\frac{1}{m+1-n}} \to 1 \quad \text{and} \quad (m+1)^{\frac{n}{m+1-n}} \to 1$$

For $\vartheta > 2$ the right side of (69) is obviously greater than 1. So, we can state for all $\vartheta > 2$ and suitable $m \gg n$ that

$$\left(0.51\,\frac{e^{n-\frac{1}{2}}}{(n-1)^n}\right)^{\frac{1}{m+1-n}}\cdot(m+1)^{\frac{n}{m+1-n}} \leq \frac{2\vartheta}{\vartheta+2}\cdot\vartheta^{\frac{1}{m+1-n}} \tag{70}$$

Consequently, for fixed n and m large enough $\tilde{t}$ can be chosen as close to $2\frac{\sqrt{n-1}}{\sqrt{2\pi}}$ as desired (remembering the substitution $t = \vartheta\frac{\sqrt{n-1}}{\sqrt{2\pi}}$ and (70)). Then asymptotically, i. e. for fixed n and m large enough, this $\tilde{t}$ is suitable for insertion into (63). The result of that consideration is

$$\begin{aligned} E[|\underline{\Delta}(T(\underline{X}_P))|] &\le (m+1)\left(\ln(1+\tilde{t}\,) + \tilde{t}\ln 2\right) \\ &= (m+1)\left(\ln\left(1+2\tfrac{\sqrt{n-1}}{\sqrt{2\pi}}\right) + 2\tfrac{\sqrt{n-1}}{\sqrt{2\pi}}\ln 2\right) \\ &\le 2(m+1)\sqrt{n-1} \end{aligned}$$

and so the following theorem is proved.

THEOREM 2.4. *For arbitrary, but fixed n and m large enough ($m \gg n$), we have*

$$E[|\underline{\Delta}(T(\underline{X}_P))|] \le const. \cdot \sqrt{n} \cdot m$$

where const. is a constant independent of n and m.

Remark 2.13. The theorem shows that $E[|\underline{\Delta}(T(\underline{X}_P))|] \le O(\sqrt{n}\,m)$. But as we have considered the asymptotic case and assumed n to be fixed we may argue that the factor $\sqrt{n}$ is a constant and this result may also be stated as $E[|\underline{\Delta}(T(\underline{X}_P))|] \le O(m)$.

We can prove something more than this asymptotic result. In particular, we are interested in finding configurations of m and n, for which the inequality (69) is satisfied with some special value for ϑ.
If we observe only values of m which are larger than $\bar{k}n - 1$, i. e. $m \ge \bar{k}n - 1$ for a fixed $\bar{k} \ge 5$ and $n \ge 3$ and if we restrict to $\vartheta > 2$, then we may argue as follows and this will lead to a successful analysis.
Recall the inequality (69), which would be sufficient for bounding $\tilde{t}$ resp. ϑ. This inequality is equivalent to

$$\begin{aligned} &\tfrac{1}{m+1-n} \cdot \left(\ln 0.51 + n\ln\left(\tfrac{m+1}{n-1}\right) + \left(n - \tfrac{1}{2}\right)\right) \le \ln\left(\tfrac{2\vartheta}{\vartheta+2}\right) + \tfrac{1}{m+1-n} \cdot \ln\vartheta \\ \Leftrightarrow \quad & \frac{\ln\left(\frac{0.51}{\vartheta}\right) + n\ln\left(\frac{m+1}{n-1}\right) + \left(n - \frac{1}{2}\right)}{m+1-n} \le \ln\left(\frac{2\vartheta}{\vartheta+2}\right). \end{aligned} \tag{71}$$

If we choose $\vartheta > 2$ then the right-hand side of (71) is positive ($\ln\left(\frac{2\vartheta}{\vartheta+2}\right) > 0$). Moreover, the left-hand side is monotonically decreasing in m and for $m \ge \bar{k}n - 1$ we have

$$\begin{aligned} \frac{\ln\left(\frac{0.51}{\vartheta}\right) + n\ln\left(\frac{m+1}{n-1}\right) + \left(n-\frac{1}{2}\right)}{m+1-n} &\le \frac{\ln\left(\frac{0.51}{\vartheta}\right) + n\ln\left(\bar{k}\frac{n}{n-1}\right) + \left(n-\frac{1}{2}\right)}{(\bar{k}-1)n} \\ &\le \tfrac{1}{\bar{k}-1}\left(\tfrac{1}{n}\ln\left(\tfrac{0.51}{\vartheta}\right) + \ln\left(\bar{k}\tfrac{n}{n-1}\right) + 1 - \tfrac{1}{2n}\right) =: \tfrac{1}{\bar{k}-1}\,q(\bar{k}, n). \end{aligned}$$

An upper bound for $q(\bar{k}, n)$ is $\ln \bar{k} + 1$, because for $\vartheta > 2$, $n \geq 3$ we have

$$\frac{1}{n} \ln\left(\frac{0.51}{\vartheta}\right) + \ln\left(\bar{k}\tfrac{n}{n-1}\right) + 1 - \frac{1}{2n} \leq -\frac{1}{n} + \ln \bar{k} + \ln\left(\tfrac{n}{n-1}\right) + 1 - \frac{1}{2n}$$
$$\leq \ln \bar{k} + 1 + \frac{1}{n-1} - \frac{3}{2n} \leq \ln \bar{k} + 1 .$$

Now, the only thing we have to do is to choose and to find a $\vartheta > 2$ such that $\frac{\ln \bar{k}+1}{\bar{k}-1} \leq \ln\left(\frac{2\vartheta}{\vartheta+2}\right)$. This is equivalent to $(e \cdot \bar{k})^{\frac{1}{\bar{k}-1}} \leq \frac{2\vartheta}{\vartheta+2}$. And we can easily show that by setting

$$\vartheta = \vartheta(\bar{k}) := \frac{2(e \cdot \bar{k})^{\frac{1}{\bar{k}-1}}}{2 - (e \cdot \bar{k})^{\frac{1}{\bar{k}-1}}},$$

we have $\frac{2\vartheta(\bar{k})}{\vartheta(\bar{k})+2} = (e \cdot \bar{k})^{\frac{1}{\bar{k}-1}}$. Finally, we use this $\vartheta(\bar{k})$ to bound $\tilde{t}$ by

$$\tilde{t} \leq \tilde{t}(\bar{k}) := \vartheta(\bar{k}) \tfrac{\sqrt{n-1}}{\sqrt{2\pi}}. \tag{72}$$

This (and (63)) leads to an upper bound for the expected complexity

$$\begin{aligned}
&E[|\underline{\Delta}(T(\underline{X}_P))|] \leq (m+1)\left(\ln(1+\tilde{t}) + \tilde{t} \ln 2\right) \\
&\leq (m+1)\left(\ln\left(1 + \vartheta(\bar{k})\tfrac{\sqrt{n-1}}{\sqrt{2\pi}}\right) + \vartheta(\bar{k})\tfrac{\sqrt{n-1}}{\sqrt{2\pi}} \ln 2\right) \\
&\leq (m+1)\left(\ln\left(1 + 48.04\tfrac{\sqrt{n-1}}{\sqrt{2\pi}}\right) + 48.04\tfrac{\sqrt{n-1}}{\sqrt{2\pi}} \ln 2\right) \\
&\leq (m+1)\left(17\sqrt{n-1}\right) = O(m\sqrt{n}) .
\end{aligned}$$

As we are ready now, we can state the final theorem.

THEOREM 2.5. *For $n \geq 3$, $m \geq 5n - 1$ the average number of steps of the* PHASE-I ALGORITHM *is at most $O(m\sqrt{n})$ under the uni-RSM.*

Theorem 2.5 gives explicit configurations of m and n, for which we can calculate an upper bound on the average number of iterations in the described way.

Remark 2.14. Concerning the different parameters note the following:
For $n \geq 3$, $\bar{k} \geq 5$, $m \geq \bar{k}n - 1$ the term $\vartheta(\bar{k})$ is well defined. $\vartheta(\bar{k})$ is motonoically decreasing in $\bar{k}$ and we have $\vartheta(\bar{k}) \leq 48.04$ and $\lim_{\bar{k}\to\infty} \vartheta(\bar{k}) = 2$, and analogously we have $\tilde{t}(\bar{k}) \leq 19.2\sqrt{n-1}$ and $\lim_{\bar{k}\to\infty} \tilde{t}(\bar{k}) = \sqrt{\frac{2}{\pi}(n-1)}$. Remember, that $\tilde{t}(\bar{k})$ represents an upper bound on $\tilde{t}$ — the value, where our lower bound on the distribution function F_T starts to be nontrivial.

3. A PHASE-II-ALGORITHM AND A TERMINATION PROCEDURE TO STOP INTERIOR-POINT-METHODS

3.1. An Interior-Point-Method for Phase II

For completeness we briefly discuss a typical IPM using the barrier function approach (compare [9] and [22]). In this section we may assume the LP to have an optimal solution, as it is guaranteed at the end of phase I. (In the case that the LP is unbounded we stop after phase I and do not proceed with phase II.)

3.1.1. Properties of Newton-Steps

Some notations have already been introduced in section 2.2.3. The barrier function for the problem (P) is defined as in (23) by

$$\phi_P(x,\mu) := -\frac{v^T x}{\mu} - \sum_{i=1}^{m} \ln(1 - a_i^T x)$$

where μ is a positive parameter. We know that the barrier function $\phi_P(x,\mu)$ achieves its minimum over X_P at a unique point $x(\mu)$, namely the analytic μ-center. The measure for the distance of an interior feasible point x to $x(\mu)$ was defined in (25) as

$$\delta(x,\mu) = ||p(x,\mu)||_{H(x,\mu)} = ||S^{-1}Ap(x,\mu)|| = \left\| \frac{Sy(s,\mu)}{\mu} - e \right\|,$$

where $p(x,\mu)$ is the Newton direction, $H(x,\mu)$ is the Hessian matrix and $g(x,\mu)$ is the gradient of $\phi_P(x,\mu)$ at x and $y(s,\mu) = \mu(S^{-1}e + S^{-2}Ap(x,\mu))$ (compare (27)), i. e.

$$g(x,\mu) = -\frac{1}{\mu}v + A^T S^{-1} e\,, \quad H(x,\mu) = A^T S^{-2} A$$
$$\text{and} \quad p(x,\mu) = -H(x,\mu)^{-1} g(x,\mu)\,.$$

An approximate μ-center $x \in \mathit{Int}\, X_P$ is characterized by $\delta(x,\mu) < 1$.

The idea of a barrier method is to approximate the μ-analytic center (the minimal point of the barrier function) for fixed μ and then reduce the barrier parameter μ and start again. For approximating the μ-analytic center we will use Newton-steps. But before presenting a corresponding algorithm we recall some fundamental properties. For proofs of the subsequent lemmata we refer to [9].

LEMMA 3.1. *1. For $x \in \mathit{Int}\, X_P$ and $d \in \mathbb{R}^n$ with $||d||_{H(x,\mu)} < 1$ we have $x + d \in \mathit{Int}\, X_P$.*

2. If $\delta(x,\mu) < 1$, then $x_+ := x + p(x,\mu) \in \mathit{Int}\, X_P$ and $\delta(x_+,\mu) \le \delta(x,\mu)^2$.

3. For $\mu_+ := (1-\bar{\eta})\mu$ with $\bar{\eta} \in (0,1)$ we have $\delta(x,\mu_+) \le \frac{1}{1-\bar{\eta}}(\delta(x,\mu) + \bar{\eta}\sqrt{m})$.

4. Let $\delta(x,\mu) \le \frac{1}{2}$, $\eta = \frac{1}{6\sqrt{m}}$, $x_+ := x + p(x,\mu)$ and $\mu_+ := (1-\eta)\mu$. Then we have $\delta(x_+,\mu_+) \le \frac{1}{2}$.

LEMMA 3.2. *If $\delta(x,\mu) \le 1$, then $y(s,\mu) = \mu(S^{-1}e + S^{-2}Ap(x,\mu))$ is a feasible point for (D) and*

$$\mu(m - \delta(x,\mu)\sqrt{m}) \le s^T y(s,\mu) \le \mu(m + \delta(x,\mu)\sqrt{m})\,.$$

Before we introduce the algorithm we want to discuss how to stop.

3.1.2. Stopping Interior-Point-Methods

For a given point $x \in Int\, X_P$ we try to find a vertex $\bar{x}$ of X_P with an improved objective function value, i. e. $v^T\bar{x} \geq v^T x$. And we can check whether or not the vertex $\bar{x}$ is the optimal vertex. The way to find such a vertex is to project the given point onto the boundary of X_P using the objective vector v as direction of the projection. So, again let

$$I(x) := \{i \mid {a_i}^T x = 1\} \subset \{1, \ldots, m\}$$

be the index set of active constraints and

$${A_I}^T := (a_{i_1}, \ldots, a_{i_j}) \quad \text{and} \quad P_I := E - {A_I}^T(A_I {A_I}^T)^{-1} A_I\,.$$

P_I is a projection matrix on the null space of A_I.

ALGORITHM 3 (PROCEDURE ROUNDING).

Input: $x \in Int\, X_P$

1. $l := 0;\ \xi_l := x;\ d_0 := v;$ $\{$*initialization*$\}$
2. $I(\xi_0) := \emptyset;$ $\{$*index set of active constraints*$\}$
3. **repeat**
4. COMPUTE $\alpha_l = \mathrm{Max}\{\alpha \mid \alpha \geq 0, \xi_l + \alpha d_l \in X_P\};$
5. $\xi_{l+1} := \xi_l + \alpha_l d_l;$
6. CHOOSE $i_{l+1} \in \{i \mid i \notin I_l, {a_i}^T \xi_{l+1} = 1\};$
7. $I_{l+1} := I_l \cup \{i_{l+1}\};$
8. COMPUTE $P_{I_{l+1}};$ $\{$*projection matrix*$\}$
9. $d_{l+1} := P_{I_{l+1}} d_l;$
10. $l := l + 1;$
11. **until** $l = n$;
12. COMPUTE $\tilde{y}^T := v^T A_{I_n}^{-1};$
13. **if** $\tilde{y} \geq 0$ **and** $\tilde{y}^T e_{(n)} = v^T \xi_n$ **then**
14. return *true* (exit); $\{$*ξ_n is the optimal vertex*$\}$
15. **else**
16. return *false* (exit); $\{$*ξ_n is a nonoptimal vertex*$\}$
17. **endif**

Output: *false:* ξ_n is a nonoptimal vertex of X_P with $v^T\xi_n > v^T\xi_0$ *or*
true: ξ_n is the optimal vertex of X_P.

THEOREM 3.1 (Complexity and Correctness of ALGORITHM 3).

1. PROCEDURE ROUNDING *terminates after at most n iterations.*

2. The effort of each iteration is at most $O(mn + n^3)$.

3. For all $n > l \geq 0$ *we have* $\xi_{l+1} \in X_P$ *and* $v^T\xi_{l+1} \geq v^T\xi_l > v^T\xi_0$.

4. If PROCEDURE ROUNDING *returns* true, *then* ξ_n *is the optimal vertex of* X_P .

5. If PROCEDURE ROUNDING *returns* false, *then* ξ_n *is a nonoptimal vertex of* X_P *with* $v^T\xi_n > v^T\xi_0$.

Proof. *Part 1:* The variable l is initialized (in line 1) with the value 0, incremented by 1 (in line 10) and we stop if $l = n$ (line 11) at the latest.

Part 2: Computing $\bar{\alpha}_l$ raises $O(mn)$ arithmetic operations because we have to take into account m constraints. The computation of the projection matrix $P_{I_{l+1}}$ can be done in $O(n^2)$ arithmetic operations using update formulas, but the calculation of $\tilde{y}$ may cost $O(n^3)$, whereas all other statements are done in at most $O(n)$ arithmetic operations.

Part 3: Because of the input and line 4/5 it is clear, that $\xi_{l+1} \in X_P$ for all l. Furthermore, we have

$$\xi_{l+1} = \xi_l + \alpha_l d_l \quad \text{and} \quad d_l = P_{I_{l-1}} d_{l-1} = P_{I_{l-1}} v\,,$$

and for $l = 0$ we can conclude that $v^T \xi_1 = v^T(\xi_0 + \alpha_0 d_0) = v^T \xi_0 + \alpha_0 v^T v > v^T \xi_0$ because $\alpha_0 > 0$ as $\xi_0 \in \mathit{Int}\, X_P$ and $\xi_1 \in \partial X_P$, and for $l \geq 1$ that $v^T \xi_{l+1} = v^T(\xi_l + \alpha_l d_l) = v^T \xi_l + \alpha_l v^T P_{I_{l-1}} v \geq v^T \xi_l > v^T \xi_0$ because $\alpha_l \geq 0$ and $P_{I_{l-1}}$ are projection matrices and positive semidefinite.

Part 4: If Procedure Rounding returns *true*, then $\tilde{y}$ satisfies $A_{I_n}^T \tilde{y} = v$, $\tilde{y} \geq 0$ and $e_{(n)}^T \tilde{y} = v^T \xi_n$. The condition of nondegeneracy guarantees that the matrix A_{I_n} is of full rank and therefore $\tilde{y}$ is well defined. Defining $\bar{x} := \xi_n$ and $\bar{y} \in \mathbb{R}^m$ by

$$\bar{y}_k := \begin{cases} \tilde{y}_j & \text{for } k = i_j \in I_n, \\ 0 & \text{otherwise;} \end{cases}$$

we see that $\bar{x} \in X_P$, $\bar{y} \in X_D$ and $v^T \bar{x} = e^T y$. So, from duality theory we conclude that $\bar{x}$ and $\bar{y}$ are optimal for (P) resp. (D). Moreover, as $A_{I_n} \xi = e_{(n)}$ we have n active (and linearly independent) constraints at ξ_n, this point ξ_n has to be a vertex of X_P.

Part 5: Procedure Rounding returns *false*, then nevertheless ξ_n is a vertex of X_P for the same reasons as in the proof of part 4 and $v^T \xi_n > v^T \xi_0$ because of part 3.

Now assume that ξ_n is optimal: Then ξ_n satisfies together with some dual optimal solution y the conditions of complementary slackness and the primal and dual objective function values at ξ_n resp. y are equal.
As a result of nondenegeneracy we know that $a_i^T x < 1$ for all $i \notin I_n$ and so the condition of complementary slackness enforces $y_i = 0$ for all $i \notin I_n$. Let y_{I_n} denote the reduced vector which contains those components y_i of y with $i \in I_n$. Then $v = A^T y = A_{I_n}^T y_{I_n}$, $y_{I_n} \geq 0$ and $v^T \xi_n = {e_{(m)}}^T y = {e_{(n)}}^T y_{I_n}$ and the reduced vector y_{I_n} satisfies the conditions that we asked for $\tilde{y}$ in the procedure and that would have forced the output *true*. This is a contradiction as we assumed the output to be *false* and therefore ξ_n could not be the optimal solution of (P). ∎

Remark 3.1. Let x_{opt} denote the optimal vertex of (P) and x_{II} the second best vertex of (P). If we start Procedure Rounding with $x \in \mathit{Int}\, X_P$ and $v^T x > v^T x_{II}$, then Procedure Rounding ends up with output *true* and $\xi_n = x_{opt}$.

3.1.3. A Phase-II-Algorithm

Now we use the information of the previous sections to formulate an interior-point-method for linear programming problems of type (P).

ALGORITHM 4 (BARRIER METHOD).

Input: x, $\mu > 0$ with $\delta(x,\mu) \le \frac{1}{2}$ and $\eta := \frac{1}{6\sqrt{m}}$.

1. $k := 0;\ x_k := x;\ \mu_k := \mu;\ s_k := e - Ax_k;$ $\{$*initialization*$\}$
2. COMPUTE ROUNDING(x_k);
3. **while** ROUNDING(x_k)= *false* **do**
4. COMPUTE $p(x_k,\mu_k) := \frac{1}{\mu_k}(A^T S_k^{-2} A)^{-1}(v - \mu_k A^T S_k^{-1} e);$
5. $x_{k+1} := x_k + p(x_k,\mu_k);$
6. $s_{k+1} := e - Ax_{k+1};$
7. $\mu_{k+1} := (1-\eta)\mu_k;$
8. $k := k+1;$
9. **endwhile**

Output

THEOREM 3.2 (Complexity and Correctness of ALGORITHM 4).

1. After k iterations the duality gap $s_k^T y(s_k,\mu_k)$ *is less than* $\mu_k(m+\sqrt{m})$.

2. Let $U := v^T x_{opt} - v^T x_{II}$. *After* K_U *iterations with*

$$K_U := \begin{cases} 0 & \textit{if } \mu_0(m+\sqrt{m}) \le U\,, \\ 6\sqrt{m}(-\ln U + \ln(\mu_0(m+\sqrt{m})) & \textit{if } \mu_0(m+\sqrt{m}) > U\,; \end{cases} \tag{73}$$

we have the following bound for the duality gap: $s_{K_U}^T y(s_{K_U},\mu_{K_U}) \le U$ *and the* BARRIER METHOD *stops.*

3. The BARRIER METHOD *terminates after at most* $O(\sqrt{m}(L+\ln m)) = O(\sqrt{m}L)$ *iterations in the worst case (using the complexity model of 2.1.4).*

4. The effort of each iteration is at most $O(mn^2 + n^4)$.

Proof. *Part 1:* Due to Lemma 3.2 and $\delta(x_k\mu_k) \le \frac{1}{2}$ we have

$$s_k^T y(s_k,\mu_k) \le \mu_k(m + \delta(x_k\mu_k)\sqrt{m}) \le \mu_k(m+\sqrt{m})\,.$$

Part 2: According to Remark 3.1 the BARRIER METHOD stops if $v^T x_K > v^T x_{II}$ or equivalently if $v^T x_{opt} - v^T x_K < v^T x_{opt} - v^T x_{II} = U$ at some iterate x_K . What we can ensure by Lemma 3.2 is

$$v^T x_{opt} - v^T x_K \le s_K^T y(s_K,\mu_K) \le \mu_K(m+\sqrt{m}) \le (1-\eta)^K \mu_0(m+\sqrt{m})\,.$$

The last inequality holds because of line 7 of the BARRIER METHOD. Now, for termination it is sufficient to choose K such that

$$(1-\eta)^K \mu_0(m+\sqrt{m}) \le U \quad \text{resp.} \quad (1-\eta)^K \le \frac{U}{\mu_0(m+\sqrt{m})}\,. \tag{74}$$

In the case that $\mu_0(m+\sqrt{m}) \le U$ it follows that $\frac{U}{\mu_0(m+\sqrt{m})} > 1$ whereas the righthand side $(1-\eta)^K$ of (74) is smaller than 1 for all $K \ge 0$. So the BARRIER

METHOD stops immediately and the vertex calculated by PROCEDURE ROUNDING is the optimal solution of (P).
In the case of $\mu_0(m+\sqrt{m}) > U$ we obtain from (74)

$$\begin{aligned}(1-\eta)^K \le \tfrac{U}{\mu_0(m+\sqrt{m})} \quad &\Leftrightarrow \quad K\ln(1-\eta) \le \ln\left(\tfrac{U}{\mu_0(m+\sqrt{m})}\right)\\ &\Leftrightarrow \quad K \ge \frac{1}{\ln(1-\eta)}(\ln U - \ln(\mu_0(m+\sqrt{m}))\,.\end{aligned}$$

So, for termination it is sufficient to choose $K \ge \frac{1}{\ln(1-\eta)}(\ln U - \ln(\mu_0(m+\sqrt{m}))$. Because of $\ln(1-\eta) \le -\eta$ and $\mu_0(m+\sqrt{m}) > U$ we have

$$\frac{1}{\ln(1-\eta)}(\ln U - \ln(\mu_0(m+\sqrt{m})) \le -\frac{1}{\eta}(\ln U - \ln(\mu_0(m+\sqrt{m}))$$

and it is sufficient (for termination) to choose any K with

$$K \ge -\frac{1}{\eta}(\ln U - \ln(\mu_0(m+\sqrt{m})) = 6\sqrt{m}(-\ln U + \ln(\mu_0(m+\sqrt{m}))\,.$$

Part 3: Using part 2 we see that the BARRIER METHOD has to stop after at most $O(\sqrt{m}(-\ln U + \ln(\mu_0(m+\sqrt{m})))$ iterations. From complexity theory we know that $U \ge O(2^{-L})$ and $\mu_0 \le 2^L$ in the worst case

$$\begin{aligned}K &= O(\sqrt{m}(-\ln U + \ln(\mu_0(m+\sqrt{m}))\\ &\le O(\sqrt{m}(L + L + \ln(m+\sqrt{m}))) = O(\sqrt{m}L)\,.\end{aligned}$$

Part 4: The computation of $\underline{H}(x_k, \mu_k)$ can be done in $O(mn^2)$ arithmetic operations and the effort of inverting this matrix and for PROCEDURE ROUNDING is at most $O(mn^2 + n^4)$. All other calculations can be done with less effort. ▮

We have seen that the complexity/number of iterations of the BARRIER METHOD depends on U (the difference between the optimal and the second best vertex) resp. $\ln U$. As we are interested not only in the worst case, but also in the average case analysis we will study the distribution of U.

3.2. Distribution of the Difference Between the Objective Values at the Best and Second Best Vertex

In this chapter we try to develop a distribution for the difference between the objective values of the two best vertices.

We start by presenting the underlying geometry. Then we will derive an exact integral representation of the distribution and continue with approximations of different parts of this integral/distribution function. In section 3.2.4 we merge the results of the previous sections to derive the desired distribution function and expectation values.

3.2.1. The Difference Between the Best and Second Best Vertex

Two Adjacent Vertices. To characterize two adjacent vertices, especially the best and the second best vertex, we make use of two sequential coordinate transformations. The first transformation replaces the points $a_1, \ldots, a_m$ by new points $b_1, \ldots, b_m$, such that $b_1^n = \ldots = b_n^n$ and that the basic simplex $conv(b_1, \ldots, b_n)$ is fully located in the level of height $h \in (0,1)$. So, we have $b_1^n = h, \ldots, b_n^n = h$.
And a second transformation maps $b_1, \ldots, b_m$ to new points $c_1, \ldots, c_m$ with the property $c_1^n = h, \ldots, c_n^n = h$ and $c_1^{n-1} = \theta, \ldots, c_{n-1}^{n-1} = \theta$, so the $(n-1)$-th coordinate of the first $n-1$ points becomes equal. The remaining degree of freedom in that transformation can be used to ascertain that $-\sqrt{1-h^2} < c_n^{n-1} < \theta < \sqrt{1-h^2}$. That means that we carry out a second rotation of $\mathbb{R}^{n-1}$, which keeps the nth coordinate axis unchanged. So we concentrate on the configuration, where c_n is "left" of $aff(c_1, \ldots, c_{n-1})$, which will in further applications be regarded as a certain rotation axis.

In our notation we use $\bar{c}_i$ for the vector $(c_i^1, \ldots, c_i^{n-1})^T \in \mathbb{R}^{n-1}$ and $\bar{\bar{c}}_i$ for $(c_i^1, \ldots, c_i^{n-2})^T \in \mathbb{R}^{n-2}$. In the new space we have a basic simplex $\Sigma(\Delta) = conv(c_i,\ i \in \Delta)$ with the basic index set $\Delta = \{1, \ldots, n\}$ and a basic solution $x_\Delta = \frac{1}{h} e_n = (0, \ldots, 0, h)^T$. x_Δ is the unique solution of $c_1^T x = 1, \ldots, c_n^T x = 1$.

It is clear that (according to Lemma 2.8) x_Δ is a vertex of X if and only if $\Sigma(\Delta)$ is a boundary simplex of $Y = conv(0, c_1, \ldots, c_m)$. And this means that $c_{n+1}^n, \ldots, c_m^n \le h$.
In the following we will work under the condition that x_Δ is the optimal vertex on X with respect to the objective $v^T x$. From the Lemma of Farkas it is then clear that in the dual space $v \in cone(c_1, \ldots, c_m)$.

Now let us think about those vertices on X, which are adjacent to x_Δ. It is known that one of these will be the second best vertex. And there is an edge connecting x_Δ with that second best vertex. Such an edge keeps $n-1$ of the restrictions $c_1^T x \le 1, \ldots, c_n^T x \le 1$ tight (as they are active in x_Δ), but it loosens one of those n restrictions. W.l.o.g. let the nth restriction be that to be loosened. So the edge under consideration has the form $\{x \in X \mid c_1^T x = 1, \ldots, c_{n-1}^T x = 1\}$. This edge is on one side bounded by x_Δ (where $c_n^T x \le 1$ is tight, too) and on the other side there are two options. Either this edge is unbounded, which means that a move away from x_Δ on our edge relaxes all other restrictions $c_n^T x \le 1, \ldots, c_m^T x \le 1$. Or it is bounded, which means that our move ends at another vertex $x_{\Delta'}$, where our move decreases some (at least one) of the slacks $1 - c_i^T x$ $(i \ge n+1)$. And x_Δ is the point, where the first of these slacks becomes 0 (the first of these restrictions becomes tight).
Again, w.l.o.g. let $c_{n+1}^T x \le 1$ be that critical restriction. In the dual space that means that $c_1, \ldots, c_{n-1}$ and c_{n+1} span another basic simplex of Y. And this is also a boundary simplex of Y, if and only if $c_n, c_{n+2}, \ldots, c_m$ lie "below" the hyperplane spanned by $c_1, \ldots, c_{n-1}$ and c_{n+1} ("below" means in the same halfspace as the origin).
We can formalize the alternative above in the following way:

1. $\{x \in X \mid c_1^T x = \cdots = c_{n-1}^T x = 1\}$ is an infinite edge of X starting at x_Δ, if and only if $conv(0, c_1, \ldots, c_{n-1})$ is a boundary simplex of Y.
This can also be characterized by the following condition:
For $k = n, n+1, \ldots, m$ it holds that

$$\begin{pmatrix} h \\ -\theta \end{pmatrix}^T \begin{pmatrix} c_k^{n-1} \\ c_k^n \end{pmatrix} \leq 0\,, \qquad \text{resp. } c_k^{n-1} h - c_k^n \theta \leq 0\,.$$

2. $\{x \in X \mid c_1^T x = \cdots = c_{n-1}^T x = 1\}$ is a bounded edge of X starting at x_Δ and ending at a vertex $x_{\Delta'}$ with $\Delta' = \{1, \ldots, n-1, n+1\}$ if and only if $conv(c_1, \ldots, c_{n-1}, c_{n+1})$ is a boundary simplex of Y.
In this case it is necessary that

$$\begin{pmatrix} h \\ -\theta \end{pmatrix}^T \begin{pmatrix} c_{n+1}^{n-1} \\ c_{n+1}^n \end{pmatrix} > 0 \qquad (\text{resp. } c_{n+1}^{n-1} h - c_{n+1}^n \theta > 0)$$

and that for all $k = n, n+2, \ldots, m$ it holds that

$$\begin{pmatrix} x_{\Delta'}{}^{n-1} \\ x_{\Delta'}{}^{n} \end{pmatrix}^T \begin{pmatrix} c_k^{n-1} \\ c_k^n \end{pmatrix} \leq 1 \text{ resp. } \frac{(h - c_{n+1}^n) c_k^{n-1} + (c_{n+1}^{n-1} - \theta) c_k^n}{c_{n+1}^{n-1} h - c_{n+1}^n \theta} \leq 1\,,$$

where $x_{\Delta'} = \dfrac{h - c_{n+1}^n}{c_{n+1}^{n-1} h - c_{n+1}^n \theta}\, e_{n-1} + \dfrac{c_{n+1}^{n-1} - \theta}{c_{n+1}^{n-1} h - c_{n+1}^n \theta}\, e_n\,.$

This reflects the fact that $x_{\Delta'}$ is the normal vector on the affine hull of $c_1, \ldots, c_{n-1}$, c_{n+1} $(= \mathit{aff}(c_1, \ldots, c_{n-1}, c_{n+1}))$ and that $c_1^T\, x_{\Delta'} = \cdots = c_{n-1}^T\, x_{\Delta'} = c_{n+1}^T\, x_{\Delta'} = 1$.

It will be useful to know that

$$h' := \frac{1}{\|\, x_{\Delta'} \,\|} = \frac{c_{n+1}^{n-1} h - c_{n+1}^n \theta}{\sqrt{(h - c_{n+1}^n)^2 + (c_{n+1}^{n-1} - \theta)^2}} \text{ and } \|\, x_{\Delta'} \,\| = \frac{1}{h'}$$

and that this is the distance from the origin to $H(c_1, \ldots, c_{n-1}, c_{n+1})$.

Now it is possible to define an index set $K_{h,\theta}$ as follows

$$K_{h,\theta} := \{k \geq n+1 \mid c_k^{n-1} h - c_k^n \theta > 0\}\,.$$

This is the index set of restrictions, whose slacks are decreased when we move away from x_Δ on our edge with direction $(0, \ldots, 0, h, -\theta)$.
Our combination of two boundary simplices can also be interpreted as a kink on the surface of Y. Therefore note that the two simplices $\Sigma(\Delta) = conv(c_1, \ldots, c_{n-1}, c_n)$ and $\Sigma(\Delta')$, resp. $conv(c_1, \ldots, c_{n-1}, c_{n+1})$ have a common side simplex $conv(c_1, \ldots, c_{n-1})$, which is the intersection of these two sets. The affine hull $\{x \mid x^{n-1} = \theta, x^n = h\}$ of this intersection set can be regarded as a kink or rotation axis. It is important to determine the kink-angle φ between the two hyperplanes $H(c_1, \ldots, c_{n-1}, c_n)$ and $H(c_1, \ldots, c_{n-1}, c_{n+1})$ as $\varphi := arc(x_\Delta, x_{\Delta'})$. From this rotation axis we can associate with every point $0, c_{n+1}, \ldots, c_m$ such a rotation angle in $(0, \pi)$.

$$\varphi_0 := \arccos \frac{-\theta}{\sqrt{h^2 + \theta^2}} \quad \text{and} \tag{75}$$

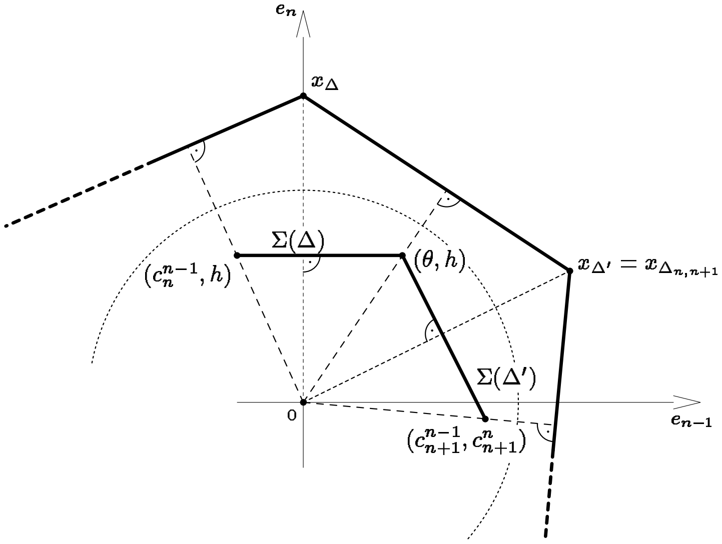

FIG. 4. After the rotations we have the following situation: The vertices x_Δ and $x_{\Delta'}$ are lying in the (e_{n-1}, e_n)-plane, and the points $c_1, \ldots, c_{n-1}$ are projected to (θ, h), c_n is projected to (c_n^{n-1}, h) on the same level, and c_{n+1} is the point which determines the adjacent vertex, resp. the adjacent boundary simplex $\Sigma(\Delta')$.

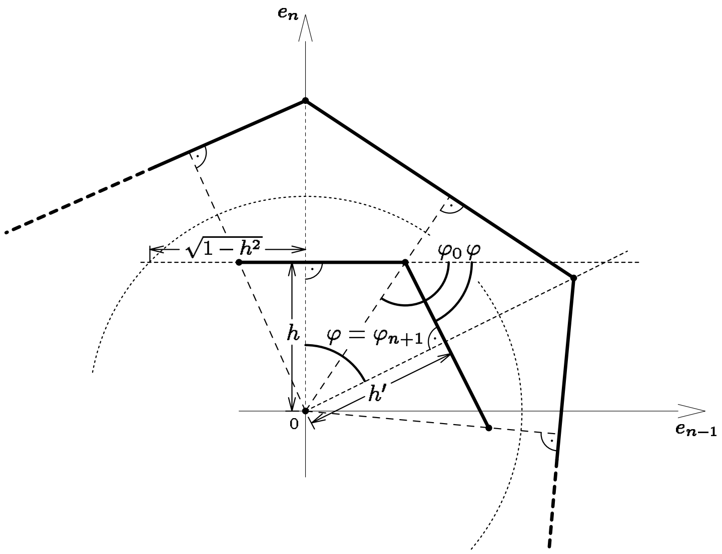

FIG. 5. This figure illustrates the heights h, h' and the angles φ_0, $\varphi = \varphi_{n+1}$ according to the situation of figure 4.

$$\varphi_k := \arccos \frac{c_k^{n-1} - \theta}{\sqrt{(h - c_k^n)^2 + (c_k^{n-1} - \theta)^2}} \qquad \text{for } k \geq n+1\,.$$

It is immediate that $conv(c_1, \ldots, c_{n-1}, c_{n+1})$ is a boundary simplex of Y if and only if

$$\varphi_{n+1} = \text{Min}\{\varphi_0, \varphi_{n+1}, \varphi_{n+2}, \ldots, \varphi_m\}\,.$$

And there is no adjacent basic simplex as boundary simplex, resp. the x_Δ-incident edge is unbounded if and only if

$$\varphi_0 = \text{Min}\{\varphi_0, \varphi_{n+1}, \varphi_{n+2}, \ldots, \varphi_m\}\,.$$

In case of a boundary simplex $conv(c_1, \ldots, c_{n-1}, c_{n+1})$ it is easy to calculate

$$h' = \frac{c_{n+1}^{n-1} h - c_{n+1}^n \theta}{\sqrt{(h - c_{n+1}^n)^2 + (c_{n+1}^{n-1} - \theta)^2}} = h \cos \varphi_{n+1} + \theta \sin \varphi_{n+1}\,. \tag{76}$$

And we have

$$x_{\Delta'} = \frac{1}{h \cos \varphi_{n+1} + \theta \sin \varphi_{n+1}} (\cos \varphi_{n+1}\, e_{n-1} + \sin \varphi_{n+1}\, e_n).$$

Note that in our stochastic model (*uni*-RSM) the intersection of the unit sphere ω_n with the hyperplane $H(c_1, \ldots, c_{n-1}, c_n)$ is a sphere of radius $\sqrt{1-h^2}$ and the intersection of ω_n with the $H(c_1, \ldots, c_{n-1}, c_{n+1})$ is a sphere of radius $\sqrt{1 - h^2 \cos \varphi_{n+1}^2 - \theta^2 \sin \varphi_{n+1}^2 - 2h\theta \sin \varphi_{n+1} \cos \varphi_{n+1}}\,.$

The Difference of Objective Values. We work under the assumption that x_Δ is the optimal vertex (which is equivalent to $v \in cone(c_1, \ldots, c_n)$).

Incident to x_Δ is the edge $\{x \in X \mid c_1^T x = \cdots = c_{n-1}^T x = 1\}$, which turns out to be either a ray of the form $x_\Delta + \mathbb{R}\,(0, \ldots, 0, h, -\theta)^T$ or a finite line of the form $[x_\Delta, x_{\Delta'}]$, where $x_{\Delta'} - x_\Delta$ is a positive multiple of $(0, \ldots, 0, h, -\theta)^T$.

Now we try to analyze the behavior of the objective function $v^T x$ when we leave x_Δ and run along that edge. In both cases it is clear that $v^T x$ will decrease, because x_Δ had been optimal. This can also be seen from a decomposition of v into three orthogonal components:

1. a multiple of $(0, \ldots, 0, \theta, h)^T$ called $\widetilde{v}$,
2. a vector of $\{x \in \mathbb{R}^n \mid x^{n-1} = x^n = 0\} \,\widehat{=}\, \mathbb{R}^{n-2}$ called $\overline{\overline{v}}$,
3. a vector orthogonal to $\{x \in \mathbb{R}^n \mid x^{n-1} = x^n = 0\}$ and $(0, \ldots, 0, \theta, h)^T$ in direction $(0, \ldots, 0, -h, \theta)^T$, called $\widehat{v}$.

So we get $v = \widetilde{v} + \overline{\overline{v}} + \widehat{v}$.

Since the direction of our edge is $(0, \ldots, 0, h, -\theta)^T$ and hence orthogonal to $\{x \in \mathbb{R}^n \mid x^{n-1} = x^n = 0\}$, it is clear that $\overline{\overline{v}}^T (0, \ldots, 0, h, -\theta)^T = 0$ and so $\overline{\overline{v}}$ does not influence the level of decrement of the objective.

Now let us think about $\widetilde{v}^T (x_\Delta - x_{\Delta'})$: This scalar product is 0, because $x_\Delta - x_{\Delta'}$ is a multiple of $(0, \ldots, 0, h, -\theta)^T$ and $\widetilde{v}$ is a multiple of $(0, \ldots, 0, \theta, h)^T$ and these

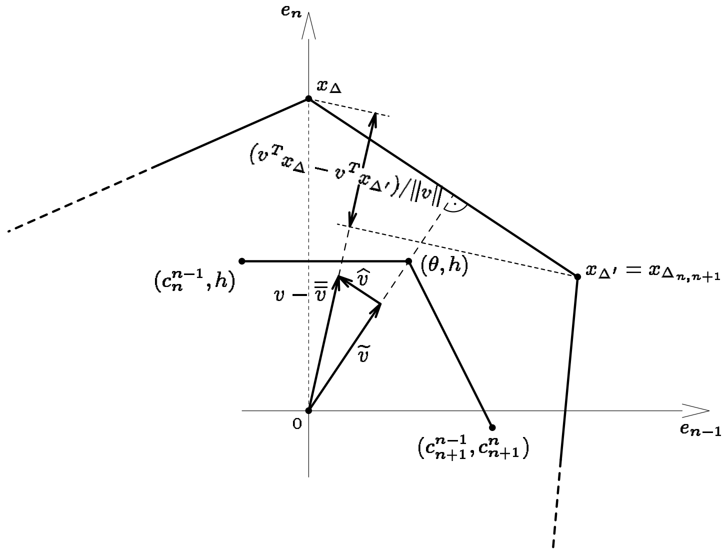

FIG. 6. Decomposition of v. As the figure shows the projection onto the (e_{n-1}, e_n)-plane, we cannot see v, but we do see $v - \overline{\overline{v}}$, which corresponds to the orthogonal projection of v onto the plane spanned by (e_{n-1}, e_n).

two vectors are orthogonal. We can determine $\widetilde{v}$ by measuring the distance of v to $\overline{H} := H(0, c_1, \ldots, c_{n-1})$:

$$dist(v, \overline{H}) = \frac{\left| \begin{pmatrix} 0_{(n-2)} \\ h \\ -\theta \end{pmatrix}^T \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \right|}{\sqrt{h^2 + \theta^2}} = \frac{|hv^{n-1} - \theta v^n|}{\sqrt{h^2 + \theta^2}} = \frac{h(v^{n-1} - \theta \frac{v^n}{h})}{\sqrt{h^2 + \theta^2}}.$$

So we know that $\widetilde{v} = \dfrac{|hv^{n-1} - \theta v^n|}{\sqrt{h^2 + \theta^2}} \begin{pmatrix} 0_{(n-2)} \\ h \\ -\theta \end{pmatrix}$.

If we apply the same distance measure to c_n, we obtain (since $c_n^n = h$ and using $\overline{\overline{c}}^T = (c_n^1, \ldots, c_n^{n-2})$)

$$dist(c_n, \overline{H}) = \frac{\left| \begin{pmatrix} 0_{(n-2)} \\ h \\ -\theta \end{pmatrix}^T \begin{pmatrix} \overline{\overline{c}} \\ c_n^{n-1} \\ h \end{pmatrix} \right|}{\sqrt{h^2 + \theta^2}} = \frac{|hc_n^{n-1} - \theta h|}{\sqrt{h^2 + \theta^2}} = \frac{h(c_n^{n-1} - \theta)}{\sqrt{h^2 + \theta^2}}.$$

Now, we see that $\dfrac{dist(v, \overline{H})}{dist(c_n, \overline{H})} = \dfrac{v^{n-1} - \theta \frac{v^n}{h}}{c_n^{n-1} - \theta} = \dfrac{\theta - v^{n-1}\frac{h}{v^n}}{\theta - c_n^{n-1}} \cdot \dfrac{v^n}{h}$.

On the other side we can exploit the fact that $v \in cone(c_1, \ldots, c_n)$ if and only if x_Δ is optimal, resp. $conv(c_1, \ldots, c_n)$ is the optimal boundary simplex. But then it is clear that $\frac{h}{v^n} v \in conv(c_1, \ldots, c_n)$. Hence $\frac{h}{v^n} v = \lambda c_n + (1-\lambda)c$ for a certain point $c \in cone(c_1, \ldots, c_{n-1})$ and $\lambda \in (0, 1)$, or equivalently $v = \frac{v^n}{h}\lambda c_n + \frac{v^n}{h}(1-\lambda)c$, where the vector $\frac{v^n}{h} c$ is an element of $\overline{H}$. We see that

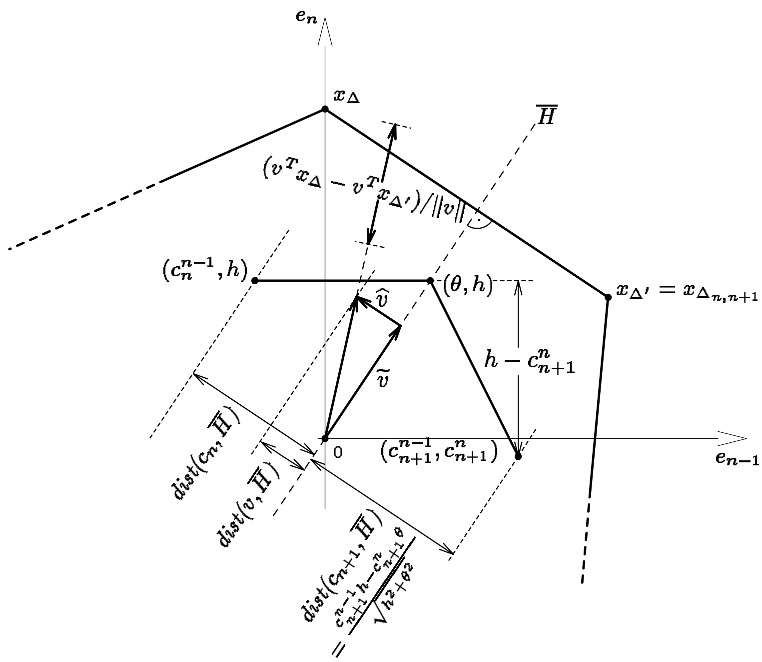

FIG. 7. Description of $v^T(x_\Delta - x_{\Delta'})$ and distances to $\overline{H}$ used in (77).

$$dist(v,\overline{H}) = dist(\frac{v^n}{h}\lambda\, c_n, \overline{H}) = \frac{v^n}{h}\lambda\, dist(c_n,\overline{H})\,,$$

because $0 \in \overline{H}$. And now it is clear that

$$\frac{\theta - v^{n-1}\frac{h}{v^n}}{\theta - c_n^{n-1}}\frac{v^n}{h} = \frac{v^n}{h}\lambda, \qquad \text{resp.} \qquad \lambda = \frac{\theta - v^{n-1}\frac{h}{v^n}}{\theta - c_n^{n-1}}\,.$$

That means that this λ is not only the internal coefficient for the intersection point of $\mathbb{R}^+ v$ in terms of $conv(c_1,\ldots,c_n)$, but also the share of the distance of that point to $\overline{H}$ compared with the maximal possible distance achievable in $conv(c_1,\ldots,c_n)$, namely $dist(c_n,\overline{H})$. This maximal distance arises when $\lambda = 1$, i.e. when we take $v := c_n \in \omega_n$. Now we can evaluate $v^T(x_\Delta - x_{\Delta'})$ as follows:

$$\begin{aligned}
v^T(x_\Delta - x_{\Delta'}) &= v^n\frac{1}{h} - v^{n-1}\frac{h - c_{n+1}^n}{c_{n+1}^{n-1}h - c_{n+1}^n\theta} - v^n\frac{c_{n+1}^{n-1} - \theta}{c_{n+1}^{n-1}h - c_{n+1}^n\theta} \\
&= \frac{h - c_{n+1}^n}{c_{n+1}^{n-1}h - c_{n+1}^n\theta}\left(\lambda\frac{v^n}{h}(\theta - c_n^{n-1})\right) \\
&= \frac{h - c_{n+1}^n}{c_{n+1}^{n-1}h - c_{n+1}^n\theta}\cdot\frac{dist(v,\overline{H})}{dist(c_n,\overline{H})}\cdot(\theta - c_n^{n-1})\,,
\end{aligned}$$

and for further considerations we change the order of the factors and we can use a fourth factor by writing $(\theta - c_n^{n-1}) = \frac{(\theta - c_n^{n-1})}{\sqrt{1-h^2}} \cdot \sqrt{1-h^2}$, so

$$v^T(x_\Delta - x_{\Delta'}) = \frac{dist(v, \overline{H})}{dist(c_n, \overline{H})} \cdot \frac{(\theta - c_n^{n-1})}{\sqrt{1-h^2}} \cdot \sqrt{1-h^2} \cdot \frac{h - c_{n+1}^n}{c_{n+1}^{n-1} h - c_{n+1}^n \theta}. \tag{77}$$

The first factor (quotient) describes a relative position of v, responsible for the superiority of x_Δ together with the factor $(\theta - c_n^{n-1})$ which strengthens that effect. As long as (θ, h) are fixed, this explains in a certain sense the deviation from $conv(c_1, \dots, c_{n-1})$ in the $(n-1)$th coordinate. And the factor $\frac{h - c_{n+1}^n}{c_{n+1}^{n-1} h - c_{n+1}^n \theta}$ does not depend on the location of v at all, but it reflects the influence of the second facet with its augmenting point c_{n+1}. We call that the "kink-factor".

The Kink-Factor. Let us have a closer look at the factor $\frac{h - c_{n+1}^n}{c_{n+1}^{n-1} h - c_{n+1}^n \theta}$ and the analogous factors $\frac{h - c_k^n}{c_k^{n-1} h - c_k^n \theta}$ for the points c_k with $k = n+2, \dots, m$.
In the case where $c_k^{n-1} h - c_k^n \theta > 0$ (and the corresponding factor is positive), c_k may replace c_n in Δ and induce a new vertex $x_{\Delta_{n,k}}$, because $k \in K_{h,\theta}$. Here $\Delta_{n,k}$ denotes the index set $\Delta \setminus \{n\} \cup \{k\}$. We are going to characterize the boundary simplex condition of $conv(c_1, \dots, c_{n-1}, c_{n+1})$ this time in terms of the kink-angle $\varphi_{n+1} = \text{Min}\{\varphi_0, \varphi_{n+1}, \dots, \varphi_m\}$, where φ_k is defined as in (75).
Since $c_k^{n-1} = \theta + (h - c_k^n) \frac{\cos \varphi_k}{\sin \varphi_k}$ the kink-factors can be described as

$$\frac{h - c_k^n}{c_k^{n-1} h - c_k^n \theta} = \frac{\sin \varphi_k}{\theta \sin \varphi_k + h \cos \varphi_k} = \frac{\sin \varphi_k}{h_k},$$

where h_k is defined analogously to (76) as

$$h_k := \frac{1}{\|x_{\Delta_{n,k}}\|} = h \sin \varphi_k + \theta \cos \varphi_k \in (0,1) \quad (\text{for } k \in K_{h,\theta}).$$

So the description of the kink-factor above leads to the following estimation:

$$\frac{h - c_k^n}{c_k^{n-1} h - c_k^n \theta} = \frac{\sin \varphi_k}{h_k} \geq \sin \varphi_k \geq \frac{2}{\pi} \varphi_k. \tag{78}$$

The Minimal Difference. So far, we have studied the objective difference, when we replace c_n by another suitable point (c_{n+1}), which means that we leave x_Δ on the corresponding edge. But there are n different edges incident to x_Δ. Each of them results from replacing one of the generators of $conv(c_1, \dots, c_n)$ by another point c_k $(k \geq n+1)$ — if possible.
The lowest objective difference will then be achieved at one of these edges. Now we must generalize our notation:
Let c_i be the generator to be replaced and $c_{k(i)}$ with $k(i) \geq n+1$ the point replacing c_i, then the new basic index set is $\Delta_{i,k(i)} = \Delta \setminus \{i\} \cup \{k(i)\}$ and $x_{\Delta_{i,k(i)}}$ the corresponding vertex (if it exists). And we write U_i for the corresponding difference of the objective function, so

$$U_i = \begin{cases} v^T(x_\Delta - x_{\Delta_{i,k(i)}}) & \text{if a vertex } x_{\Delta_{i,k(i)}} \text{ exists,} \\ \infty & \text{if no such vertex exists.} \end{cases}$$

If a vertex $x_{\Delta_{i,k(i)}}$ exists then we have a formulation of $v^T(x_\Delta - x_{\Delta_{i,k(i)}})$ as a product analogously to (77). And the minimal objective difference results as $\mathrm{Min}_{i=1,\ldots,n} U_i$. W.l.o.g. we may study U_n and assume that $k(n) = n+1$. This will lead to an estimation for the probability that $\mathrm{Min}_{i=1,\ldots,n} U_i \leq \varepsilon \quad \forall \varepsilon \geq 0$.

3.2.2. The Distribution of the Objective Difference (Part 1)

An Exact Integral Quotient for the Distribution. In this chapter we try to analyze the three (resp. four) factors from a probabilistic point of view.
We work under the guaranteed assumption that $\|x_{opt}\| = \|x_\Delta\| \leq \frac{1}{q}$ for a fixed constant $q \in (0,1)$. This assumption may not be true for all constellations of the vectors a_i, $i = 1, \ldots, m$, and v but the probability that this condition is satisfied tends to one in the asymptotic case. Remark 3.2 will give explicit bounds on that probability. But it has a second useful implication. If we allow q to vary with the dimension pairs (m,n) – and regard it as fixed value only as long as we are in the same (m,n)-class – then we find a sequence of $q(m,n)$ values tending to 1 for $m \to \infty$ and n fixed such that even then the probability of our event tends to 1.

We start our considerations in the original configuration of the vectors a_i ($i = 1, \ldots, m$). Our interest is directed towards a conditional probability of the following kind. For this we want to derive an integral formula and we set $U = \mathrm{Min}_{i\in\Delta} U_i$ and consider $\varepsilon > 0$.

$$\begin{aligned} F_U(\varepsilon) &:= P(U \leq \varepsilon \mid x_\Delta \text{ is the optimal vertex} \wedge \|x_\Delta\| \leq \tfrac{1}{q}) \\ &= \frac{P(U \leq \varepsilon \wedge x_\Delta \text{ is the optimal vertex} \wedge \|x_\Delta\| \leq \frac{1}{q})}{P(x_\Delta \text{ is the optimal vertex} \wedge \|x_\Delta\| \leq \frac{1}{q})} \\ &= \frac{\displaystyle\int_{\mathbb{R}^n}^{(m+1)} I_{\{U\leq\varepsilon\}} I_{\{x_\Delta \text{ is the optimal vertex}\}} I_{\{\|x_\Delta\|\leq\frac{1}{q}\}} dF(v)dF(a_1)\cdots dF(a_m)}{\displaystyle\int_{\mathbb{R}^n}^{(m+1)} I_{\{x_\Delta \text{ is the optimal vertex}\}} I_{\{\|x_\Delta\|\leq\frac{1}{q}\}} dF(v)dF(a_1)\cdots dF(a_m)}. \end{aligned}$$

Note that $I_{\{x_\Delta \text{ is the optimal vertex}\}} = I_{\{x_\Delta \text{ is a vertex}\}} \cdot I_{\{v\in cone(a_i, i\in\Delta)\}}$. Therefore,

$$F_U(\varepsilon) = \frac{\displaystyle\int_{\mathbb{R}^n}^{(m+1)} I_{\{U\leq\varepsilon\}} I_{\{x_\Delta \text{ is a vertex}\}} I_{\{v\in cone(a_i,i\in\Delta)\}}}{\displaystyle\int_{\mathbb{R}^n}^{(m+1)} I_{\{x_\Delta \text{ is a vertex}\}} I_{\{v\in cone(a_i,i\in\Delta)\}}}$$

$$\frac{I_{\{\|x_\Delta\|\leq\frac{1}{q}\}} dF(v)dF(a_1)\cdots dF(a_m)}{I_{\{\|x_\Delta\|\leq\frac{1}{q}\}} dF(v)dF(a_1)\cdots dF(a_m)}.$$

Again, w.l.o.g. it is feasible to choose $\Delta = \{1, \dots, n\}$. Our first coordinate transformation $a_i \longrightarrow b_i$, $b_1^n = \dots = b_n^n = h$ delivers

$$F_U(\varepsilon) = \frac{\int\limits_0^1 \int\limits_{\mathbb{R}^{n-1}} \int\limits_{\mathbb{R}^n}^{(n)} \int\limits_{\mathbb{R}^n}^{(m-n)} I_{\{h \geq q\}} \prod_{k=n+1}^{m} I_{\{b_k^n \leq h\}} I_{\{U \leq \varepsilon\}} I_{\{v \in cone(b_i, i \in \Delta)\}}}{\int\limits_0^1 \int\limits_{\mathbb{R}^{n-1}}^{(n)} \int\limits_{\mathbb{R}^n} G(h)^{m-n} I_{\{h \geq q\}} I_{\{v \in cone(b_i, i \in \Delta)\}} f(v) dv}$$

$$\frac{f(v)dv \prod_{k=n+1}^{m} f(b_k) db_{n+1} \cdots db_m |\operatorname{Det}(B)| \prod_{k=1}^{n} f(b_k) d\bar{b}_1 \cdots d\bar{b}_n dh}{|\operatorname{Det}(B)| \prod_{k=1}^{n} f(b_k) d\bar{b}_1 \cdots d\bar{b}_n dh},$$

where B is the same matrix as in (53) and $\lambda_{n-1}(\omega_n) \operatorname{Det}(B)$ is the Jacobian of this first transformation and the marginal distribution function G was introduced in (48). In the second transformation we replace the b_i's by c_i's such that $c_1^{n-1} = \dots = c_{n-1}^{n-1} = \theta$ and $c_n^{n-1} < \theta$ and $c_1^n = \dots = c_n^n = h$.
The Jacobian of that transformation is $\lambda_{n-2}(\omega_{n-1}) \operatorname{Det}(\overline{C})$, where

$$\overline{C} := \begin{pmatrix} c_1^1 & \cdots & c_1^{n-2} & 1 \\ \vdots & & \vdots & \vdots \\ c_{n-1}^1 & \cdots & c_{n-1}^{n-2} & 1 \end{pmatrix}.$$

The matrix B now turns into a matrix $C = \begin{pmatrix} c_1^1 & \cdots & c_1^{n-2} & \theta & 1 \\ \vdots & & \vdots & \vdots & \vdots \\ c_n^1 & \cdots & c_n^{n-2} & \theta & 1 \end{pmatrix}$.

And we know that

$$\begin{aligned} |\operatorname{Det}(B)| &= |\operatorname{Det}(C)| = \lambda_{n-1}(conv(c_1, \dots, c_n))(n-1)! \\ &= |\theta - c_n^{n-1}| \lambda_{n-2}(conv(c_1, \dots, c_{n-1}))(n-2)! \\ |\operatorname{Det}(\overline{C})| &= \lambda_{n-2}(conv(c_1, \dots, c_{n-1}))(n-2)! \\ \Rightarrow \quad & |\operatorname{Det}(\overline{C})| \, |\operatorname{Det}(C)| = \\ & \quad = |\theta - c_n^{n-1}| \Big(\lambda_{n-2}(conv(c_1, \dots, c_{n-1}))\Big)^2 \Big((n-2)!\Big)^2. \end{aligned}$$

After changing the order of integrations (Fubini), we arrive at

$$F_U(\varepsilon) = \frac{\int\limits_q^1 \int\limits_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int\limits_{\mathbb{R}^{n-1}} |\theta - c_n^{n-1}| I_{\{c_n^{n-1} \leq \theta\}} \int\limits_{\mathbb{R}^n}^{(m-n)} \prod_{k=n+1}^{m} I_{\{c_k^n \leq h\}}}{\int\limits_q^1 G(h)^{m-n} \int\limits_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int\limits_{\mathbb{R}^{n-1}} |\theta - c_n^{n-1}| I_{\{c_n^{n-1} \leq \theta\}}}$$

$$\frac{\int\limits_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(c_1,\dots,c_{n-1}))^2 \int\limits_{\mathbb{R}^n} I_{\{U\le\varepsilon\}} I_{\{v\in cone(c_1,\dots,c_n)\}} f(v)dv}{\int\limits_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(c_1,\dots,c_{n-1}))^2 V(c_1,\dots,c_n)}$$

$$\frac{\prod_{i=1}^{n-1} f(c_i)d\bar{\bar{c}}_1\cdots d\bar{\bar{c}}_{n-1} \prod_{k=n+1}^{m} f(c_k)dc_{n+1}\cdots dc_m f(c_n)d\bar{c}_n d\theta dh}{\prod_{i=1}^{n-1} f(c_i)d\bar{\bar{c}}_1\cdots d\bar{\bar{c}}_{n-1} f(c_n)d\bar{c}_n d\theta dh},$$

where we use $V(c_1,\dots,c_n)$ for

$$\frac{\lambda_{n-1}(cone(c_1,\dots,c_n)\cap\omega_n)}{\lambda_{n-1}(\omega_n)} = \int\limits_{\mathbb{R}^n} I_{\{v\in cone(c_1,\dots,c_n)\}} f(v)dv\,.$$

This is the final stage where we are able to give an exact formula for $F_U(\varepsilon)$. In the following we shall deal with approximations of that distribution function resp. probability.

An Estimation of F_U in Four Separate Factors. Let us estimate $F_U(\varepsilon)$ for arbitrary values $\varepsilon\in(0,1]$.

Our first observation is that $\{U\le\varepsilon\} = \bigcup_{i=1}^n\{U_i\le\varepsilon\}$ and therefore for the indicator functions we have $I_{\{U\le\varepsilon\}} \le \sum_{i=1}^n I_{\{U_i\le\varepsilon\}}$. The rotation symmetry ascertains that the events $\{U_i\le\varepsilon\}$ are distributed identically and hence

$$\int\limits_{\mathbb{R}^n} I_{\{U\le\varepsilon\}} I_{\{v\in cone(c_1,\dots,c_n)\}} f(v)dv \le n \int\limits_{\mathbb{R}^n} I_{\{U_n\le\varepsilon\}} I_{\{v\in cone(c_1,\dots,c_n)\}} f(v)dv\,.$$

So we are allowed to concentrate on U_n (and let c_{n+1} be the replacing vector). Remember that

$$\{U_n\le\varepsilon\} \Leftrightarrow \left\{\frac{dist(v,\overline{H})}{dist(c_n,\overline{H})}\cdot\frac{(\theta-c_n^{n-1})}{\sqrt{1-h^2}}\cdot\sqrt{1-h^2}\cdot\frac{h-c_{n+1}^n}{c_{n+1}^{n-1}h-c_{n+1}^n\theta}\le\varepsilon\right\}$$

$$\Leftrightarrow \left\{\frac{dist(v,\overline{H})}{dist(c_n,\overline{H})}\cdot\frac{(\theta-c_n^{n-1})}{\sqrt{1-h^2}}\cdot\sqrt{1-h^2}\cdot\frac{\sin\varphi_{n+1}}{\theta\sin\varphi_{n+1}+h\cos\varphi_{n+1}}\le\varepsilon\right\}.$$

Now the validity of $\{U_n\le\varepsilon\}$ depends on four factors. And it is clear that for any choice of $\alpha_1,\alpha_2,\alpha_3,\alpha_4\ge 0$ such that $\alpha_1+\alpha_2+\alpha_3+\alpha_4\le 1$ we have the following inequality:

$$\begin{aligned} I_{\{U_n\le\varepsilon\}} \;\le\; & I_{\{\frac{dist(v,\overline{H})}{dist(c_n,\overline{H})}\le\varepsilon^{\alpha_1}\}} + I_{\{\frac{(\theta-c_n^{n-1})}{\sqrt{1-h^2}}\le\varepsilon^{\alpha_2}\}} \\ & + I_{\{\sqrt{1-h^2}\le\varepsilon^{\alpha_3}\}} + I_{\{\frac{\sin\varphi_{n+1}}{\theta\sin\varphi_{n+1}+h\cos\varphi_{n+1}}\le\varepsilon^{\alpha_4}\}}\,. \end{aligned}$$

This inequality is immediate: if all events on the right side happen to fail, then the event on the left side happens to fail, too.

In the view of the role of U_n corresponding to U it is clear that

$$F_U(\varepsilon) \leq n\,\frac{\mathcal{I}_1}{\mathcal{I}_0} + n\,\frac{\mathcal{I}_2}{\mathcal{I}_0} + n\,\frac{\mathcal{I}_3}{\mathcal{I}_0} + n\,\frac{\mathcal{I}_4}{\mathcal{I}_0}\,. \tag{79}$$

$$\text{Here,}\quad \mathcal{I}_0 := \int_q^1 G(h)^{m-n} \int_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int_{\mathbb{R}^{n-1}} |\theta - c_n^{n-1}| I_{\{c_n^{n-1} \leq \theta\}} \int_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(c_1,\ldots,c_{n-1}))^2 V(c_1,\ldots,c_n) \prod_{i=1}^{n-1} f(c_i) d\bar{\bar{c}}_1 \cdots d\bar{\bar{c}}_{n-1} f(c_n) d\bar{c}_n d\theta dh\,,$$

recall that $V(c_1,\ldots,c_n) = \frac{1}{\lambda_{n-1}(\omega_n)} \int\limits_{cone(c_1,\ldots,c_n)\cap\omega_n} 1\,dv\,.$

$$\mathcal{I}_1 := \int_q^1 G(h)^{m-n} \int_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int_{\mathbb{R}^{n-1}} |\theta - c_n^{n-1}| I_{\{c_n^{n-1} \leq \theta\}} \int_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(c_1,\ldots,c_{n-1}))^2 \int\limits_{v \in cone(c_1,\ldots,c_n)\cap\omega_n} I_{\{\frac{dist(v,\overline{H})}{dist(c_n,\overline{H})} \leq \varepsilon^{\alpha_1}\}} dv \prod_{i=1}^{n-1} f(c_i) d\bar{\bar{c}}_1 \cdots d\bar{\bar{c}}_{n-1} f(c_n) d\bar{c}_n d\theta dh\,,$$

$$\mathcal{I}_2 := \int_q^1 G(h)^{m-n} \int_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int_{\mathbb{R}^{n-1}} |\theta - c_n^{n-1}| I_{\{c_n^{n-1} \leq \theta\}} \int_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(c_1,\ldots,c_{n-1}))^2 V(c_1,\ldots,c_n) I_{\{\frac{(\theta - c_n^{n-1})}{\sqrt{1-h^2}} \leq \varepsilon^{\alpha_2}\}} \prod_{i=1}^{n-1} f(c_i) d\bar{\bar{c}}_1 \cdots d\bar{\bar{c}}_{n-1} f(c_n) d\bar{c}_n d\theta dh\,,$$

$$\mathcal{I}_3 := \int_q^1 G(h)^{m-n} I_{\{\sqrt{1-h^2} \leq \varepsilon^{\alpha_3}\}} \int_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int_{\mathbb{R}^{n-1}} |\theta - c_n^{n-1}| I_{\{c_n^{n-1} \leq \theta\}} \int_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(c_1,\ldots,c_{n-1}))^2 V(c_1,\ldots,c_n)$$

$$\prod_{i=1}^{n-1} f(c_i) d\bar{\bar{c}}_1 \cdots d\bar{\bar{c}}_{n-1} f(c_n) d\bar{c}_n d\theta dh \,,$$

$$\mathcal{I}_4 := \int_q^1 \int_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int_{\mathbb{R}^{n-1}} |\theta - c_n^{n-1}| I_{\{c_n^{n-1} \leq \theta\}}$$

$$\int_{\mathbb{R}^n}^{(m-n)} \prod_{k=n+1}^{m} I_{\{c_k^n \leq h\}} I_{\{K_{h,\theta} \neq \emptyset \text{ and } \mathrm{Min}_{k \in K_{h,\theta}} \frac{\sin \varphi_k}{\theta \sin \varphi_k + h \cos \varphi_k} \leq \varepsilon^{\alpha_4}\}}$$

$$\int_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(c_1, \ldots, c_{n-1}))^2 V(c_1, \ldots, c_n)$$

$$\prod_{i=1}^{n-1} f(c_i) d\bar{\bar{c}}_1 \cdots d\bar{\bar{c}}_{n-1} \prod_{k=n+1}^{m} f(c_k) dc_{n+1} \cdots dc_m f(c_n) d\bar{c}_n d\theta dh \,.$$

Remember that

$$I_{\{K_{h,\theta} \neq \emptyset \text{ and } \mathrm{Min}_{k \in K_{h,\theta}} \frac{\sin \varphi_k}{\theta \sin \varphi_k + h \cos \varphi_k} \leq \varepsilon^{\alpha_4}\}} = I_{\{\frac{\sin \varphi_{n+1}}{\theta \sin \varphi_{n+1} + h \cos \varphi_{n+1}} \leq \varepsilon^{\alpha_4}\}}$$

as we have assumed that $k(n) = n+1$ and that U_n is the minimal objective difference.

3.2.3. *The Four Quotients*

In this section we will derive approximations for the distributions of the four different quotients. As this requires various arguments from integration and measure theory, the section will be quite technical. So, the reader may skip this part and go on with section 3.2.4, where the results of all the approximations are summarized.

The Relative Distance. To study the distribution of the distance of v to $\overline{H} = H(0, c_1, \ldots, c_{n-1})$ under the condition that $\mathbb{R}^+ v \cap conv(c_1, \ldots, c_n) \neq \emptyset$, we can employ results from [7] about the relation between the spherical measures $V(c_1, \ldots, c_n)$ and $W(c_1, \ldots, c_{n-1}) := \frac{\lambda_{n-2}(cone(c_1,\ldots,c_{n-1}) \cap \omega_n)}{\lambda_{n-2}(\omega_{n-1})}$. Both figures are spherical measures of the cones spanned by the corresponding set of vectors.
If z is defined as the normal vector on $\overline{H}$ (positively directed towards c_n), then for any $w \in \omega_n \cap \overline{H}$ and $c_n \in \omega_n$ it is known that

$$\begin{aligned} V(c_1, \ldots, c_n) \lambda_{n-1}(\omega_n) &= \lambda_{n-1}(cone(c_1, \ldots, c_n) \cap \omega_n) \\ &= \int_{cone(c_1,\ldots,c_n) \cap \omega_n} \Upsilon(c_n, z, w) \lambda_{n-2}(d_{n-1} w) \end{aligned}$$

with $\Upsilon(c_n, z, w) = \dfrac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \dfrac{dist(c_n, \overline{H})}{(1 - (c_n^T w)^2)^{\frac{n-1}{2}}} \displaystyle\int_{c_n^T w}^1 \sqrt{1-h^2}^{\,n-3} dh \,.$
When we partition the ground area $cone(c_1, \ldots, c_{n-1}) \cap \omega_n$ into small sets $M(w_i)$, where $M(w_i)$ denotes a neighbourhood of w_i, we induce a corresponding partition of $cone(c_1, \ldots, c_n) \cap \omega_n$ into small spherical sets $cone(M(w_i), c_n) \cap \omega_n$.

But to the latter set our formula can be applied, too.

Ignoring all the constants including the position of c_n, $\overline{H}$ and w, the varying influence of the spherical effect is described in the last integral.
This gives a chance to describe a sufficiently good internal distribution of points on the "stripe" $cone(M_i(w), c_n) \cap \omega_n$.
First we note that $\int\limits_{c_n^T w}^{1} \sqrt{1-h^2}^{\,n-3} dh = \int\limits_{0}^{\beta} \sin\gamma^{n-2} d\gamma$, where $\beta = \measuredangle(c_n, w)$ resp. $\cos(c_n, w) = c_n^T w$ and $h = \cos\gamma$.

This reflects the movement on a geodetic circle on the ball from w, $(\eta = 0)$ to c_n, $(\eta = \beta)$. If we move on such a circle induced by w and c_n, we increase the distance to $\overline{H}$ until an angle of $\frac{\pi}{2}$ is traversed. Afterwards we decrease the distance symmetrically.

We are interested in the share of those points on our spherical stripe, whose distance to $\overline{H}$ is below a certain proportion of $dist(c_n, \overline{H})$, e.g. less than $\tau \cdot dist(c_n, \overline{H})$ for a $\tau \in (0,1)$.
Since in the case $\beta < \frac{\pi}{2}$ the maximal distance on our move is $dist(c_n, \overline{H})$, and since for $\beta > \frac{\pi}{2}$ it is even higher we should distinguish these two cases and treat them separately.

1. Case $\beta < \frac{\pi}{2}$:
If we move away from w by an angle $\arcsin(\tau \sin\beta)$, then we are exactly at the borderline, where we leave the region with distance less than $\tau\, dist(c_n, \overline{H})$. Afterwards, the distance still increases up to the angle β and remains above the level in question.

So we are interested in the following proportion:

$$P(dist(x, \overline{H}) \le \tau\, dist(c_n, \overline{H})) = \frac{\int\limits_{0}^{\arcsin(\tau\sin\beta)} \sin(\beta-\eta)^{n-2} d\eta}{\int\limits_{0}^{\beta} \sin(\beta-\eta)^{n-2} d\eta} .$$

It is well known that $\sin(\tau\beta) \ge \tau \sin\beta$ for $\tau < 1$, hence
$\tau\beta = \arcsin(\sin(\tau\beta)) \ge \arcsin(\tau\sin\beta)$, which implies

$$P(dist(x, \overline{H}) \le \tau\, dist(c_n, \overline{H})) \le \frac{\int\limits_{0}^{\tau\beta} \sin(\beta-\eta)^{n-2} d\eta}{\int\limits_{0}^{\beta} \sin(\beta-\eta)^{n-2} d\eta} .$$

If we replace the terms $\sin(\beta-\eta)$ by $(\beta-\eta)$ we multiply with a factor $\frac{\beta-\eta}{\sin(\beta-\eta)}$, which is an increasing function of $(\beta-\eta)$ for $\beta-\eta \in (0, \frac{\pi}{2})$ and a decreasing function of η for $\eta \in (0,\beta)$. Hence such a transformation would strengthen the critical region ($\eta \le \tau\beta$). Therefore

$$\begin{aligned} P(dist(x,\overline{H}) \le \tau\, dist(c_n,\overline{H})) &\le \frac{\int_0^{\tau\beta} (\beta-\eta)^{n-2} d\eta}{\int_0^{\beta} (\beta-\eta)^{n-2} d\eta} = \frac{\frac{-1}{n-1}(\beta-\eta)^{n-1}\Big|_0^{\tau\beta}}{\frac{-1}{n-1}(\beta-\eta)^{n-1}\Big|_0^{\beta}} \\ &= \frac{\beta^{n-1} - ((1-\tau)\beta)^{n-1}}{\beta^{n-1}} = 1-(1-\tau)^{n-1} \end{aligned}$$

2. Case $\beta > \frac{\pi}{2}$:
Exactly at $\arcsin(\tau \sin\beta)$ we leave the critical region, then the distance increases until we have an angle $\frac{\pi}{2}$ and after that it returns down to the level $dist(c_n,\overline{H})$ at β.

But $\sin\beta = \sin(\pi-\beta) = \sin\overline{\beta}$ with $\overline{\beta} = \pi - \beta$. So $\arcsin(\sin\beta) = \overline{\beta} < \frac{\pi}{2}$, and it is clear that $\arcsin(\tau\sin\overline{\beta}) \le \tau\overline{\beta}$. Our proportion is

$$\begin{aligned} &P(dist(x,\overline{H}) \le \tau\, dist(c_n,\overline{H})) = \\ &= \frac{\int_0^{\arcsin(\tau\sin\beta)} \sin(\beta-\eta)^{n-2} d\eta}{\int_0^{\beta} \sin(\beta-\eta)^{n-2} d\eta} \le \frac{\int_0^{\tau\overline{\beta}} \sin(\beta-\eta)^{n-2} d\eta}{\int_0^{\beta} \sin(\beta-\eta)^{n-2} d\eta}. \end{aligned}$$

Again, we replace $\sin(\beta-\eta)$ by $(\beta-\eta)$, which strengthens the role of the interval $[0,\tau\sin\overline{\beta}]$. So we have

$$\begin{aligned} P(dist(x,\overline{H}) \le \tau\, dist(c_n,\overline{H})) &\le \frac{\int_0^{\tau\overline{\beta}} (\beta-\eta)^{n-2} d\eta}{\int_0^{\beta} (\beta-\eta)^{n-2} d\eta} = \frac{(\beta^{n-1} - (\beta-\tau\overline{\beta})^{n-1})}{\beta^{n-1}} \\ &\le \frac{\beta^{n-1} - (\beta-\tau\beta)^{n-1}}{\beta^{n-1}} = 1-(1-\tau)^{n-1}. \end{aligned}$$

This relation is true for every stripe of the partition of the spherical simplex $cone(c_1,\ldots,c_n)\cap\omega_n$, so it must be true for the total set, too.

We are allowed to exploit this insight when we evaluate the first quotient $\frac{\mathcal{I}_1}{\mathcal{I}_0}$. This makes the estimation very simple and it means handling the divergence of the two innermost integrals

$$\frac{\int_{v\in cone(c_1,\ldots,c_n)\cap\omega_n} I_{\{\frac{dist(v,\overline{H})}{dist(c_n,\overline{H})} \le \varepsilon^{\alpha_1}\}} dv}{\int_{v\in cone(c_1,\ldots,c_n)\cap\omega_n} 1 dv} = 1-(1-\varepsilon^{\alpha_1})^{n-1}.$$

We conclude that

$$\frac{\mathcal{I}_1}{\mathcal{I}_0} \le 1-(1-\varepsilon^{\alpha_1})^{n-1} \le (n-1)\varepsilon^{\alpha_1}. \tag{80}$$

The Relative Extension of the Simplex. Now we want to estimate the quotient $\frac{\mathcal{I}_2}{\mathcal{I}_0}$. The only difference between the numerator and denominator lies in the indicator $I_{\{\frac{(\theta-c_n^{n-1})}{\sqrt{1-h^2}}\leq\varepsilon^{\alpha_2}\}}$ which appears in $\mathcal{I}_2$ but not in $\mathcal{I}_0$.

We should discuss some inner integrals in $\mathcal{I}_2$ resp. $\mathcal{I}_0$ and derive upper resp. lower bounds. First, look at

$$\int\limits_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(c_1,\ldots,c_{n-1}))^2 V(c_1,\ldots,c_n)\prod_{i=1}^{n-1} f(c_i)d\bar{\bar{c}}_1\cdots d\bar{\bar{c}}_{n-1}$$

which appears in $\mathcal{I}_2$ and in $\mathcal{I}_0$ as well.
An upper bound results from

$$\begin{aligned}
V(c_1,\ldots,c_n) &= \frac{\lambda_n(cone(c_1,\ldots,c_n)\cap\Omega_n)}{\lambda_n(\Omega_n)} \leq \frac{\lambda_n(conv(0,\frac{1}{h}c_1,\ldots,\frac{1}{h}c_n))}{\lambda_n(\Omega_n)} \\
&= \frac{1}{h^n}\frac{\lambda_n(conv(0,c_1,\ldots,c_n))}{\lambda_n(\Omega_n)} = \frac{1}{h^n}\cdot\frac{h}{n}\cdot\frac{\lambda_{n-1}(conv(c_1,\ldots,c_n))}{\lambda_n(\Omega_n)} \\
&= \frac{|\theta-c_n^{n-1}|}{h^{n-1}n(n-1)}\cdot\frac{\lambda_{n-2}(conv(c_1,\ldots,c_{n-1}))}{\lambda_n(\Omega_n)}.
\end{aligned}$$

So it is clear that the integral above is bounded from above by

$$\frac{|\theta-c_n^{n-1}|}{h^{n-1}n(n-1)\lambda_n(\Omega_n)} \cdot \int\limits_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(c_1,\ldots,c_{n-1}))^3\prod_{i=1}^{n-1} f(c_i)d\bar{\bar{c}}_1\cdots d\bar{\bar{c}}_{n-1}.$$

And a lower bound results from

$$\begin{aligned}
V(c_1,\ldots,c_n) &\geq \frac{\lambda_n(conv(0,c_1,\ldots,c_n)\cap\Omega_n)}{\lambda_n(\Omega_n)} \\
&= \frac{h|\theta-c_n^{n-1}|}{n(n-1)}\cdot\frac{\lambda_{n-2}(conv(c_1,\ldots,c_{n-1}))}{\lambda_n(\Omega_n)},
\end{aligned}$$

hence we have a lower bound for the inner integral of

$$\frac{h|\theta-c_n^{n-1}|}{n(n-1)\lambda_n(\Omega_n)}\cdot\int\limits_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(c_1,\ldots,c_{n-1}))^3\prod_{i=1}^{n-1} f(c_i)d\bar{\bar{c}}_1\cdots d\bar{\bar{c}}_{n-1}.$$

Our bounds had produced a "pre-factor" of $\frac{1}{h^n}$ and the second quotient can be bounded from above in the following way (after reducing the fraction by the factor

$\frac{1}{n(n-1)\lambda_n(\Omega_n)}$):

$$\frac{\mathcal{I}_2}{\mathcal{I}_0} \leq \frac{\int_q^1 G(h)^{m-n} \int_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int_{\mathbb{R}^{n-1}} I_{\{c_n^{n-1}\leq\theta\}}|\theta - c_n^{n-1}| I_{\{\frac{(\theta-c_n^{n-1})}{\sqrt{1-h^2}}\leq\varepsilon^{\alpha_2}\}} \frac{|\theta-c_n^{n-1}|}{h^{n-1}}}{\int_q^1 G(h)^{m-n} \int_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int_{\mathbb{R}^{n-1}} I_{\{c_n^{n-1}\leq\theta\}}|\theta - c_n^{n-1}| \cdot h|\theta - c_n^{n-1}|}$$

$$\frac{\int_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(c_1,\ldots,c_{n-1}))^3 \prod_{i=1}^{n-1} f(c_i)d\bar{\bar{c}}_1\cdots d\bar{\bar{c}}_{n-1}f(c_n)d\bar{c}_n d\theta dh}{\int_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(c_1,\ldots,c_{n-1}))^3 \prod_{i=1}^{n-1} f(c_i)d\bar{\bar{c}}_1\cdots d\bar{\bar{c}}_{n-1}f(c_n)d\bar{c}_n d\theta dh}$$

$$= \frac{\int_q^1 G(h)^{m-n}\frac{1}{h^{n-1}} \int_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int_{\mathbb{R}^{n-1}} I_{\{c_n^{n-1}\leq\theta\}} I_{\{\frac{(\theta-c_n^{n-1})}{\sqrt{1-h^2}}\leq\varepsilon^{\alpha_2}\}} \cdot |\theta - c_n^{n-1}|^2}{\int_q^1 G(h)^{m-n}h \int_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int_{\mathbb{R}^{n-1}} I_{\{c_n^{n-1}\leq\theta\}}|\theta - c_n^{n-1}|^2}$$

$$\frac{(1-h^2-\theta^2)^{\frac{1}{2}(3(n-2)+(n-1)(n-4))} \int_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(\frac{\bar{\bar{c}}_1}{||\bar{\bar{c}}_1||},\ldots,\frac{\bar{\bar{c}}_{n-1}}{||\bar{\bar{c}}_{n-1}||}))^3}{(1-h^2-\theta^2)^{\frac{1}{2}(3(n-2)+(n-1)(n-4))} \int_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(\frac{\bar{\bar{c}}_1}{||\bar{\bar{c}}_1||},\ldots,\frac{\bar{\bar{c}}_{n-1}}{||\bar{\bar{c}}_{n-1}||}))^3}$$

$$\frac{\prod_{i=1}^{n-1} \bar{f}(\bar{\bar{c}}_i|\theta,h)d\bar{\bar{c}}_1\cdots d\bar{\bar{c}}_{n-1}f(c_n)d\bar{c}_n d\theta dh}{\prod_{i=1}^{n-1} \bar{f}(\bar{\bar{c}}_i|\theta,h)d\bar{\bar{c}}_1\cdots d\bar{\bar{c}}_{n-1}f(c_n)d\bar{c}_n d\theta dh}.$$

The identity of these quotients of integrals can be explained as follows:
When h and θ are fixed, then $\bar{\bar{c}}_1,\ldots,\bar{\bar{c}}_{n-1}$ are positioned on a sphere of radius $\sqrt{1-h^2-\theta^2}$ and they are uniformly distributed, too (and independently of h,θ). The density function $f(c_i)$ can then be factorized by the two-dimensional marginal density $f(\theta,h) = \frac{\lambda_{n-3}(\omega_{n-2})}{\lambda_{n-1}(\omega_n)}(1-h^2-\theta^2)^{\frac{1}{2}(n-4)}$ and by the conditional densities $\bar{f}(\bar{\bar{c}}_i|\theta,h)$ in the form $f(c_i) = f(\theta,h)\bar{f}(\bar{\bar{c}}_i|\theta,h)$.
The insertion of this yields the above quotient.
The analogous factorization of the density $f(c_n)$ and

$$\begin{aligned}\lambda_{n-2}(conv(c_1,\ldots,c_{n-1})) &= \lambda_{n-2}(\,conv(\bar{\bar{c}}_1,\ldots,\bar{\bar{c}}_{n-1}))\\ &= (1-h^2-\theta^2)^{\frac{1}{2}(n-2)}\lambda_{n-2}(conv(\frac{\bar{\bar{c}}_1}{||\bar{\bar{c}}_1||},\ldots,\frac{\bar{\bar{c}}_{n-1}}{||\bar{\bar{c}}_{n-1}||}))\end{aligned}$$

makes a further simplification of the quotient possible.
(The simplification is the reduction of the quotient by the factors depending on n only and by the factor
$\int_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(\frac{\bar{\bar{c}}_1}{||\bar{\bar{c}}_1||},\ldots,\frac{\bar{\bar{c}}_{n-1}}{||\bar{\bar{c}}_{n-1}||}))^3 \prod_{i=1}^{n-1} \bar{f}(\bar{\bar{c}}_i|\theta,h)d\bar{\bar{c}}_1\cdots d\bar{\bar{c}}_{n-1}$.
Note, that this factor resp. integral does not depend on θ and h any longer.)

So we obtain

$$\frac{\int\limits_q^1 G(h)^{m-n}\frac{1}{h^{n-1}} \int\limits_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int\limits_{-\sqrt{1-h^2}}^{\theta} I_{\{\frac{(\theta-c_n^{n-1})}{\sqrt{1-h^2}}\leq\varepsilon^{\alpha_2}\}} \cdot |\theta - c_n^{n-1}|^2}{\int\limits_q^1 G(h)^{m-n} h \int\limits_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int\limits_{-\sqrt{1-h^2}}^{\theta} |\theta - c_n^{n-1}|^2}$$

$$\frac{(1-h^2-\theta^2)^{\frac{1}{2}(3(n-2)+(n-1)(n-4))}(1-h^2-(c_n^{n-1})^2)^{\frac{1}{2}(n-4)}dc_n^{n-1}d\theta dh}{(1-h^2-\theta^2)^{\frac{1}{2}(3(n-2)+(n-1)(n-4))}(1-h^2-(c_n^{n-1})^2)^{\frac{1}{2}(n-4)}dc_n^{n-1}d\theta dh},$$

and for further considerations it is recommended to perform a substitution

$$\xi := \frac{c_n^{n-1}}{\sqrt{1-h^2}} \quad \text{resp.} \quad c_n^{n-1} = \xi\sqrt{1-h^2} \quad \text{and} \quad \frac{dc_n^{n-1}}{d\xi} = \sqrt{1-h^2},$$
$$\zeta := \frac{\theta}{\sqrt{1-h^2}} \quad \text{resp.} \quad \theta = \zeta\sqrt{1-h^2} \quad \text{and} \quad \frac{d\theta}{d\zeta} = \sqrt{1-h^2}.$$

What we obtain is of the form

$$\frac{\int\limits_q^1 G(h)^{m-n}\frac{1}{h^{n-1}}(1-h^2)^{\frac{1}{2}(n^2-n-2)} \int\limits_{-1}^{1} (1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)}}{\int\limits_q^1 G(h)^{m-n} h (1-h^2)^{\frac{1}{2}(n^2-n-2)} \int\limits_{-1}^{1} (1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)}}$$

$$\frac{\int\limits_{-1}^{\zeta} I_{\{(\zeta-\xi)\leq\varepsilon^{\alpha_2}\}} \cdot |\zeta-\xi|^2(1-\xi^2)^{\frac{1}{2}(n-4)}d\xi d\zeta dh}{\int\limits_{-1}^{\zeta} |\zeta-\xi|^2(1-\xi^2)^{\frac{1}{2}(n-4)}d\xi d\zeta dh}$$

and this is identical to

$$\frac{\int\limits_q^1 G(h)^{m-n}\frac{1}{h^{n-1}}(1-h^2)^{\frac{1}{2}(n^2-n-2)}dh}{\int\limits_q^1 G(h)^{m-n} h (1-h^2)^{\frac{1}{2}(n^2-n-2)}dh}$$

$$\cdot \frac{\int\limits_{-1}^{1} (1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)} \int\limits_{-1}^{\zeta} I_{\{(\zeta-\xi)\leq\varepsilon^{\alpha_2}\}} \cdot |\zeta-\xi|^2(1-\xi^2)^{\frac{1}{2}(n-4)}d\xi d\zeta}{\int\limits_{-1}^{1} (1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)} \int\limits_{-1}^{\zeta} |\zeta-\xi|^2(1-\xi^2)^{\frac{1}{2}(n-4)}d\xi d\zeta},$$

because the inner integrals do not depend on h any longer. But this is the same quotient as (because of symmetry)

$$\frac{\int_q^1 G(h)^{m-n}\frac{1}{h^{n-1}}(1-h^2)^{\frac{1}{2}(n^2-n-2)}dh}{\int_q^1 G(h)^{m-n}h(1-h^2)^{\frac{1}{2}(n^2-n-2)}dh}$$

$$\frac{\int_0^1(1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)}\int_{-1}^1 I_{\{|\zeta-\xi|\le\varepsilon^{\alpha_2}\}}\cdot|\zeta-\xi|^2(1-\xi^2)^{\frac{1}{2}(n-4)}d\xi d\zeta}{\int_0^1(1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)}\int_{-1}^1|\zeta-\xi|^2(1-\xi^2)^{\frac{1}{2}(n-4)}d\xi d\zeta}.$$

In the numerator we will estimate $1-\xi^2\le 1$, and in the denominator we shall calculate $|\zeta-\xi|^2=\zeta^2-2\zeta\xi+\xi^2$ and recognize that the mixed term $-2\zeta\xi$ is redundant for integration as a result of odd symmetry. So we have an upper bound of

$$\frac{\int_q^1 G(h)^{m-n}\frac{1}{h^{n-1}}(1-h^2)^{\frac{1}{2}(n^2-n-2)}dh}{\int_q^1 G(h)^{m-n}h(1-h^2)^{\frac{1}{2}(n^2-n-2)}dh} \tag{81}$$

$$\frac{\int_0^1(1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)}\int_{\zeta-\varepsilon^{\alpha_2}}^{\zeta+\varepsilon^{\alpha_2}}|\zeta-\xi|^2d\xi d\zeta}{\int_0^1(1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)}\left(\zeta^2\int_{-1}^1(1-\xi^2)^{\frac{1}{2}(n-4)}d\xi+\int_{-1}^1\xi^2(1-\xi^2)^{\frac{1}{2}(n-4)}d\xi\right)d\zeta}.$$

For the inner integral of the numerator we use $\int_{-\varepsilon^{\alpha_2}}^{\varepsilon^{\alpha_2}}\kappa^2d\kappa=\frac{2}{3}\varepsilon^{3\alpha_2}$ and the two inner integrals in the denominator are approximated in the following way:

$$\begin{aligned}
&\int_0^1(1-\zeta^2)^{\frac{(n^2-2n-2)}{2}}\left(\zeta^2\int_{-1}^1(1-\xi^2)^{\frac{(n-4)}{2}}d\xi+\int_{-1}^1\xi^2(1-\xi^2)^{\frac{(n-4)}{2}}d\xi\right)d\zeta\\
&=\int_0^1(1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)}\left(\zeta^2\frac{\Gamma(\frac{1}{2})\Gamma(\frac{(n-4)}{2}+1)}{\Gamma(\frac{(n-4)}{2}+\frac{3}{2})}+\frac{\Gamma(\frac{3}{2})\Gamma(\frac{(n-4)}{2}+1)}{\Gamma(\frac{(n-4)}{2}+\frac{5}{2})}\right)d\zeta\\
&=\int_0^1(1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)}d\zeta\\
&\quad\cdot\left(\frac{\Gamma(\frac{3}{2})\Gamma(\frac{1}{2}(n^2-2n-2)+\frac{3}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2}(n^2-2n-2)+\frac{5}{2})}\frac{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2}n-1)}{\Gamma(\frac{1}{2}n-\frac{1}{2})}+\frac{\Gamma(\frac{3}{2})\Gamma(\frac{1}{2}n-1)}{\Gamma(\frac{1}{2}n+\frac{1}{2})}\right)\\
&=\int_0^1(1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)}d\zeta\cdot\frac{\sqrt{\pi}}{2}\left(\frac{\Gamma(\frac{n^2}{2}-n+\frac{1}{2})\Gamma(\frac{n}{2}-1)}{\Gamma(\frac{n^2}{2}-n+\frac{3}{2})\Gamma(\frac{n}{2}-\frac{1}{2})}+\frac{\Gamma(\frac{n}{2}-1)}{\Gamma(\frac{n}{2}+\frac{1}{2})}\right)
\end{aligned}$$

$$\begin{aligned}
&\geq \int_0^1 (1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)} d\zeta \cdot \frac{\sqrt{\pi}}{2}\left(\frac{1}{\frac{n^2}{2}-n+\frac{1}{2}}+\frac{1}{\frac{n}{2}-\frac{1}{2}}\right)\frac{1}{\sqrt{\frac{n}{2}-1}} \\
&= \int_0^1 (1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)} d\zeta \cdot \frac{\sqrt{2\pi}n}{(n-1)^2\sqrt{n-2}} \\
&\geq \int_0^1 (1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)} d\zeta \cdot \frac{\sqrt{2\pi}}{(n-1)^{3/2}} .
\end{aligned}$$

These results can be inserted into the quotient at the stage (81) and we obtain

$$\frac{\mathcal{I}_2}{\mathcal{I}_0} \leq \frac{\int_q^1 G(h)^{m-n}\frac{1}{h^{n-1}}(1-h^2)^{\frac{1}{2}(n^2-n-2)}dh}{\int_q^1 G(h)^{m-n}h(1-h^2)^{\frac{1}{2}(n^2-n-2)}dh} \cdot$$

$$\frac{\int_0^1 (1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)} \int_{\zeta-\varepsilon^{\alpha_2}}^{\zeta+\varepsilon^{\alpha_2}} |\zeta-\xi|^2 d\xi d\zeta}{\int_0^1 (1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)}\left(\zeta^2\int_{-1}^1(1-\xi^2)^{\frac{(n-4)}{2}}d\xi + \int_{-1}^1 \xi^2(1-\xi^2)^{\frac{(n-4)}{2}}d\xi\right)d\zeta}$$

$$\leq \frac{\int_q^1 G(h)^{m-n}\frac{1}{h^{n-1}}(1-h^2)^{\frac{1}{2}(n^2-n-2)}dh \int_0^1(1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)}d\zeta\cdot\frac{2}{3}\varepsilon^{3\alpha_2}}{\int_q^1 G(h)^{m-n}h(1-h^2)^{\frac{1}{2}(n^2-n-2)}dh \int_0^1(1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)}d\zeta\cdot\sqrt{2\pi}\frac{1}{(n-1)^{\frac{3}{2}}}}$$

$$\leq \frac{\sqrt{2}}{3\sqrt{\pi}}\varepsilon^{3\alpha_2}(n-1)^{\frac{3}{2}}\frac{\int_q^1 G(h)^{m-n}\frac{1}{h^{n-1}}(1-h^2)^{\frac{1}{2}(n^2-n-2)}dh}{\int_q^1 G(h)^{m-n}h(1-h^2)^{\frac{1}{2}(n^2-n-2)}dh} .$$

And this can (in the view of $h \geq q$) be simplified to

$$\begin{aligned}
\frac{\mathcal{I}_2}{\mathcal{I}_0} &\leq \frac{\sqrt{2}}{3\sqrt{\pi}}\varepsilon^{3\alpha_2}(n-1)^{\frac{3}{2}}\frac{1}{q^n} && (82)\\
\Rightarrow \quad P\left(\frac{|\theta-c_n^{n-1}|}{\sqrt{1-h^2}}\leq\varepsilon^{\alpha_2}\right) &\leq \frac{\sqrt{2}}{3\sqrt{\pi}}\varepsilon^{3\alpha_2}(n-1)^{\frac{3}{2}}\frac{1}{q^n} .
\end{aligned}$$

The Radius of the Spherical Support. Since in the last quotient we have developed bounds in terms of $\sqrt{1-h^2}$, which is the radius of the ball we meet in level h, it is now necessary to ask for the probability that this radius itself will be small.

This can be expressed by

$$P(\sqrt{1-h^2}\leq\varepsilon^{\alpha_3}) = \frac{\mathcal{I}_3}{\mathcal{I}_0}$$

$$= \frac{\int_q^1 G(h)^{m-n} I_{\{\sqrt{1-h^2} \leq \varepsilon^{\alpha_3}\}} \int_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int_{\mathbb{R}^{n-1}} |\theta - c_n^{n-1}| I_{\{c_n^{n-1} \leq \theta\}}}{\int_q^1 G(h)^{m-n} \int_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int_{\mathbb{R}^{n-1}} |\theta - c_n^{n-1}| I_{\{c_n^{n-1} \leq \theta\}}}$$

$$\frac{\int_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(c_1, \dots, c_{n-1}))^2 V(c_1, \dots, c_n)}{\int_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(c_1, \dots, c_{n-1}))^2 V(c_1, \dots, c_n)}$$

$$\frac{\prod_{i=1}^{n-1} f(c_i) d\bar{\bar{c}}_1 \cdots d\bar{\bar{c}}_{n-1} f(c_n) d\bar{c}_n d\theta dh}{\prod_{i=1}^{n-1} f(c_i) d\bar{\bar{c}}_1 \cdots d\bar{\bar{c}}_{n-1} f(c_n) d\bar{c}_n d\theta dh}.$$

We are now allowed to replace $V(c_1, \dots, c_n)$ by $\lambda_{n-2}(conv(c_1, \dots, c_{n-1})) \cdot \frac{h\|\theta - c_n^{n-1}\|}{n(n-1)\lambda_n(\Omega_n)}$ because every facet-area element is extended when we go to the sphere. But this extension factor increases when we look at a lower level of h and at the corresponding counterpart-element at the lower level. Note that the internal structure in each level of h (the "internal distribution") is identical for every level.
But when this extension is stronger for small values of h, then dropping this extension will help the high values of h. And this will increase the value of the high interval and the quotient.
Arguing and simplifying as in the section before leads to

$$\frac{\mathcal{I}_3}{\mathcal{I}_0} \leq \frac{\int_q^1 G(h)^{m-n} h I_{\{\sqrt{1-h^2} \leq \varepsilon^{\alpha_3}\}} \int_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int_{-\sqrt{1-h^2}}^{\theta} |\theta - c_n^{n-1}|^2}{\int_q^1 G(h)^{m-n} h \int_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int_{-\sqrt{1-h^2}}^{\theta} |\theta - c_n^{n-1}|^2}$$

$$\frac{(1-h^2-\theta^2)^{\frac{1}{2}(3(n-2)+(n-1)(n-4))}(1-h^2-(c_n^{n-1})^2)^{\frac{1}{2}(n-4)} dc_n^{n-1} d\theta dh}{(1-h^2-\theta^2)^{\frac{1}{2}(3(n-2)+(n-1)(n-4))}(1-h^2-(c_n^{n-1})^2)^{\frac{1}{2}(n-4)} dc_n^{n-1} d\theta dh}$$

$$\leq \frac{\int_{\sqrt{1-(\varepsilon^{\alpha_3})^2}}^1 G(h)^{m-n} h (1-h^2)^{\frac{1}{2}(n^2-n-2)} \int_{-1}^1 (1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)}}{\int_q^1 G(h)^{m-n} h (1-h^2)^{\frac{1}{2}(n^2-n-2)} \int_{-1}^1 (1-\zeta^2)^{\frac{1}{2}(n^2-2n-2)}}$$

$$\frac{\int_{-1}^{\zeta} |\zeta - \xi|^2 (1-\xi^2)^{\frac{1}{2}(n-4)} d\xi d\zeta dh}{\int_{-1}^{\zeta} |\zeta - \xi|^2 (1-\xi^2)^{\frac{1}{2}(n-4)} d\xi d\zeta dh} \quad \text{for } q \leq \sqrt{1-(\varepsilon^{\alpha_3})^2}$$

$$= \frac{\int\limits_{\sqrt{1-(\varepsilon^{\alpha_3})^2}}^{1} G(h)^{m-n} h(1-h^2)^{\frac{1}{2}(n^2-n-2)} dh}{\int\limits_{q}^{1} G(h)^{m-n} h(1-h^2)^{\frac{1}{2}(n^2-n-2)} dh} \,.$$

We take into regard that

$$g(h) = \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)}(1-h^2)^{\frac{n-3}{2}},$$
$$g_2(h) = \frac{\lambda_{n-2}(\omega_{n-1})}{(n-1)\lambda_{n-1}(\omega_n)}(1-h^2)^{\frac{n-1}{2}} = \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \int_h^1 (1-\sigma^2)^{\frac{n-3}{2}} \sigma d\sigma,$$
$$1 - G(h) = \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \int_h^1 (1-\sigma^2)^{\frac{n-3}{2}} d\sigma \,.$$

And from [7] we know that $g_2(h) \le 1 - G(h)$ and $\lim_{h\to 1} \frac{g_2(h)}{1-G(h)} = 1$ monotone, and moreover $(1-G(h))\left(h + (1-h)\frac{2}{n+1}\right) \le g_2(h)$ and
$g_2(h) \le (1-G(h))\left(h + (1-h)\frac{2\lambda_{n-2}(\omega_{n-1})}{(n-1)\lambda_{n-1}(\omega_n)}\right)$.
So we can modify the last quotient (after reducing factors depending on n only) to

$$\frac{\int\limits_{\sqrt{1-(\varepsilon^{\alpha_3})^2}}^{1} G(h)^{m-n} h g(h)(1-h^2)^{\frac{1}{2}(n^2-2n+1)} dh}{\int\limits_{q}^{1} G(h)^{m-n} h g(h)(1-h^2)^{\frac{1}{2}(n^2-2n+1)} dh}$$
$$= \frac{\int\limits_{\sqrt{1-(\varepsilon^{\alpha_3})^2}}^{1} G(h)^{m-n} h g(h) g_2(h)^{n-1} dh}{\int\limits_{q}^{1} G(h)^{m-n} h g(h) g_2(h)^{n-1} dh} \,.$$

Since $\dfrac{h(1-G(h))}{g_2(h)} = \dfrac{\int_h^1 (1-\sigma^2)^{\frac{n-3}{2}} h d\sigma}{\int_h^1 (1-\sigma^2)^{\frac{n-3}{2}} \sigma d\sigma} = \dfrac{\int_h^1 (1-\sigma^2)^{\frac{n-3}{2}} d\sigma}{\int_h^1 (1-\sigma^2)^{\frac{n-3}{2}} \frac{\sigma}{h} d\sigma}$ is an increasing function of h, we can strengthen the influence of the interval $[\sqrt{1-(\varepsilon^{\alpha_3})^2}, 1]$ by a replacement of g_2 by $h(1-G(h))$. So we know that

$$\frac{\mathcal{I}_3}{\mathcal{I}_0} \le \frac{\int\limits_{\sqrt{1-(\varepsilon^{\alpha_3})^2}}^{1} G(h)^{m-n} h^n g(h)(1-G(h))^{(n-1)} dh}{\int\limits_{q}^{1} G(h)^{m-n} h^n g(h)(1-G(h))^{(n-1)} dh}$$

$$\leq \frac{1}{q^n} \frac{\int_0^{Y(\sqrt{1-(\varepsilon^{\alpha_3})^2})} (1-Y)^{m-n} Y^{n-1} dY}{\int_0^{Y(q)} (1-Y)^{m-n} Y^{n-1} dY}, \tag{83}$$

by substituting $Y = 1 - G(h)$ and where $Y(q) = 1 - G(q)$ and $Y(\sqrt{1-(\varepsilon^{\alpha_3})^2}) = 1 - G(\sqrt{1-(\varepsilon^{\alpha_3})^2})$.
For a random variable $\tilde{Y}$ on $[0, Y(q)]$ let us define a probability $\tilde{P}(\tilde{Y} < \rho)$ by

$$\frac{\int_0^{\rho} (1-\tilde{y})^{m-n} \tilde{y}^{n-1} d\tilde{y}}{\int_0^{Y(q)} (1-\tilde{y})^{m-n} \tilde{y}^{n-1} d\tilde{y}}.$$

For an estimation of that probability we can employ Markov's inequality in the following form for $\rho = 1 - G(\sqrt{1-(\varepsilon^{\alpha_3})^2})$

$$\tilde{P}\Big(\tilde{Y} < 1 - G(\sqrt{1-(\varepsilon^{\alpha_3})^2})\Big) = \tilde{P}\left(\frac{1}{\tilde{Y}} > \frac{1}{1 - G(\sqrt{1-(\varepsilon^{\alpha_3})^2})}\right)$$

$$\leq \frac{\int_0^{Y(q)} (1-\tilde{y})^{m-n} \tilde{y}^{n-1-1} d\tilde{y}}{\int_0^{Y(q)} (1-\tilde{y})^{m-n} \tilde{y}^{n-1} d\tilde{y}} \cdot (1 - G(\sqrt{1-(\varepsilon^{\alpha_3})^2})).$$

If we can rely on

$$\int_0^{Y(q)} (1-\tilde{y})^{m-n} \tilde{y}^{n-1} d\tilde{y} \geq \frac{1}{2} \int_0^1 (1-\tilde{y})^{m-n} \tilde{y}^{n-1} d\tilde{y} \tag{84}$$

then we can conclude (using (83) and taking Y by $\tilde{Y}$):

$$q^n \cdot P(\sqrt{1-h^2} = q^n \frac{\mathcal{I}_3}{\mathcal{I}_0} \leq \varepsilon^{\alpha_3}) \leq \tilde{P}\big(Y < 1 - G(\sqrt{1-(\varepsilon^{\alpha_3})^2}\,)\big)$$

$$\leq 2 \cdot \frac{\int_0^1 (1-\tilde{y})^{m-n} \tilde{y}^{n-2} d\tilde{y}}{\int_0^1 (1-\tilde{y})^{m-n} \tilde{y}^{n-1} d\tilde{y}} \cdot \big(1 - G(\sqrt{1-(\varepsilon^{\alpha_3})^2}\,)\big)$$

$$= 2\frac{m}{n-1}\big(1 - G(\sqrt{1-(\varepsilon^{\alpha_3})^2}\,)\big).$$

So $P(\sqrt{1-h^2} \leq \varepsilon^{\alpha_3}) \leq \frac{2}{q^n}\frac{m}{n-1}(1 - G(\sqrt{1-(\varepsilon^{\alpha_3})^2}))$.
We should evaluate $G(\sqrt{1-(\varepsilon^{\alpha_3})^2})$ as good as possible.
Let $\bar{h} = \sqrt{1-(\varepsilon^{\alpha_3})^2}$. Then

$$1 - G(\bar{h}) \leq \frac{g_2(\bar{h})}{\bar{h} + (1-\bar{h})\frac{2}{n+1}} = \frac{(1-\bar{h}^2)^{\frac{n-1}{2}}}{\bar{h} + (1-\bar{h})\frac{2}{n+1}} \frac{\lambda_{n-2}(\omega_{n-1})}{(n-1)\lambda_{n-1}(\omega_n)}$$

$$\leq \frac{(\varepsilon^{\alpha_3})^{n-1}}{\frac{2}{n+1} + \sqrt{1-(\varepsilon^{\alpha_3})^2 \frac{n-1}{n+1}}} \cdot \frac{1}{\sqrt{(n-1)2\pi}}.$$

Hence, under the assumption (84)

$$\frac{\mathcal{I}_3}{\mathcal{I}_0} \leq \frac{2}{q^n} \frac{m}{n-1} \frac{1}{\sqrt{(n-1)2\pi}} \frac{(\varepsilon^{\alpha_3})^{n-1}}{\frac{2}{n+1} + \sqrt{1-(\varepsilon^{\alpha_3})^2 \frac{n-1}{n+1}}}.$$

Let us now derive a sufficient condition for the assumption (84) namely

$$\int_0^{1-G(q)} (1-\tilde{y})^{m-n} \tilde{y}^{n-1} d\tilde{y} \geq \frac{1}{2} \int_0^1 (1-\tilde{y})^{m-n} \tilde{y}^{n-1} d\tilde{y}.$$

To satisfy this, q has to be small enough. Again, we use the Markov-Inequality to prove

$$\frac{\int_{1-G(q)}^{1} (1-\tilde{y})^{m-n} \tilde{y}^{n-1} d\tilde{y}}{\int_0^1 (1-\tilde{y})^{m-n} \tilde{y}^{n-1} d\tilde{y}} \leq \frac{1}{2} \text{ for a suitable } q.$$

For another random variable $\hat{Y}$ we define the probability $\hat{P}(\hat{Y} \geq 1-G(q))$ by $\frac{\int_{1-G(q)}^{1} (1-\hat{y})^{m-n} \hat{y}^{n-1} d\hat{y}}{\int_0^1 (1-\hat{y})^{m-n} \hat{y}^{n-1} d\hat{y}}$, then

$$\hat{P}(\hat{Y} \geq 1-G(q)) \leq \frac{\int_0^1 (1-\hat{y})^{m-n} \hat{y}^{n-1+1} d\hat{y}}{\int_0^1 (1-\hat{y})^{m-n} \hat{y}^{n-1} d\hat{y} \cdot (1-G(q))} = \frac{n}{m+1} \frac{1}{(1-G(q))}.$$

So it is sufficient that the q in question satisfies $2\frac{n}{m+1} \leq 1-G(q)$, resp. $G(q) \leq 1-\frac{2n}{m+1}$. This can be satisfied for a fixed $q > 0$ if $\frac{2n}{m+1} \leq \frac{1}{2}$ or equivalently $4n \leq m+1$. So, as $\frac{2n}{m+1} \leq 1-G(q)$ is sufficient, and since $1-G(q) \geq g_2(q)$ also $g_2(q) \geq \frac{2n}{m+1}$ is sufficient. This means q should satisfy $\frac{2n}{m+1} \leq \frac{\lambda_{n-2}(\omega_{n-1})}{(n-1)\lambda_{n-1}(\omega_n)}(1-q^2)^{\frac{n-1}{2}}$ or equivalently $\frac{2n}{m+1} \frac{(n-1)\lambda_{n-1}(\omega_n)}{\lambda_{n-2}(\omega_{n-1})} \leq (1-q^2)^{\frac{n-1}{2}}$.
Sufficient is also $\frac{2n(n-1)}{m+1}\sqrt{\frac{2\pi}{n-2}} \leq (1-q^2)^{\frac{n-1}{2}}$ and $\frac{2^{3/2}\sqrt{\pi}n^{3/2}}{m+1}\sqrt{\frac{n-1}{n-2}} \leq (1-q^2)^{\frac{n-1}{2}}$, and therefore

$$\left(\frac{2^{3/2}\sqrt{\pi}n^{3/2}}{m+1}\sqrt{\frac{n-1}{n-2}}\right)^{\frac{2}{n-1}} \leq (1-q^2)$$

$$\Leftrightarrow \quad q \leq \sqrt{1-\left(\frac{2^{3/2}\sqrt{\pi}n^{3/2}}{m+1}\sqrt{\frac{n-1}{n-2}}\right)^{\frac{2}{n-1}}} =: \tilde{q}(m,n).$$

Note that the borderline $\tilde{q}(m,n)$ increases monotonically to 1, if $\frac{m+1}{4n}$ increases to infinity, because G^{-1} is a monotonically increasing function.

In other cases, namely where $m+1 \leq 4n$, we simply calculate the bound for $m+1 = 4n$ and use this for an estimation of the corresponding interval. This is feasible because increasing m leads to a support for low values of $\sqrt{1-h^2}$.

Our result is now:

$$P(\sqrt{1-h^2} \leq \varepsilon^{\alpha_3}) \leq \begin{cases} \frac{2m}{q^n(n-1)}(1-G(\sqrt{1-(\varepsilon^{\alpha_3})^2})) & \text{if } m+1 \geq 4n\,, \\ \frac{2(4n-1)}{q^n(n-1)}(1-G(\sqrt{1-(\varepsilon^{\alpha_3})^2})) & \text{if } m+1 < 4n\,. \end{cases} \quad (85)$$

Recall that

$$1 - G(\sqrt{1-(\varepsilon^{\alpha_3})^2}) \leq \frac{(\varepsilon^{\alpha_3})^{n-1}}{\frac{2}{n+1} + \sqrt{1-(\varepsilon^{\alpha_3})^2}\frac{n-1}{n+1}} \cdot \frac{1}{\sqrt{(n-1)2\pi}}\,. \quad (86)$$

The Kink Factor. The numerator in the fourth quotient is

$$\mathcal{I}_4 := \int_q^1 \int_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int_{\mathbb{R}^{n-1}} |\theta - c_n^{n-1}| I_{\{c_n^{n-1}\leq\theta\}}$$
$$\int_{\mathbb{R}^n}^{(m-n)} \prod_{k=n+1}^{m} I_{\{c_k^n\leq h\}} I_{\{K_{h,\theta}\neq\emptyset \text{ and } \mathrm{Min}_{k\in K_{h,\theta}} \frac{\sin\varphi_k}{\theta\sin\varphi_k + h\cos\varphi_k}\leq\varepsilon^{\alpha_4}\}}$$
$$\prod_{k=n+1}^{m} f(c_k)dc_{n+1}\cdots dc_m \int_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(c_1,\ldots,c_{n-1}))^2 V(c_1,\ldots,c_n)$$
$$\prod_{i=1}^{n-1} f(c_i)d\bar{\bar{c}}_1\cdots d\bar{\bar{c}}_{n-1} f(c_n)d\bar{c}_n d\theta dh\,.$$

$\mathcal{I}_0$ is identical, but there is no indicator $I_{\{K_{h,\theta}\neq\emptyset \text{ and } \mathrm{Min}_{k\in K_{h,\theta}} \frac{\sin\varphi_k}{\theta\sin\varphi_k + h\cos\varphi_k}\leq\varepsilon^{\alpha_4}\}}$. Let us study the situation, where $c_1,\ldots,c_n$ and consequently h, θ are fixed. The varying part of the integrals is then

$$\int_{\mathbb{R}^n}^{(m-n)} \prod_{k=n+1}^{m} I_{\{c_k^n\leq h\}} I_{\{K_{h,\theta}\neq\emptyset \text{ and } \mathrm{Min}_{k\in K_{h,\theta}} \frac{\sin\varphi_k}{\theta\sin\varphi_k + h\cos\varphi_k}\leq\varepsilon^{\alpha_4}\}}$$
$$\prod_{k=n+1}^{m} f(c_k)dc_{n+1}\cdots dc_m\,.$$

From a geometric point of view we should introduce a rotation angle for a movement around the point $\begin{pmatrix}\theta\\h\end{pmatrix}$ (two-dimensional) associated with that angle. Note that for $\varphi = 0$ means that we take the direction of e_{n-1} (resp. e_1) and $\varphi = \pi$ would lead in the opposite direction $-e_{n-1}$.

We deal with such rotations only as long as $\varphi \leq \varphi_{h,\theta}$, where $\varphi_{h,\theta}$ is the angle of the direction to the origin, induced by $\theta\sin\varphi_{h,\theta} + h\cos\varphi_{h,\theta} = 0$.

From (78) we know that

$$\frac{\sin \varphi_k}{\theta \sin \varphi_k + h \cos \varphi_k} \geq \sin \varphi_k \geq \frac{2}{\pi} \varphi_k ,$$

where the first quotient is an increasing function of φ_k as long as the denominator is positive (resp. as long as we are in the region of $K_{h,\theta}$) and we have the upper bound

$$\begin{aligned} & I_{\{K_{h,\theta} \neq \emptyset \text{ and } \mathrm{Min}_{k \in K_{h,\theta}} \frac{\sin \varphi_k}{\theta \sin \varphi_k + h \cos \varphi_k} \leq \varepsilon^{\alpha_4}\}} \leq I_{\{\mathrm{Min}_{k \in K_{h,\theta}} \frac{2}{\pi} \varphi_k \leq \varepsilon^{\alpha_4}\}} \\ = \; & 1 - I_{\{\mathrm{Min}_{k \in K_{h,\theta}} \frac{2}{\pi} \varphi_k \geq \varepsilon^{\alpha_4}\}} = 1 - \prod_{k \in K_{h,\theta}} I_{\{\frac{2}{\pi} \varphi_k \geq \varepsilon^{\alpha_4}\}} \\ \leq \; & 1 - \prod_{n+1}^{m} I_{\{\frac{2}{\pi} \varphi_k \geq \varepsilon^{\alpha_4}\}} . \end{aligned}$$

Hence we have for the partial integral

$$\begin{aligned} & \int_{\mathbb{R}^n}^{(m-n)} \prod_{k=n+1}^{m} I_{\{c_k^n \leq h\}} I_{\{K_{h,\theta} \neq \emptyset \text{ and } \mathrm{Min}_{k \in K_{h,\theta}} \frac{\sin \varphi_k}{\theta \sin \varphi_k + h \cos \varphi_k} \leq \varepsilon^{\alpha_4}\}} \\ & \prod_{k=n+1}^{m} f(c_k) dc_{n+1} \cdots dc_m \\ \leq \; & \int_{\mathbb{R}^n}^{(m-n)} \prod_{k=n+1}^{m} I_{\{c_k^n \leq h\}} \Big(1 - \prod_{n+1}^{m} I_{\{\frac{2}{\pi} \varphi_k \geq \varepsilon^{\alpha_4}\}} \Big) \prod_{k=n+1}^{m} f(c_k) dc_{n+1} \cdots dc_m \\ = \; & G(h)^{m-n} - \bigg(\int_{\mathbb{R}^n} I_{\{c_k^n \leq h \wedge \frac{2}{\pi} \varphi_k \geq \varepsilon^{\alpha_4}\}} f(c_{n+1}) dc_{n+1} \bigg)^{m-n} . \end{aligned}$$

Note, that we can treat each vector c_k $(k > n)$ separately and that for symmetry reasons c_{n+1} may replace each c_k . Let us define

$$\widetilde{G}(\eta, h, \theta) := \int_{\mathbb{R}^n} I_{\{c_{n+1}^n \leq h \varphi_{n+1} \geq \eta\}} f(c_{n+1}) dc_{n+1} .$$

Then the proof of the following Lemma is immediate.

LEMMA 3.3. *Suppose that* $h \in (0,1)$, $\theta \in (-\sqrt{1-h^2}, \sqrt{1-h^2})$, $\eta \in (0, \pi)$. *Then*

1. $\widetilde{G}(\eta, h, \theta)$ *is monotonically decreasing with* η *for fixed* (h, θ) *and monotonically increasing with* θ *for fixed* (η, h) .

2. $\widetilde{G}(\eta, h, -\sqrt{1-h^2}) = G(\mathsf{h}(\eta))$ *with* $\mathsf{h}(\eta) = h \cos \eta - \sqrt{1-h^2} \sin \eta$.

3. $G(h)^{m-n} - \widetilde{G}(\eta, h, \theta)^{m-n} \leq G(h)^{m-n} - G(\mathsf{h}(\eta))^{m-n}$
$\leq \frac{\sqrt{2} \lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \cdot \eta (m-n) G(h)^{m-n-1} \leq 2 \sqrt{\frac{\pi}{n-2}} \cdot \eta (m-n) G(h)^{m-n-1}$.

Proof. *Part 1:* The first monotonicity is trivial, because we have even the monotonicity in the inclusion relation.

The second part is also simple, because for fixed η, h, θ the bound for our feasible region is given by a certain hyperplane. If we increase θ and keep η, h fixed, then we use a corresponding parallel hyperplane. And the stripe between the two hyperplanes has become feasible, too. So also this relation relies on monotonicity in the inclusion relation.

Part 2: If we use a kink-angle η at $\begin{pmatrix} -\sqrt{1-h^2} \\ h \end{pmatrix}$ then this induces a hyperplane through $\begin{pmatrix} -\sqrt{1-h^2} \\ h \end{pmatrix}$, which has a distance to the origin of $h\cos\eta - \sqrt{1-h^2}\sin\eta = \mathsf{h}(\eta)$.

This new hyperplane (taken as a restriction) makes the restriction $c^n_{n+1} \leq h$ redundant, because f

$$\begin{aligned} & x^{n-1}\sin\eta + x^n\cos\eta \leq h\cos\eta - \sqrt{1-h^2}\sin\eta \\ \Leftrightarrow \quad & (x^{n-1} + \sqrt{1-h^2})\sin\eta \leq (h - x^n)\cos\eta \end{aligned}$$

Since $\eta \in (0, \frac{\pi}{2})$ we have $\sin\eta > 0$, $\cos\eta > 0$ and therefore

$$\frac{\sin\eta}{\cos\eta}(x^{n-1} + \sqrt{1-h^2}) \leq (h - x^n).$$

In the case that $x^{n-1} \geq -\sqrt{1-h^2}$ the righthand side of the inequality above is nonnegative and it follows that $x^n \leq h$. The case that $x^{n-1} < -\sqrt{1-h^2}$ and $x^n \geq h$ cannot occur as all points are lying inside of the unit sphere.

This means that the region below the new hyperplane is the feasible region and that $\widetilde{G}(\eta, h, -\sqrt{1-h^2}) = G(\mathsf{h}(\eta))$.

Part 3:

$$\begin{aligned} G(h)^{m-n} - \widetilde{G}(\eta,h,\theta)^{m-n} &\leq G(h)^{m-n} - \widetilde{G}(\eta,h,-\sqrt{1-h^2})^{m-n} \\ &\leq G(h)^{m-n} - G(\mathsf{h}(\eta))^{m-n}. \end{aligned}$$

If we identify $G(\mathsf{h}(\eta))$ with $(1-\rho)G(h)$ for some $\rho \in (0,1)$ (recall that $\mathsf{h}(\eta) < \eta$), then it is clear (with Bernoulli's inequality) that

$$\begin{aligned} & G(h)^{m-n} - G(\mathsf{h}(\eta))^{m-n} = G(h)^{m-n}(1-(1-\rho)^{m-n}) \\ \leq \; & G(h)^{m-n}(1-(1-(m-n)\rho)) = G(h)^{m-n}(m-n)\rho. \end{aligned}$$

As we see that

$$\begin{aligned} G(h)^{m-n} - \widetilde{G}(\eta,h,\theta) &\leq G(h)^{m-n} - G(\mathsf{h}(\eta))^{m-n} \\ &= G(h)^{m-n}\left(1 - \underbrace{\left(\frac{G(\mathsf{h}(\eta))}{G(h)}\right)}_{=1-\rho}{}^{m-n}\right) \\ &\leq G(h)^{m-n}(m-n)\left(1 - \frac{G(\mathsf{h}(\eta))}{G(h)}\right) \\ &= G(h)^{m-n-1}(m-n)(G(h) - G(\mathsf{h}(\eta))). \end{aligned}$$

We know that

$$\begin{aligned}
G(h) - G(\mathsf{h}(\eta)) &= \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \int_{\mathsf{h}(\eta)}^{h} (1-\sigma^2)^{\frac{n-3}{2}} d\sigma \\
&\leq \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \left(1-\mathsf{h}(\eta)^2\right)^{\frac{n-3}{2}} (h - \mathsf{h}(\eta)) \\
&= \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \left(1-\mathsf{h}(\eta)^2\right)^{\frac{n-3}{2}} (h(1-\cos\eta) + \sqrt{1-h^2}\sin\eta)) \\
&\leq \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \cdot 1 \cdot (h\eta + \sqrt{1-h^2}\eta) \\
&\leq \sqrt{\frac{2\pi}{n-2}}\sqrt{2}\eta = 2\sqrt{\frac{\pi}{n-2}}\eta .
\end{aligned}$$

This is a uniform result for all locations of (h, θ) and proves the lemma. ∎

Now return to the analysis of the partial integral.

$$\begin{aligned}
&\int_{\mathbb{R}^n}^{(m-n)} \prod_{k=n+1}^{m} I_{\{c_k^n \leq h\}} I_{\{K_{h,\theta} \neq \emptyset \text{ and } \mathrm{Min}_{k \in K_{h,\theta}} \frac{\sin\varphi_k}{\theta \sin\varphi_k + h\cos\varphi_k} \leq \varepsilon^{\alpha_4}\}} \\
&\qquad \prod_{k=n+1}^{m} f(c_k) dc_{n+1} \cdots dc_m \\
&\leq G(h)^{m-n} - \widetilde{G}(\frac{\pi}{2}\varepsilon^{\alpha_4}, h, \theta)^{m-n} \\
&\leq G(h)^{m-n-1}(m-n)2\sqrt{\frac{\pi}{n-2}}\frac{\pi}{2}\varepsilon^{\alpha_4} .
\end{aligned}$$

So we have to compare the upper bound for $\mathcal{I}_4$

$$\begin{aligned}
\mathcal{I}_4 \leq\ & (m-n)\sqrt{\frac{\pi}{n-2}}\pi\varepsilon^{\alpha_4} \cdot \\
& \int_q^1 G(h)^{m-n-1} \int_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int_{\mathbb{R}^{n-1}} |\theta - c_n^{n-1}| I_{\{c_n^{n-1} \leq \theta\}} \\
& \quad \int_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(c_1, \ldots, c_{n-1}))^2 V(c_1, \ldots, c_n) \\
& \quad \prod_{i=1}^{n-1} f(c_i) d\bar{\bar{c}}_1 \cdots d\bar{\bar{c}}_{n-1} f(c_n) d\bar{c}_n d\theta dh
\end{aligned}$$

with the denominator

$$\mathcal{I}_0 = \int_q^1 G(h)^{m-n} \int_{-\sqrt{1-h^2}}^{\sqrt{1-h^2}} \int_{\mathbb{R}^{n-1}} |\theta - c_n^{n-1}| I_{\{c_n^{n-1} \leq \theta\}}$$

$$\int_{\mathbb{R}^{n-2}}^{(n-1)} \lambda_{n-2}(conv(c_1,\ldots,c_{n-1}))^2 V(c_1,\ldots,c_n)$$
$$\prod_{i=1}^{n-1} f(c_i)d\bar{\bar{c}}_1 \cdots d\bar{\bar{c}}_{n-1} f(c_n)d\bar{c}_n d\theta dh\,.$$

Now assume that we have a fixed configuration of $c_1,\ldots,c_n$ and h,θ. Then the pointwise quotient of the two integrals is less than

$$\frac{(m-n)\sqrt{\frac{\pi}{n-2}}\pi\varepsilon^{\alpha_4}}{G(h)} \le 2\cdot(m-n)\sqrt{\frac{\pi}{n-2}}\pi\varepsilon^{\alpha_4} \tag{87}$$

since $G(h) > \frac{1}{2}$.

This is the desired upper bound for $\frac{\mathcal{I}_4}{\mathcal{I}_0}$.

3.2.4. *The Distribution of the Objective Difference (Part 2)*

After the geometric discussion of the difference between the best and the second best vertex in section 3.2.1 and the technical estimation of different quotients in 3.2.2 and 3.2.3 we return to the estimation of the distribution function F_U.

An Estimation for the Distribution Function. Let us summarize what we know about our four quotients from (80), (82), (85), (86) and (87)

$$\begin{aligned}
\frac{\mathcal{I}_1}{\mathcal{I}_0} &= P\left(\frac{dist(v,\overline{H})}{dist(c_n,\overline{H})} \le \varepsilon^{\alpha_1}\right) \le (n-1)\varepsilon^{\alpha_1}\,,\\
\frac{\mathcal{I}_2}{\mathcal{I}_0} &= P\left(\frac{\theta - c_n^{n-1}}{\sqrt{1-h^2}} \le \varepsilon^{\alpha_2}\right) \le \frac{\sqrt{2}}{3\sqrt{\pi}}(n-1)^{\frac{3}{2}}\frac{1}{q^n}\varepsilon^{3\alpha_2}\,,\\
\frac{\mathcal{I}_3}{\mathcal{I}_0} &= P(\sqrt{1-h^2} \le \varepsilon^{\alpha_3}) \le \frac{2}{q^n}\frac{m}{n-1}(1-G(\sqrt{1-\varepsilon^{2\alpha_3}}))\\
&\le \frac{2}{q^n}\frac{m}{n-1}\frac{1}{\sqrt{2\pi(n-1)}}\frac{1}{\frac{2}{n-1}+\sqrt{1-\varepsilon^{2\alpha_3}}\frac{n-1}{n+1}}\varepsilon^{(n-1)\alpha_3}\\
&\le \frac{\sqrt{2}}{\sqrt{\pi}}\frac{m}{(n-1)^{\frac{3}{2}}}\frac{1}{q^n}\frac{1}{\frac{2}{n-1}+q\frac{n-1}{n+1}}\varepsilon^{(n-1)\alpha_3}\\
&\qquad \text{(as we assume } q < \sqrt{1-\varepsilon^{2\alpha_3}}\text{)}\,,\\
\frac{\mathcal{I}_4}{\mathcal{I}_0} &= P\left(\operatorname*{Min}_{k\in K_{h,\theta}} \frac{\sin\varphi_k}{\theta\sin\varphi_k + h\cos\varphi_k} \le \varepsilon^{\alpha_4}\right) \le 2\cdot(m-n)\sqrt{\frac{\pi}{n-2}}\pi\varepsilon^{\alpha_4}\,.
\end{aligned}$$

Altogether we have

$$\begin{aligned}
&\frac{\mathcal{I}_1}{\mathcal{I}_0}+\frac{\mathcal{I}_2}{\mathcal{I}_0}+\frac{\mathcal{I}_3}{\mathcal{I}_0}+\frac{\mathcal{I}_4}{\mathcal{I}_0}\\
&\le (n-1)\varepsilon^{\alpha_1} + \frac{\sqrt{2}}{3\sqrt{\pi}}(n-1)^{\frac{3}{2}}\frac{1}{q^n}\varepsilon^{3\alpha_2}
\end{aligned}$$

$$+\frac{\sqrt{2}}{\sqrt{\pi}}\frac{m}{(n-1)^{\frac{3}{2}}}\frac{1}{q^n}\frac{1}{\frac{2}{n-1}+q^{\frac{n-1}{n+1}}}\varepsilon^{(n-1)\alpha_3}+2(m-n)\sqrt{\frac{\pi}{n-2}}\pi\varepsilon^{\alpha_4}\,,$$

and for the distribution function F_U follows

$$\begin{aligned}
F_U(\varepsilon) &= P(U\le\varepsilon \mid x_\Delta \text{ is optimal vertex} \wedge \|x_\Delta\|\le\tfrac{1}{q}) \\
&\le n\cdot\left(\frac{\mathcal{I}_1}{\mathcal{I}_0}+\frac{\mathcal{I}_2}{\mathcal{I}_0}+\frac{\mathcal{I}_3}{\mathcal{I}_0}+\frac{\mathcal{I}_4}{\mathcal{I}_0}\right) \\
&\le n(n-1)\varepsilon^{\alpha_1}+\frac{\sqrt{2}n(n-1)^{3/2}}{3\sqrt{\pi}}\frac{1}{q^n}\varepsilon^{3\alpha_2} \\
&\quad+\frac{\sqrt{2}}{\sqrt{\pi}}\frac{mn}{(n-1)^{\frac{3}{2}}}\frac{1}{q^n}\frac{1}{\frac{2}{n-1}+q^{\frac{n-1}{n+1}}}\varepsilon^{(n-1)\alpha_3}+(m-n)\frac{2n\pi^{\frac{3}{2}}}{\sqrt{n-2}}\varepsilon^{\alpha_4} \\
&= s_1(n)\varepsilon^{\alpha_1}+s_2(n,q)\varepsilon^{3\alpha_2}+s_3(m,n,q)\varepsilon^{(n-1)\alpha_3}+s_4(m,n)\varepsilon^{\alpha_4}
\end{aligned}$$

where $s_1(n):=n(n-1)\,,\qquad s_2(n,q):=\dfrac{\sqrt{2}n(n-1)^{3/2}}{3\sqrt{\pi}}\dfrac{1}{q^n}\,,$

$$s_3(m,n,q):=\frac{\sqrt{2}}{\sqrt{\pi}}\frac{mn}{(n-1)^{\frac{3}{2}}}\frac{1}{q^n}\frac{1}{\frac{2}{n-1}+q^{\frac{n-1}{n+1}}} \quad\text{and}$$

$$s_4(m,n):=(m-n)\frac{2n\pi^{\frac{3}{2}}}{\sqrt{n-2}}\,.$$

Now, we want to maximize the region of validity in terms of ε, where this upper bound is smaller than 1 (via variation of α_1, α_2, α_3 and α_4). That means we have to solve a nonlinear optimization problem:

$$\begin{aligned}
\min\quad & -\varepsilon \quad (\Leftrightarrow \quad \max\ \varepsilon) \\
\text{s.t.}\quad & s_1\varepsilon^{\alpha_1}+s_2\varepsilon^{3\alpha_2}+s_3\varepsilon^{(n-1)\alpha_3}+s_4\varepsilon^{\alpha_4}\le 1 \\
& \alpha_1+\alpha_2+\alpha_3+\alpha_4\le 1,\quad \alpha_i\ge 0,\quad 0\le\varepsilon\le 1\,.
\end{aligned}$$

Here and in the following discussion we use $s_1,\ldots,s_4$ for $s_1(n)$, $s_2(n,q)$, $s_3(m,n,q)$ and $s_4(m,n)$ for abbreviation. In general good candidates for (local) minima are points $(\alpha_1,\alpha_2,\alpha_3,\alpha_4,\varepsilon)$ that satisfy the Karush-Kuhn-Tucker condition for this problem. It is easy to verify that the KKT-point is

$$\begin{aligned}
\bar\varepsilon &= \left(s_1(3s_2)^{1/3}((n-1)s_3)^{1/(n-1)}s_4(\tfrac{7}{3}+\tfrac{1}{n-1})^{\frac{7}{3}+\frac{1}{n-1}}\right)^{-1} \\
\bar\alpha_1 &= -\frac{1}{\ln\bar\varepsilon}\cdot\ln\left(s_1(\tfrac{7}{3}+\tfrac{1}{n-1})\right) \\
\bar\alpha_2 &= -\frac{1}{3\ln\bar\varepsilon}\cdot\ln\left(3s_2(\tfrac{7}{3}+\tfrac{1}{n-1})\right) \\
\bar\alpha_3 &= -\frac{1}{(n-1)\ln\bar\varepsilon}\cdot\ln\left((n-1)s_3(\tfrac{7}{3}+\tfrac{1}{n-1})\right) \\
\bar\alpha_4 &= -\frac{1}{\ln\bar\varepsilon}\cdot\ln\left(s_4(\tfrac{7}{3}+\tfrac{1}{n-1})\right)\,.
\end{aligned} \tag{88}$$

Moreover it is easy to verify that this point $(\bar{\alpha}_1, \bar{\alpha}_2, \bar{\alpha}_3, \bar{\alpha}_4, \bar{\varepsilon})$ satisfies some second order conditions (compare [5], p. 573) that will guarantee the KKT-point to be a local minimum.

So we have proven the following upper bound for $F_U(x)$:

$$\begin{aligned} F_U(\varepsilon) &= P(U \le \varepsilon \mid x_\Delta \text{ is optimal vertex} \wedge \|x_\Delta\| \le \tfrac{1}{q}) \\ &\le s_1\varepsilon^{\bar{\alpha}_1} + s_2\varepsilon^{3\bar{\alpha}_2} + s_3\varepsilon^{(n-1)\bar{\alpha}_3} + s_4\varepsilon^{\bar{\alpha}_4} \end{aligned} \tag{89}$$

for $0 < q < \sqrt{1-\bar{\varepsilon}^{2\bar{\alpha}_3}}$, resp. for $0 < q < \sqrt{1-2m^{-2/(n-1)}}$ (as $\bar{\varepsilon}^{2\bar{\alpha}_3}$ can be bounded by $O(m^{-2/(n-1)})$ for any $q \in (0,1)$ although $\bar{\varepsilon}$ and $\bar{\alpha}_3$ depend on q itself, compare (93)). And we can interpret

$$\widetilde{F}(\varepsilon) := \begin{cases} 0 & \forall \varepsilon < 0\,, \\ s_1\varepsilon^{\bar{\alpha}_1} + s_2\varepsilon^{3\bar{\alpha}_2} + s_3\varepsilon^{(n-1)\bar{\alpha}_3} + s_4\varepsilon^{\bar{\alpha}_4} & \forall \varepsilon \in [0, \bar{\varepsilon}]\,, \\ 1 & \forall \varepsilon > \bar{\varepsilon}\,; \end{cases} \tag{90}$$

as a distribution function and obviously we have $F_U(\varepsilon) \le \widetilde{F}(\varepsilon)$ for all ε.

3.3. The Average Number of Iterations until Stopping

In this section we want to calculate the average number of iterations until stopping.

As a probabilistic model we use again the specialization (38) of the rotation-symmetry-model. From Theorem 3.2 we know that (compare (73))

$$K_U \;\le\; \begin{cases} 0 & \text{if } \mu_0(m+\sqrt{m}) \le U\,, \\ 6\sqrt{m}(-\ln U + \ln(\mu_0(m+\sqrt{m})) & \text{if } \mu_0(m+\sqrt{m}) > U\,; \end{cases}$$

where μ_0 is the barrier parameter corresponding to the starting point,

is an upper bound on the number of iterations for the Barrier Method and therefore an upper bound on the average number of iterations can be calculated as $E[K_U]$.

If we would know the exact distribution function $F_U(\varepsilon)$ of U we could simply calculate the expectation value as $E[K_U] = \int K_U dF_U$. Unfortunately, we only know an upper bound (89) on the distribution function $F_U(\varepsilon)$ and this upper bound is restricted to the condition that $\|x_{opt}\| \le \frac{1}{q}$ for a (so far not specified) parameter $q \in (0,1)$.

For further considerations we will restrict to this condition and we want to use the approximate distribution function $\widetilde{F}$ as defined in (90).

One advantage is that we can derive an upper bound on μ_0 if $\|x_{opt}\| \le \frac{1}{q}$, namely $\mu_0 \le 3(1+\frac{1}{q})$, by using Remark 2.11 and 2.12 with $\tau_{var} = \frac{1}{2}$ and $\underline{\eta} = \underline{\tau} = 0.36$. This leads to an upper bound on K_U, denoted by $\overline{K}_U$:

$$K_U \le \overline{K}_U := \begin{cases} 0 & \text{if } 3(1+\frac{1}{q})(m+\sqrt{m}) \le U, \\ 6\sqrt{m}\Big(-\ln U + \ln(3(1+\frac{1}{q})(m+\sqrt{m}))\Big) & \\ \qquad \text{if } 3(1+\frac{1}{q})(m+\sqrt{m}) > U; & \end{cases}$$

and we conclude $E[K_U] \le E[\overline{K}_U]$, where the expectation value is still calculated with respect to the exact distribution function F_U. But we can use any approximate distribution function $\widetilde{F}$ with the property $F_U(\varepsilon) \le \widetilde{F}(\varepsilon)$ for all $\varepsilon \ge 0$ to derive an upper bound on the expectation value. Such an approximate distribution function $\widetilde{F}$ puts more weight on the small values of U than the exact distribution function and this variation of weights leads to

$$E_{F_U}[\overline{K}_U] \le E_{\widetilde{F}}[\overline{K}_U]$$

as the small values of U cause a large number of iterations. Analogously to paragraph 2.3.2 $E_{F_U}[\,.\,]$ denotes the expectation value with respect to the exact distribution function F_U and $E_{\widetilde{F}}[\,.\,]$ is the expectation value with respect to the approximate distribution function $\widetilde{F}$.

Before we start to calculate $E_{\widetilde{F}}[\overline{K}_U]$ we introduce $\overline{u}(q,m) := 3(1+\frac{1}{q})(m+\sqrt{m})$ for abbreviation. So, we obtain

$$\begin{aligned} E_{\widetilde{F}}[\overline{K}_U] &= \int_0^{\overline{u}(q,m)} 6\sqrt{m}(-\ln\varepsilon + \ln(\overline{u}(q,m))d\widetilde{F}(\varepsilon) \\ &= 6\sqrt{m}\int_0^{\overline{u}(q,m)} \ln\left(\frac{\overline{u}(q,m)}{\varepsilon}\right) d\widetilde{F}(\varepsilon) \\ &= -6\sqrt{m}\left(\ln\left(\frac{\varepsilon}{\overline{u}(q,m)}\right)\widetilde{F}(\varepsilon)\right)\Big|_0^{\overline{u}(q,m)} + 6\sqrt{m}\int_0^{\overline{u}(q,m)} \varepsilon^{-1}\widetilde{F}(\varepsilon)d\varepsilon \\ &= -6\sqrt{m}\left(\ln 1\cdot\widetilde{F}(\overline{u}(q,m)) - \lim_{\varepsilon\to 0}\ln\left(\frac{\varepsilon}{\overline{u}(q,m)}\right)\widetilde{F}(\varepsilon)\right) \\ &\quad + 6\sqrt{m}\int_0^{\overline{u}(q,m)} \varepsilon^{-1}\widetilde{F}(\varepsilon)d\varepsilon\,. \end{aligned}$$

With the rules of l'Hospital we obtain $\lim_{\varepsilon\to 0}\ln\left(\frac{\varepsilon}{\overline{u}(q,m)}\right)\widetilde{F}(\varepsilon) = 0$ and since $\widetilde{F}(\varepsilon) = 1$ for all $\varepsilon > \bar{\varepsilon}$, resp. $\widetilde{F}(\varepsilon) = s_1\varepsilon^{\bar{\alpha}_1} + s_2\varepsilon^{3\bar{\alpha}_2} + s_3\varepsilon^{(n-1)\bar{\alpha}_3} + s_4\varepsilon^{\bar{\alpha}_4}$ for $\varepsilon \in [0,\bar{\varepsilon}]$, and $\overline{u}(q,m) \ge 1 \ge \bar{\varepsilon}$ we can proceed in the following way:

$$\begin{aligned} E_{\widetilde{F}}[\overline{K}_U] &\le 6\sqrt{m}\int_0^{\bar{\varepsilon}} \varepsilon^{-1}\widetilde{F}(\varepsilon)d\varepsilon + 6\sqrt{m}\int_{\bar{\varepsilon}}^{\overline{u}(q,m)} \varepsilon^{-1}\cdot 1\,d\varepsilon \\ &= 6\sqrt{m}\int_0^{\bar{\varepsilon}} \varepsilon^{-1}(s_1\varepsilon^{\bar{\alpha}_1} + s_2\varepsilon^{3\bar{\alpha}_2} + s_3\varepsilon^{(n-1)\bar{\alpha}_3} + s_4\varepsilon^{\bar{\alpha}_4})d\varepsilon + 6\sqrt{m}\ln\varepsilon\Big|_{\bar{\varepsilon}}^{\overline{u}(q,m)} \\ &= 6\sqrt{m}\Big(\frac{s_1}{\bar{\alpha}_1}\bar{\varepsilon}^{\bar{\alpha}_1} + \frac{s_2}{3\bar{\alpha}_2}\bar{\varepsilon}^{3\bar{\alpha}_2} + \frac{s_3}{(n-1)\bar{\alpha}_3}\bar{\varepsilon}^{(n-1)\bar{\alpha}_3} + \frac{s_4}{\bar{\alpha}_4}\bar{\varepsilon}^{\bar{\alpha}_4}\Big) \\ &\quad + 6\sqrt{m}\,(\ln\overline{u}(q,m) - \ln\bar{\varepsilon}). \end{aligned}$$

At this point we could stop, as we have found an upper bound on the average number of steps. But we are interested in an upper bound in terms of n and m, the dimensions of the optimization problem. So, we go on and we will prove the following upper bound

$$E_{\widetilde{F}}[\overline{K}_U] \leq O(\sqrt{m}(|\ln q| + \ln n + \ln m)).$$

Therefore, we conclude from (88) that

$$\begin{aligned} \bar{\varepsilon}^{\bar{\alpha}_1} &= \Big(s_1\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big)\Big)^{-1}, \qquad \bar{\varepsilon}^{3\bar{\alpha}_2} = \Big(3s_2\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big)\Big)^{-1}, \\ \bar{\varepsilon}^{(n-1)\bar{\alpha}_3} &= \Big((n-1)s_3\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big)\Big)^{-1} \text{ and } \bar{\varepsilon}^{\bar{\alpha}_4} = \Big(s_4\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big)\Big)^{-1} \end{aligned} \tag{91}$$

and

$$\begin{aligned} \frac{s_1}{\bar{\alpha}_1}\bar{\varepsilon}^{\bar{\alpha}_1} &= \frac{-\ln\bar{\varepsilon}}{\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big)\ln\Big(s_1\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big)\Big)}, \\ \frac{s_2}{3\bar{\alpha}_2}\bar{\varepsilon}^{3\bar{\alpha}_2} &= \frac{-\ln\bar{\varepsilon}}{\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big)3\ln\Big(3s_2\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big)\Big)}, \\ \frac{s_3}{(n-1)\bar{\alpha}_3}\bar{\varepsilon}^{(n-1)\bar{\alpha}_3} &= \frac{-\ln\bar{\varepsilon}}{\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big)(n-1)\ln\Big((n-1)s_3\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big)\Big)}, \\ \frac{s_4}{\bar{\alpha}_4}\bar{\varepsilon}^{\bar{\alpha}_4} &= \frac{-\ln\bar{\varepsilon}}{\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big)\ln\Big(s_4\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big)\Big)}. \end{aligned}$$

For $q \in (0,1)$ and assuming $m > n \geq 3$ we easily obtain lower bounds on some of the figures which appear in the denominators of the quotients above, especially

$$\begin{aligned} & s_1\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big) = n(n-1)\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big) \geq 3\cdot 2\cdot\frac{7}{3}+3 = 17 \\ & 3s_2\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big) = 3\frac{\sqrt{2}n(n-1)^{3/2}}{3\sqrt{\pi}}\frac{1}{q^n}\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big) \\ & = \frac{\sqrt{2}n(n-1)^{1/2}}{\sqrt{\pi}q^n}\big(\tfrac{7(n-1)}{3}+1\big) \geq \frac{\sqrt{2}\cdot 3\cdot\sqrt{2}}{\sqrt{\pi}\,1^n}\big(\tfrac{7\cdot 2}{3}+1\big) \geq 19 \\ & (n-1)s_3\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big) = (n-1)\frac{\sqrt{2}}{\sqrt{\pi}}\frac{mn}{(n-1)^{\frac{3}{2}}}\frac{1}{q^n}\frac{1}{\frac{2}{n-1}+q\frac{n-1}{n+1}}\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big) \\ & \geq \frac{\sqrt{2}}{\sqrt{\pi}}\frac{mn}{(n-1)^{\frac{1}{2}}}\cdot 1\cdot\frac{1}{2}\cdot\frac{7}{3} \geq \frac{\sqrt{2}}{\sqrt{\pi}}\cdot\frac{4\cdot 3}{\sqrt{2}}\cdot\frac{7}{6} \geq 7.5 \\ & s_4\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big) = (m-n)\frac{2n\pi^{\frac{3}{2}}}{\sqrt{n-2}}\big(\tfrac{7}{3}+\tfrac{1}{n-1}\big) \\ & \qquad \geq 1\cdot\frac{2\cdot 3\pi^{\frac{3}{2}}}{1}\frac{7}{3} \geq 77. \end{aligned}$$

It follows that $\frac{s_1}{\bar\alpha_1}\bar\varepsilon^{\bar\alpha_1} \le -(\ln\bar\varepsilon)/\left(\frac{7}{3}\ln 17\right)$, $\frac{s_2}{3\bar\alpha_2}\bar\varepsilon^{3\bar\alpha_2} \le -(\ln\bar\varepsilon)/(7\ln 19)$, $\frac{s_3\bar\varepsilon^{(n-1)\bar\alpha_3}}{(n-1)\bar\alpha_3} \le -(\ln\bar\varepsilon)/\left(\frac{14}{3}\ln 7.5\right)$ and $\frac{s_4}{\bar\alpha_4}\bar\varepsilon^{\bar\alpha_4} \le -(\ln\bar\varepsilon)/\left(\frac{7}{3}\ln 77\right)$, and therefore

$$\begin{aligned}
E_{\widetilde{F}}[\overline{K}_U] &\le 6\sqrt{m}(-\ln\bar\varepsilon)\Big(\frac{1}{\frac{7}{3}\ln 17}+\frac{1}{7\ln 19}+\frac{1}{\frac{14}{3}\ln 7.5}+\frac{1}{\frac{7}{3}\ln 77}\Big) \\
&\quad + 6\sqrt{m}(\ln\overline{u}(q,m) - \ln\bar\varepsilon) \\
&\le 9\sqrt{m}\,|\ln\bar\varepsilon| + 6\sqrt{m}\ln\overline{u}(q,m)\,. \qquad (92)
\end{aligned}$$

To derive an upper bound that depends on m, n and q only we will approximate $|\ln\bar\varepsilon|$ and $\ln\overline{u}(q,m)$. Remembering $q \in (0,1)$ we obtain

$$\begin{aligned}
\ln\overline{u}(q,m) &= \ln\left(3\left(1+\tfrac{1}{q}\right)(m+\sqrt{m})\right) \\
&\le \ln\left(\tfrac{6}{q}(m+\sqrt{m})\right) = \ln 6 + |\ln q| + \ln(m+\sqrt{m})
\end{aligned}$$

and

$$\begin{aligned}
|\ln\bar\varepsilon| &= \ln\Bigg(n(n-1)\cdot\Big(3\frac{\sqrt{2}n(n-1)^{3/2}}{3\sqrt{\pi}}\frac{1}{q^n}\Big)^{1/3} \\
&\qquad \cdot\Big((n-1)\frac{\sqrt{2}}{\sqrt{\pi}}\frac{mn}{(n-1)^{\frac{3}{2}}}\frac{1}{q^n}\frac{1}{\frac{2}{n-1}+q\frac{n-1}{n+1}}\Big)^{1/(n-1)} \\
&\qquad \cdot(m-n)\frac{2n\pi^{\frac{3}{2}}}{\sqrt{n-2}}\cdot\Big(\frac{7}{3}+\frac{1}{n-1}\Big)^{(\frac{7}{3}+\frac{1}{n-1})}\Bigg) \\
&= \ln\Bigg(\frac{2^{\frac{1}{6}}n^{\frac{4}{3}}(n-1)^{\frac{3}{2}}}{\pi^{\frac{1}{6}}q^{\frac{n}{3}}}\cdot\frac{2^{\frac{1}{2(n-1)}}m^{\frac{1}{(n-1)}}n^{\frac{1}{(n-1)}}(n-1)^{\frac{1}{2(n-1)}}}{\pi^{\frac{1}{2(n-1)}}q^{\frac{n}{(n-1)}}(2+q\frac{(n-1)^2}{n+1})^{\frac{1}{(n-1)}}} \\
&\qquad \cdot\frac{2\pi^{\frac{3}{2}}n(m-n)}{\sqrt{n-2}}\cdot\Big(\frac{7}{3}+\frac{1}{n-1}\Big)^{(\frac{7}{3}+\frac{1}{n-1})}\Bigg) \\
&\le \ln\left(\frac{2^{\frac{7}{6}}\pi^{\frac{4}{3}}3^3}{q^{\frac{n}{3}+\frac{n}{n-1}}}\cdot\frac{n^{\frac{7}{3}+\frac{1}{n-1}}(n-1)^{\frac{3}{2}+\frac{1}{2(n-1)}}}{\sqrt{n-2}}m^{\frac{1}{n-1}}(m-n)\right) \\
&\le \ln\left(\frac{813}{q^{\frac{n}{3}+\frac{n}{n-1}}}\cdot n^{\frac{10}{3}}m^{\frac{1}{n-1}}(m-n)\right) \\
&= \ln 813 + \left(\tfrac{n}{3}+\tfrac{n}{n-1}\right)|\ln q| + \tfrac{10}{3}\ln n + \tfrac{1}{n-1}\ln m + \ln(m-n)\,.
\end{aligned}$$

Insertion of these upper bounds on $|\ln\bar\varepsilon|$ and $\ln\overline{u}(q,m)$ into (92) delivers

$$\begin{aligned}
&E_{\widetilde{F}}[\overline{K}_U] \le 9\sqrt{m}\,|\ln\bar\varepsilon| + 6\sqrt{m}\ln\overline{u}(q,m) \\
&\le 9\sqrt{m}\left(\ln 813 + \left(\tfrac{n}{3}+\tfrac{n}{n-1}\right)|\ln q| + \tfrac{10}{3}\ln n + \tfrac{1}{n-1}\ln m + \ln(m-n)\right) \\
&\quad + 6\sqrt{m}\left(\ln 6 + |\ln q| + \ln(m+\sqrt{m})\right) \\
&= O(\sqrt{m}(n|\ln q| + \ln n + \ln m))\,.
\end{aligned}$$

Finally we have to prove an upper bound on $\bar{\varepsilon}^{2\bar{\alpha}_3}$ as announced in addition to (89). This is to verify that there is a nonempty region in (0,1) for the parmeter q with $0 < q < \sqrt{1 - \bar{\varepsilon}^{2\bar{\alpha}_3}}$.

$$\begin{aligned}
\varepsilon^{2\bar{\alpha}_3} &= \left(\varepsilon^{(n-1)\bar{\alpha}_3}\right)^{2/(n-1)} = \left((n-1)s_3(\tfrac{7}{3} + \tfrac{1}{n-1})\right)^{-2/(n-1)} \qquad \text{(using (91))} \\
&= \left((n-1)\frac{\sqrt{2}}{\sqrt{\pi}}\frac{mn}{(n-1)^{\frac{3}{2}}}\frac{1}{q^n}\frac{1}{\frac{2}{n-1} + q\frac{n-1}{n+1}}(\tfrac{7}{3} + \tfrac{1}{n-1})\right)^{-2/(n-1)} \\
&= \left(\frac{\sqrt{\pi}}{\sqrt{2}}\frac{(n-1)^{\frac{1}{2}}q^n}{mn}\frac{\frac{2}{n-1} + q\frac{n-1}{n+1}}{(\frac{7}{3} + \frac{1}{n-1})}\right)^{\frac{2}{n-1}} \leq \left(\frac{\sqrt{\pi}}{\sqrt{2}}\frac{(n-1)^{\frac{1}{2}}1}{mn}\frac{\frac{2}{n-1} + \frac{n-1}{n+1}}{(\frac{7}{3} + \frac{1}{n-1})}\right)^{\frac{2}{n-1}} \\
&\leq \left(\frac{11}{10}\frac{1}{mn^{\frac{1}{2}}}\right)^{2/(n-1)} \leq 2m^{-\frac{2}{n-1}} \qquad (93)
\end{aligned}$$

This proves the following theorem

THEOREM 3.3. *For $n \geq 3$, $m > n$ and $q \in (0,1)$ with $\|x_{opt}\| \leq \frac{1}{q}$ the average number of steps of the* BARRIER METHOD *under the uni-RSM is at most $O(\sqrt{m}(n|\ln q| + \ln n + \ln m))$.*

Remark 3.2. We want to point out that the condition $\|x^{opt}\| \leq \frac{1}{q}$ may not be satisfied for all problems (P), but the probability that this condition is satisfied tends to 1 for $m \to \infty$ and an appropiate choice for q.

The probability of the complementary event, i. e. $P(\|x^{opt}\| \geq \frac{1}{q})$ can be bounded from above by $0.8n\binom{m}{n}G(q)^{m-n}$ (analogously to (59)). If $q \in (0,1)$ is a constant independent of n and m, then it is easy to see that $\lim_{m\to\infty} P(\|x^{opt}\| \geq \frac{1}{q}) \to 0$. But we can also choose q as a function of n and m. Let $\bar{q}(\eta) := \sqrt{1 - (\frac{\lambda_{n-1}(\omega_n)(n-1)}{\lambda_{n-2}(\omega_{n-1})}\eta)^{2/(n-1)}}$, $\bar{q}$ is well defined for $\eta \in (0, \frac{4}{5\sqrt{n}})$ and $n \geq 3$. For $\bar{q}(\eta)$ we have $G(\bar{q}(\eta)) \leq 1 - g_2(\bar{q}(\eta)) = \eta$ and $\lim_{\eta\to 0} \bar{q}(\eta) = 1$. Now we look at $\bar{\eta} = \frac{2n \ln m}{m-n}$, which is a proper choice for η if $m \gg n$, and get $G(\bar{q}(\bar{\eta}))^{m-n} \leq (1-\bar{\eta})^{m-n} \leq m^{-2n}$. Therefore

$$P(\|x^{opt}\| \geq \tfrac{1}{q}) \leq 0.8n\binom{m}{n}m^{-2n} \to 0 \text{ for } m \to \infty,\ n \text{ fixed.}$$

4. CONCLUDING REMARKS

Finally, we want to discuss the results of our average case analysis.

We have seen that the complexity of the phase I algorithm is $O(m\sqrt{n})$ in the average case with $m \gg n$. And the result of the average complexity of phase II can be simplified to $O(\sqrt{m}\ln m)$ in the case $m \gg n$. So we have an (overall) average case complexity of $O(m\sqrt{n}) + O(\sqrt{m}\ln m)$.

Looking at the result for phase I we want to remark that the factor m in $O(m\sqrt{n})$ is caused by a rather crude estimation of the phase I barrier function values via the Cauchy-Schwarz inequality in Lemma 2.5. So this bound may be improvable. In phase II the average number of steps is at most $O(\sqrt{m}\ln m)$. Here, the factor

$\sqrt{m}$ depends only on the worst case reduction rate η of the IPM. If we would use a long-step method with reduction rate $\eta = O(1)$ then we get an average number of steps of at most $O(m \ln m)$.

Our phase II result of the average number of steps until termination with our termination PROCEDURE ROUNDING can be compared to the results in Anstreicher et al. [4], Todd et al. [20] and Ye [21] which were based on a probabilistic model in Todd [19]. Ye proves an average number of steps of at most $O(\sqrt{m} \ln m)$ (in our notation) with high probability, using the finite termination scheme (with projections) as described in Mehrotra and Ye [14]. Anstreicher et al. [4] showed that an upper bound of $O(m \ln m)$ for the average number of steps of an infeasible interior point method combined with the same termination scheme holds. So we see that all the results are of the same order although they were based on different probabilistic models and different termination schemes.

What is not done yet, is the average case analysis of a single step of IPMs. There is only one approach to this kind of analysis by Mizuno, Todd and Ye [15]. They assumed some distributions on internal figures in each iteration and assumed independence between different iterations. But these assumptions are inconsistent with the assumption of a distribution for the original problem data. So, their probabilistic analysis is not rigorous. The main difficulty of such a "single step"-average case analysis is the dependency between successive iterations. The rigorous handling of the dependency will play a leading part in the average case analysis of a single step. And this part of the analysis will hopefully complete the average case analysis of IPMs.

Evidently, the behaviour in both phases is significantly better than our so far ensured results show. And a potential analysis of the "average reduction" may demonstrate the reason why this reduction could be much more effective in reality than in the worst case. The usage of the "worst case reduction rate" had been sufficient to prove the strong polynomiality, but there is still a gap to the real behaviour. Such an efficient analysis of the reduction rate will give us a chance to compare the behaviour of interior-point-methods with that of the simplex-method and to show that IPMs perform very well in that comparison.

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. I. Adler, N. Megiddo, and M. J. Todd. New results on the average behavior of simplex algorithms. *Bulletin of the American Mathematical Society*, 11:378–382, 1984.

3. K.M. Anstreicher, J. Ji, F.A. Potra, and Y. Ye. Average performance of a self-dual interior-point algorithm for linear programming. In P.M. Pardalos, editor, *Complexity in Numerical Optimization*, pages 1–15. World Scientific Publishing Co., 1993.

4. K.M. Anstreicher, J. Ji, F.A. Potra, and Y. Ye. Probabilistic analysis of an infeasible primal-dual algorithm for linear programming. *Mathematics of Operations Research*, 24:176–192, 1999.

5. M.S. Bazaraa, H.D. Sherali, and C.M. Shetty, *Nonlinear programming. Theory and algorithms. 2. ed.* Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley, New York, 1993.
6. K. H. Borgwardt. *The Simplex Method — A Probabilistic Analysis.* Springer-Verlag, Heidelberg, 1987.
7. K. H. Borgwardt. Ideas leading to a better bound on the average number of pivot steps for solving an LP. In *Operations Research Proceedings 1997*, pages 1–12. Springer Verlag, 1998. GOR-Tagung, Jena.
8. K. H. Borgwardt. A sharp upper bound for the expected number of shadow-vertices in LP-polyhedra under orthogonal projection on two-dimension planes. *Mathematics of Operations Research*, 24(4):925–984, 1999.
9. D. den Hertog. *Interior Point Approach to Linear, Quadratic and Convex Programming: Algorithms and Complexity*, volume 277 of *Mathematics and Its Applications.* Kluwer Academic Publishers, 1994.
10. 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.
11. P. Huhn. A phase-1-algorithm for interior-point-methods: Worst-case and average-case behaviour. In *Operations Research Proceedings 1998*, pages 103–112. Springer Verlag, 1999. GOR-Conference, Zurich.
12. P. Huhn and K.H. Borgwardt. An upper bound for the average number of iterations required in phase II of an interior-point-method. In *Operations Research Proceedings 1997*, pages 19–24. Springer Verlag, 1998. GOR-Conference, Jena.
13. N. K. Karmarkar. A new polynomial–time algorithm for linear programming. *Combinatorica*, 26:373–395, 1984.
14. S. Mehrotra and Y. Ye. Finding an interior point in the optimal face of linear programs. *Mathematical Programming*, 62A:497–515, 1993.
15. S. Mizuno, M.J. Todd, and Y. Ye. On adaptive-step primal-dual interior-point algorithms for linear programming. *Mathematics of Operations Research*, 18:964–981, 1993.
16. C.H. Papadimitriou and K. Steiglitz. *Combinatorial Optimization, Algorithms and Complexity.* Prentice-Hall, New Jersey, 1982.
17. S. Smale. On the average number of steps of the simplex method. *Mathematical Programming*, 27:241–262, 1983.
18. M.J. Todd. Polynomial Expected Behavior of a Pivoting Algorithm for Linear Complementarity and Linear Programming Problems. *Mathematical Programming*, 35:173–192, 1986.
19. M.J. Todd. Probabilistic models for linear programming. *Mathematics of Operations Research*, 16:671–693, 1991. Erratum, 23:767-768.
20. M.J. Todd, L. Tuncel and Y. Ye. Characterizations, bounds, and probabilistic analysis of two compexity measures for linear programming. *Mathematical Programming*, 90:59-69, 2001.
21. Y. Ye. Toward probabilistic analysis of interior-point algorithms for linear programming. *Mathematics of Operations Research*, 19:38–52, 1994.
22. Y. Ye. *Interior Point Algorithms: Theory and Analysis.* Wiley-Interscience. Wiley, 1997.