A sharp upper bound for the expected number of shadow vertices in LP-polyhedra under orthogonal projection on two-dimensional planes

Karl-Heinz Borgwardt

A Sharp Upper Bound For The Expected Number Of Shadow Vertices In LP-Polyhedra Under Orthogonal Projection On Two-Dimensional Planes [*][†]

Quadratic in the number of variables and sublinear in the number of restrictions of LPs, generated under the rotation symmetry model

Karl Heinz Borgwardt

Institut für Mathematik der Universität Augsburg

86135 Augsburg, Universitätsstraße 14

E-Mail: borgwardt@math.uni-augsburg.de

September 23, 1999

Abstract

Consider a set $\{a_1, \ldots, a_m, u, v\} \subset \mathbb{R}^n$ satisfying a mild nondegeneracy condition. When the polyhedron $X = \{x \mid a_1^T x \leq 1, \ldots, a_m^T x \leq 1\} \subset \mathbb{R}^n$ is orthogonally projected on the two-dimensional plane $\text{lin}(u, v)$, then some of its vertices may be mapped on vertices of the two-dimensional image of X. These vertices of X will be called shadow vertices under that projection. Their number will be denoted by S.

This figure S has an outstanding relevance for the complexity of a special Simplex-Method, the dimension-by-dimension method, for solving LPs of the type

$$\text{maximize } v^T x \text{ s.t. } a_1^T x \leq 1, \ldots, a_m^T x \leq 1 \text{ where } x, v, a_1, \ldots, a_m \in \mathbb{R}^n \text{ and } m \geq n.$$

This method applies the "shadow vertex algorithm" in $n - 1$ stages of increasing dimension, by constructing a Simplex-Path that visits only shadow vertices. In each stage specific versions of the vectors u and v are used to define the projection plane and to control the application of the "shadow vertex algorithm." So S delivers an upper bound for the length of such a path in one stage.

We want to study the average-case complexity of these algorithms and are therefore interested in $E_{m,n}(S)$, the expected value of S. For that purpose we assume a distribution of the linear programming problems corresponding to the rotation symmetry model (RSM):

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

---

[*]AMS-Classification: 90C05, 68Q25, 60D05, 52A22

[†]Key-Words: Linear Programming, Average-Case Complexity of Algorithms, Stochastic Geometry

In this paper we improve the upper bounds for $E_{m,n}(S)$ from our polynomiality-proofs of 1982 and 1987. There we had found an upper bound of $m^{\frac{1}{n-1}} \cdot n^3 \cdot Const.$, valid for all RSM-distributions and all pairs $(m \geq n)$.

Based on a refined evaluation technique for space-angles, on a reorganization and on substantial modifications of the old proof, we are now able to confirm a bound

$$E_{m,n}(S) \leq m^{\frac{1}{n-1}} \cdot n^2 \cdot Const. \text{ for all pairs } (m,n) \text{ with } m \geq n.$$

This result had been desired since 1982, because then a bound $m^{\frac{1}{n-1}} \cdot n^2 \cdot Const.$ had been derived exclusively for the situation, where n is fixed and m tends to infinity. All the time the question had been open, whether the asymptotical behaviour would be better than the behaviour in moderate dimensions, or whether the bound for the moderate case was crude.

In addition, we know – from an asymptotic analysis in 1987 that for a special RSM-distribution, namely the uniform distribution on the unit sphere, there is a lower bound for $E_{m,n}(S)$ of type $m^{\frac{1}{n-1}} \cdot n^2 \cdot Const.$ The combination of this result and the new upper bound shows that the new bound is sharp.

1 The number of shadow vertices and its algorithmic relevance

The motivation for writing this paper and for doing this study came from the question for the (average) complexity of a solution method for canonical LPs of the type

$$\text{maximize } v^T x \text{ subject to } a_1^T x \leq 1, \ldots, a_m^T x \leq 1, \tag{1}$$
$$\text{where } x, v, a_1, \ldots, a_m \in \mathbb{R}^n \text{ and } m \geq n.$$

For the purpose of averaging we assume a distribution of the linear programming problems corresponding to the rotation symmetry model (RSM):

$$\text{The vectors } a_1, \ldots, a_m, v \text{ are distributed on } \mathbb{R}^n \setminus \{0\} \text{ indepen-} \tag{2}$$
$$\text{dently, identically and symmetrically under rotations.}$$

This question has been investigated in earlier papers (cf. Borgwardt (1982b), Borgwardt (1987), etc.). The new bounds on the average number of shadow vertices, as derived in this paper, will help to improve those results on the average complexity of (1).

The core of the old and the new analysis of the average complexity is the answer to a certain subproblem, which can be given more precisely in this paper. In this first section, we will explain this subproblem and its relevance for the average complexity. After that, we shall totally focus on that subproblem, which can be seen as a mathematical problem in its own right.

We use the abbreviations lin, conv, cone, dist for linear hull, convex hull, convex cone and distance respectively. When we write $\text{conv}(x, y, \ldots)$ we mean the set of convex combinations formed with the help of $x, y, \ldots$. The same applies to cone, lin.

The question attacked directly in this paper can be seen from three different points of view. Suppose that $\{a_1, \ldots, a_m, u, v\} \subset \mathbb{R}^n$ with $m \geq n$ is a set of given points, which meets the following nondegeneracy-assumption

$$\text{Nondegeneracy is valid, if any } n \text{ vectors out of } \{a_1, \ldots, a_m, u, v\} \text{ are linearly inde-} \tag{3}$$
$$\text{pendent, and if any } n+1 \text{ vectors out of } \{a_1, \ldots, a_m, u, v\} \text{ are in general position.}$$

Let the subset $\{a_1, \ldots, a_m\}$ and the subset $\{u, v\}$ respectively induce a (nonempty) polyhedron and a two-dimensional plane, i.e.

$$X := \{ x \mid a_1^T x \leq 1, \ldots, a_m^T x \leq 1 \} \subset R^n \quad \text{and} \quad \text{lin}(u, v). \tag{4}$$

We write $\{a_1, \ldots, a_m; u, v\}$ to separate both parts and to indicate that in both subsets internal permutations are irrelevant for the resulting configuration of X and $\text{lin}(u, v)$.
Consider an orthogonal projection

$$\Pi_{u,v} : \mathbb{R}^n \to \text{lin}(u, v), \quad \text{mapping} \quad X \text{ into } \text{lin}(u, v). \tag{5}$$

$\Pi_{u,v}(X)$, the image of X, is a two-dimensional polyhedron. And some of the vertices of X may be mapped on vertices of $\Pi_{u,v}(X)$, which motivates the name "shadow vertices," precisely:

$$\text{A vertex } x_* \text{ of } X \text{ is called shadow vertex of } X \text{ with regard to } u \text{ and}$$
$$v, \text{ if the orthogonal projection } \Pi_{u,v} \text{ on } \text{lin}(u, v) \text{ maps } x_* \text{ on a vertex} \tag{6}$$
$$\text{of the two-dimensional image } \Pi_{u,v}(X) \text{ of } X \text{ in } \text{lin}(u, v).$$

3

1) A purely geometrical question is:
How many shadow vertices are induced by the projection $\Pi_{u,v}$ mapping X on $\mathrm{lin}(u,v)$?

We regard this number $S = S(a_1, \ldots, a_m; u, v)$ as a functional of the given $(m+2)$-element set of input vectors. Note that the concentration on the configuration $m \geq n$ is natural, because for $m < n$ the number of shadow vertices is 0, as there cannot be vertices at all.

2) From the view of parametric programming the interpretation is as follows:
Consider the set of all LPs

$$\left\langle \text{maximize } w^T x \quad \text{subject to } x \in X \right\rangle \text{ with } 0 \neq w \in \mathrm{lin}(u,v). \tag{7}$$

Here we ask for the number of all vertices of X being extremal with respect to any objective $w^T x$, such that $0 \neq w \in \mathrm{lin}(u,v)$ is an arbitrary nontrivial linear combination of u and v. This extremality is just an equivalent formulation of the shadow vertex property. So the answer to that question is again $S = S(a_1, \ldots, a_m; u, v)$.
At this point we may restrict our set of possible objective gradients to a subcone of $\mathrm{lin}(u,v)$, as for instance to $\mathrm{cone}(u_*, v_*)$, where $u_*, v_* \in \mathrm{lin}(u,v)$:
Solve all LPs of the form

$$\left\langle \text{maximize } w^T x \quad \text{subject to } x \in X \right\rangle \text{ with } 0 \neq w \in \mathrm{cone}(u_*, v_*) \subset \mathrm{lin}(u,v). \tag{8}$$

Note that $u_* = u$ and $v_* = v$ would describe a special case here.
For these two parametric programming problems, we may count those vertices, that are optimal at least for one choice of w. We know

$$S = S(a_1, \ldots, a_m; u, v) = \{x_* | x_* \text{ optimal vertex of } X \text{ for some } w \in \mathrm{lin}(u,v)\}.$$

Accordingly we define the cardinality of the set of optimal vertices for the subcone as

$$s_{u_*, v_*}(a_1, \ldots, a_m; u, v) := \#\{x_* | x_* \text{ optimal vertex of } X \text{ for some } w \in \mathrm{cone}(u_*, v_*)\}.$$

And now it is clear that

$$s_{u_*, v_*}(a_1, \ldots, a_m; u, v) \leq S \qquad \text{and} \qquad s_{u,v}(a_1, \ldots, a_m; u, v) \leq S. \tag{9}$$

So far, we understand that S helps to bound qualitative figures on parametric programming problems (as the number of "vertex-solutions").

3) Now we come to the third point of view, which asks for the complexity of algorithms. Suppose that we have vectors $u_*, v_* \in \mathrm{lin}(u,v)$, and that a vertex x_1 of X, maximizing $u_*^T x$ on X, is available. Then x_1 must be a shadow vertex with respect to $\mathrm{lin}(u,v)$. Assume that our aim is to maximize $v_*^T x$ on X, starting from x_1. An optimal vertex for that problem would be a shadow vertex, as well. It is clear that in $\mathrm{lin}(u,v)$ it will be possible to improve $v_*^T x$ by running along the boundary of $\Pi_{u,v}(X)$ over a connected path, which improves the objective monotonically and visits only the images of shadow vertices. Of course, the

number of visited shadow vertices equals $s_{u_*,v_*}(a_1,\ldots,a_m;u,v)$ and is bounded from above by S.

But now it is important that – due to nondegeneracy – the preimage of that path (before being mapped by $\Pi_{u,v}$) is a Simplex-Path in the original space and on the full dimensional polyhedron X over the corresponding shadow vertices, improving the objective $v_*^T x$.

This is just the geometrical idea for the shadow vertex algorithm.

And it is clear that this fits into the general framework of the Simplex-Method for solving (1), as it defines a variant for Phase II.

It is easy to construct that path. We start in x_1, the optimal vertex for $u_*^T x$, where $0 \neq u_* \in \text{lin}(u,v)$. To optimize $v_*^T x$ with $0 \neq v_* \in \text{lin}(u,v)$, we create a family of objectives $(u_* + \gamma \cdot v_*)^T x$ with $\gamma \in [0,\infty)$, start at $\gamma = 0$, increase γ continuously and follow the sequence of temporarily optimal vertices. (For implementation it suffices to update the representation of both objective gradients u_*, v_* under the current basis and to observe the γ-weighted combination of both. At certain values of γ vertex-exchanges (pivot steps) will be necessary to stay optimal. Then the corresponding representation-coordinates determine the pivot element.) This is just a realization of the parametric Simplex-Variant of Gass and Saaty, adapted to the current situation. Hence the number of shadow vertices S is a natural upper bound for the number of vertices visited on this path from x_1 to x_s, which means $s = s_{u_*,v_*} \leq S$.

This is the point where we can finally turn to average-case analysis.

In this paper we modify the polynomiality-proof of Borgwardt (1987) and improve the upper bound for $E_{m,n}(S)$, the expected number of shadow vertices (under orthogonal projection on planes $\text{lin}(u,v)$) of feasible polyhedra X in linear programming problems of the type

$$
\begin{aligned}
&\text{maximize } v^T x \quad \text{s. t. } x \in X = \{x \mid a_1^T x \leq 1, \ldots, a_m^T x \leq 1\} \\
&\text{with } x, u, v, a_1, \ldots, a_m \in \mathrm{I\!R}^n, \\
&\text{where } u \text{ is an augmenting vector for defining the projection-plane } \text{lin}(u,v).
\end{aligned}
\tag{10}
$$

We assume a distribution of input sets under the rotation symmetry model (RSM) by extending the definition (2) :

$$
\begin{aligned}
&\text{The vectors } a_1, \ldots, a_m, v \text{ and the auxiliary vector } u \text{ are distributed on} \\
&\mathrm{I\!R}^n \setminus \{0\} \text{ independently, identically and symmetrically under rotations.}
\end{aligned}
\tag{11}
$$

This assumption justifies concentration on the nondegenerate cases, which are almost sure.

We remark that under RSM for problems of type (10) holds: $E_{m,n}(S) \approx 4 \cdot E_{m,n}(s_{u,v})$.

Note that rotation symmetry does not apply to the distribution of the (total) input sets $\{a_1, \ldots, a_m; u, v\} \in \mathrm{I\!R}^{(m+2) \times n}$, but for the identical distributions of the single points a_i over $\mathrm{I\!R}^n$. These distributions characterize different cases in our stochastic analysis (we call that "special distributions"). When we speak of "our distribution," we mean one of those $\mathrm{I\!R}^n$-distributions. For denoting the corresponding distribution-function we use F. If our distribution has a density, then the corresponding density function will be denoted

by f. Without loss of generality we can suppose, that our distribution has a density (cf. Borgwardt (1987)) and that its support is the unit ball.

Under rotation symmetry , the characterization of these special distributions is very simple, because the density of different points with identical Euclidean norm must coincide. Hence every direction $d \in \mathbb{R}^n$ will be attained with the same likelihood. Note that every point $x \in R^n$ can be written as $\|x\| \cdot d(x)$ with $\|d(x)\| = 1$. So the special distribution can be uniquely characterized by specification of the so-called "radial distribution function" $\mathcal{F}$.

$$\text{Let } \mathcal{F} \text{ stand for the probability } \mathcal{F}(r) := P(\|x\| \leq r) \text{ for } r \in [0, \infty]. \tag{12}$$

We introduce notation for the unit ball resp. the unit sphere in $\mathbb{R}^k$, namely Ω_k resp. ω_k

$$\Omega_k := \{x \mid \|x\| \leq 1, \ x \in \mathbb{R}^k\} \qquad \text{and} \qquad \omega_k := \{x \mid \|x\| = 1, \ x \in \mathbb{R}^k\}. \tag{13}$$

When we use $\Omega_k(r)$ or $\omega_k(r)$, we mean the corresponding ball or sphere of radius r.

λ_k stands for the k-dimensional Lebesgue-measure. We note that

$$\lambda_n\{\Omega_n\} = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n+2}{2})} \ , \quad \lambda_{n-1}\{\omega_n\} = \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}, \quad \text{and} \quad n \cdot \lambda_n\{\Omega_n\} = \lambda_{n-1}\{\omega_n\}. \tag{14}$$

When we write $Const.$ or C, then we mean a constant real value, independent of m and n. $C(n)$ will denote n-dependent reals, which are independent of m.

In Borgwardt (1982b) and Borgwardt (1987) it was shown that there is an upper bound for all $m \geq n$ and for all rotation symmetric distributions of the form

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

In Borgwardt (1994a) resp. Borgwardt (1994b) we improved that result to

$$E_{m,n}(S) \leq m^{\frac{1}{n-1}} \cdot n^{\frac{5}{2}} \cdot Const.$$

by analyzing the relation between certain spherical angles more precisely. In this paper we give a new proof, based on techniques of Borgwardt (1987). The main ideas of Borgwardt (1994a) resp. Borgwardt (1994b) are included and combined with more precise calculations replacing crude estimations of Borgwardt (1987). This leads to another saved factor of $O(\sqrt{n})$. We obtain

Theorem 1

For all distributions according to our rotation symmetry model (11) we have

$$E_{m,n}(S) \leq m^{\frac{1}{n-1}} \cdot n^2 \cdot Const. \qquad \square \tag{15}$$

Theorem 1 is a long desired final step in closing the gap between observed asymptotic results and theoretical bounds for moderate dimensions. From Borgwardt (1982a) we had known an (asymptotic) upper bound of

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

6

For 15 years that discrepancy between bounds for general dimension-pairs and bounds for the asymptotic case remained an open problem. Also in numerical experiments (compare Borgwardt, Damm, Donig and Joas (1993)) the mean values did not exceed $m^{\frac{1}{n-1}} \cdot n^2 \cdot 2$.

Another theoretical enforcement for the hope for quadratic behaviour in the lower dimension came from Küfer (1996). For the special case of uniform distribution on ω_n, he proved the existence of an expansion for $E_{m,n}(S)$ in a series of products. From the very complicated calculation of the first coefficients he could conclude truncated-series-estimations, as e.g.

$$E_{m,n}(S) = \quad C_1(n) \cdot m^{\frac{1}{n-1}} + C_2(n) \cdot m^{-\frac{1}{n-1}} + O(m^{-\frac{2}{n-1}}) \quad \text{for fixed } n \geq 3$$

with $C_1(n), C_2(n)$ such that $C_1(n) \to 2n^2$ for $n \to \infty$ and $C_2(n) < 0 \; \forall \; n$.

Unfortunately, it could not be proven that the first terms dominate the rest-term already for moderate m. So again, only the asymptotic bound for this distribution was confirmed.

On the other hand, also an asymptotic <u>lower</u> bound for this special distribution, the uniform distribution on the unit sphere, is known from Borgwardt (1987). This is

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

Seeing this in the light of Theorem 1 we conclude

Corollary 1

The bound of Theorem 1 is sharp in that sense, that it cannot be reduced uniformly for general distributions (with exception of lower constants). We have constants $C_1, C_2 > 0$, such that

$$C_1 \geq \sup_{F \in RSM} \; \sup_n \; \limsup_{m \to \infty} \; \frac{E_{m,n}^F}{m^{\frac{1}{n-1}} \cdot n^2} \geq C_2, \tag{16}$$

where $F \in RSM$ indicates that the distribution fits into the rotation symmetry model. □

For applications of the shadow vertex algorithm in the dimension-by-dimension method it will be necessary to study the situation where u resp. v are fixed vectors, i.e. to replace u by a fixed vector $\overline{u}$ and v by $\overline{v}$. Then the question arises, whether our results will remain valid under that fixation or conditioning.

Corollary 2

The bound in Theorem 1 applies even to the conditional expectation values, when u is replaced by a fixed vector $\overline{u}$ or when both random vectors (u, v) are replaced by a fixed pair $(\overline{u}, \overline{v})$. □

<u>Proof</u>

In RSM it is possible to rotate $\mathbb{R}^n$, such that u is sent to $\overline{u}$. Such a rotation (applied to all input vectors) has no impact on S and on all the stochastic figures in RSM. For the same reason, we can even fix the total plane $\lin(u, v)$ to $\lin(\overline{u}, \overline{v})$, because our figure S depends on the position of the plane exclusively, not on the specific location of u and v in this plane. □

Note that conditioning is possible only now, after averaging under RSM. As long as we consider the deterministic case, this would be absolutely false. And we mention, that the

second claim in the corollary cannot be applied to expected values of figures like $s_{u,v}$, where not only the position of the plane, but also the relative position of v to u is relevant.

Let us recall the motivating quantity, the average complexity for solving the canonical LP (1) completely. We employ a special version of the Simplex-Method, the dimension-by-dimension method and analyze its behaviour. Our three interpretations for the subproblem will help to perform that analysis. Now we should clarify how we organize Phase I in our LP (1). In order to make use of the shadow vertex idea and to meet the stochastic conditions of our model in all stages of the method, we employ the following (n-stage) method. Let us denote

$$\Pi_k(x) := \begin{pmatrix} x^1 \\ \vdots \\ x^k \end{pmatrix} \text{ for } x = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} \in \mathbb{R}^n. \tag{17}$$

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

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

$$\text{maximize } \Pi_k(v)^T \Pi_k(x) \quad \text{s. t.} \quad \Pi_k(a_i)^T \Pi_k(x) \le 1 \text{ for } i = 1, \ldots, m$$
$$\text{where} \quad x, v, a_1, \ldots, a_m \in \mathbb{R}^n \text{ and } m \ge n. \tag{18}$$

The feasible region of I_k will be called X_k. It is – due to nondegeneracy – a cute polyhedron (containing vertices). Note that I_n coincides with the original LP (1). For a complete solution we may apply the following algorithm.

Dimension-By-Dimension Method

1. Set $k = 1$ and find a vertex of X_1.

2. If existing, find the optimal vertex $(\hat{x}^1)$ of I_1. Else go to 6).

3. If $k = n$, then go to 7). Set $k = k + 1$.

4. For a $k \in \{2, \ldots, n\}$ the solution of I_{k-1} may be available. We call it

   $\begin{bmatrix} \hat{x}^1 \\ \vdots \\ \hat{x}^{k-1} \end{bmatrix}$. Then $\begin{bmatrix} \hat{x}^1 \\ \vdots \\ \hat{x}^{k-1} \\ 0 \end{bmatrix}$ is feasible for I_k on an edge of X_k. Determine $\begin{bmatrix} \tilde{x}^1 \\ \vdots \\ \tilde{x}^k \end{bmatrix}$, which

   denotes a vertex incident to that edge. It is even a shadow vertex under projection on the plane $\lim (\Pi_k(e_k), \Pi_k(v))$.

5. Use $\lim (\Pi_k(e_k), \Pi_k(v))$ as projection plane and start the shadow vertex algorithm in

   $\begin{bmatrix} \tilde{x}^1 \\ \vdots \\ \tilde{x}^k \end{bmatrix}$ to find an optimal vertex $\begin{bmatrix} \hat{x}^1 \\ \vdots \\ \hat{x}^k \end{bmatrix}$ for I_k. Go to 3) if $\hat{x}$ exists, else to 6).

6. A solution of the complete problem does not exist. STOP

7. The vector $\hat{x} \in \mathbb{R}^n$ is the solution of I_n (the original problem). STOP $\qquad \square$

Now it should be clear that in this method the shadow vertex algorithm is applied in dimensions $k = 2, \ldots, n$ and that these applications determine its complexity, since the effort for the other calculations is ignorable. So we have in stage k objective gradients $u_*^{(k)}, v_*^{(k)}$ in $\operatorname{lin}\left(\Pi_k(e_k), \Pi_k(v)\right)$ and by use of the shadow vertex algorithm the "distance" between the $u_*^{(k)}$-solution and the $v_*^{(k)}$-solution has to be bridged. Hence all the estimations and results mentioned above apply and we have

$$\sum_{k=2}^{n} s_{u_*^{(k)}, v_*^{(k)}}(a_1, \ldots, a_m; u, v) \leq \sum_{k=2}^{n} S^{(k)}. \tag{19}$$

It makes sense to concentrate on one stage, namely stage $k = n$, and to analyze $S^{(n)}(a_1, \ldots, a_m; u, v)$. This is, what we call Phase II of our method. Phase I consists of the stages $k = 1, \ldots, n-1$. We ignore the stage-number n and write S. Also the role of $u_*^{(k)}, v_*^{(k)}$ becomes irrelevant now, because we concentrate on the upper bounds.

Our average-case analysis of the dimension-by-dimension method profits from the new result. Note that the (truncated) input vectors of I_k in (18), as $\Pi_k(a_i), \Pi_k(v)$, are distributed according to RSM, too. So the stochastic assumptions apply to each use of the shadow vertex algorithm in step 5) with u replaced by $\Pi_k(e_k)$ and v replaced by $\Pi_k(v)$.

For bounding the expected value of the total number of pivot steps s_t, we simply have to sum up over all $n-1$ applications of the shadow vertex algorithm. A simple upper bound would result from summation over all values $E_{m,k}(S)$ for $k = 2, \ldots, n$, resp. their upper bounds given in Theorem 1 with n replaced by k. But that upper bound on the expected value of s_t creates a false impression, as we would ignore that most of the RSM-distributions over $\mathbb{R}^k$ are no possible projection-distributions from $\mathbb{R}^n$ under Π_k. Remember that the input vectors of I_k are truncated versions of randomly generated vectors in $\mathbb{R}^n$. So they follow a so-called "projection-distribution." This shows that we have to deal only with a <u>subset</u> of the distributions studied for Theorem 1. With elementary techniques on projected distributions presented in (Borgwardt (1987), pp. 179-186) we can exploit our new result and show

Theorem 2
For the expected value of the number of shadow vertices in stage k we know that

$$E_{m,n}(S^{(k)}) \leq m^{\frac{1}{n-1}} \cdot n \cdot k \cdot Const. \tag{20}$$

And the expected number of pivot steps s_t of the dimension-by-dimension method for complete solution of (1) satisfies

$$E_{m,n}(s_t) \leq m^{\frac{1}{n-1}} \cdot n^3 \cdot Const. \qquad \square \tag{21}$$

2   Basics of the evaluation

The task of this section is to give a report on the fundamentals of the approach in Borgwardt (1987), as far as they can be used for our new, refined study, too. Recall our assumption of nondegeneracy from (3), which assures the validity of the following facts.
Since we have to count a certain subset of the vertices of X and since the impact of the random data $a_1, \ldots, a_m$ on X is rather indirect and complicated to describe, we can profit from shifting our analysis from the primal space to the corresponding dual space (with respect to our scalar product), where the vectors a_i are generated. There we find the polar polyhedron to X, defined by $\{y \in R^n \mid y^T x \le 1 \ \forall \ x \in X\}$. It can easily be seen that in our case the polar polyhedron coincides with the following polytope:

$$Y \ := \ \text{conv}(0, a_1, \ldots, a_m). \tag{22}$$

Since the face-lattice of Y reflects the essential properties of the face-lattice of X in an antisymmetric way, we can carry out our analysis here, as well. The key to our translation lies in the following one-to-one correspondence.

$$\text{Let} \ \triangle \ \text{be an } n\text{-element index set} \ \ \{\triangle^1, \triangle^2, \ldots, \triangle^n\} \ \subset \ \{1, \ldots, m\}. \tag{23}$$

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

$$a_{\triangle^1}^T x \ = \ 1, \ldots, a_{\triangle^n}^T x \ = \ 1. \tag{24}$$

That means that in $x_\triangle$ exactly n restrictions are active, hence $x_\triangle$ is one of $\binom{m}{n}$ *basic solutions* of (1). On the other side, $\triangle$ uniquely defines a simplex $\text{conv}(a_{\triangle^1}, \ldots, a_{\triangle^n})$ in Y. Only the $\binom{m}{n}$ basic solutions are candidates for being shadow vertices of X. For becoming a shadow vertex, $x_\triangle$ must simultaneously satisfy two conditions (cf. Borgwardt (1987),p.64)

1. it must be a vertex of X, which is equivalent to $a_i^T x_\triangle \ \le \ 1 \ \forall \, i \notin \triangle$,   (25)
2. it must optimize some objective $w^T x$ on X, where $w \in \text{lin}(u, v) \cap \omega_n$.   (26)

Due to the Lemma of Farkas this can equivalently be expressed with the polar polyhedron. $\text{conv}(a_{\triangle^1}, \ldots, a_{\triangle^n})$ corresponds to a shadow vertex $x_\triangle$, iff simultaneously (cf. Borgwardt (1987),pp.69–74)

1. $\text{conv}(a_{\triangle^1}, \ldots, a_{\triangle^n})$ is a facet of $Y \ = \ \text{conv}(0, a_1, \ldots, a_m)$,   (27)
2. $\text{conv}(a_{\triangle^1}, \ldots, a_{\triangle^n}) \ \cap \ \text{lin}(u, v) \ne \emptyset$.   (28)

As S denotes the number of shadow vertices, we can use that characterization in the space of the a_i's to formulate an integral representation for $E_{m,n}(S)$ by means of stochastic geometry (cf. Borgwardt (1987),p.122). It describes the probability, that a certain basic solution of problem (10) actually is a shadow vertex, and it multiplies that probability with the number of candidates, i.e. $\binom{m}{n}$.

Throughout the paper, I will stand for an indicator function and P for probability.

$$E_{m,n}(S) = \binom{m}{n} \cdot \int\limits_{\mathbb{R}^n} \cdots \int\limits_{\mathbb{R}^n} I(\operatorname{conv}(a_1,...,a_n) \text{ is a facet of } \operatorname{conv}(0,a_1,...,a_m)) \cdot$$
$$\cdot I(\operatorname{conv}(a_1,...,a_n) \text{ is intersected by } \operatorname{lin}(u,v)) \cdot$$
$$\cdot f(a_1)\cdots f(a_n)f(a_{n+1})\cdots f(a_m)f(u)f(v)\, da_1 \ldots da_m\, du\, dv \quad (29)$$

The intersection-event of the plane with a simplex $\operatorname{conv}(a_1,\ldots,a_n)$ can easier be described via intersection-events with side-simplices $\operatorname{conv}(a_1,\ldots,a_{i-1},a_{i+1},\ldots,a_n)$. We make use of:

Each intersection of $\operatorname{lin}(u,v)$ with $\operatorname{conv}(a_1,\ldots,a_n)$ produces exactly two intersected simplex sides of type $\operatorname{conv}(a_1,\ldots,a_{i-1},a_{i+1},\ldots,a_n)$. And for a fixed $(n-2)$-dimensional simplex $\operatorname{conv}(a_1,\ldots,a_{n-1})$ we have due to RSM (cf. Borgwardt (1987),p.129)

$$P(\operatorname{lin}(u,v) \text{ intersects } \operatorname{conv}(a_1,\ldots,a_{n-1})) = 2 \cdot W(a_1,\ldots,a_{n-1}), \quad (30)$$

where $W(a_1,\ldots,a_{n-1})$ is defined as the spherical angle of $\operatorname{cone}(a_1,\ldots,a_{n-1})$, i.e.

$$W(a_1,\ldots,a_{n-1}) := \frac{\lambda_{n-1}\{\operatorname{cone}(a_1,\ldots,a_{n-1}) \cap \Omega_n\}}{\lambda_{n-1}\{\Omega_{n-1}\}} = \frac{\lambda_{n-2}\{\operatorname{cone}(a_1,\ldots,a_{n-1}) \cap \omega_n\}}{\lambda_{n-2}\{\omega_{n-1}\}} \quad (31)$$

Hence $\quad P(\operatorname{lin}(u,v) \text{ intersects } \operatorname{conv}(a_1,\ldots,a_n)) = \sum\limits_{i=1}^{n} W(a_1,\ldots,a_{i-1},a_{i+1},\ldots,a_n). \quad (32)$

Exploiting the rotation symmetry of f, (29) amounts to an integral-formula (cf. Borgwardt (1987),p.124)

$$E_{m,n}(S) = \binom{m}{n} n \int\limits_{\mathbb{R}^n} \ldots \int\limits_{\mathbb{R}^n} G(h[a_1,...,a_n])^{m-n} W(a_1,...,a_{n-1}) f(a_1) \cdots f(a_n) da_1 ... da_n. \quad (33)$$

Here, $h[a_1,...,a_n]$ is the distance from the origin to the hyperplane containing $\{a_1,...,a_n\}$. We use $G(h)$ for the marginal distribution function of our distribution along the n-th coordinate and $g(h)$ for the marginal density function. That means

$$G(h) := P\{x \mid x^n \le h\} = \int\limits_{\mathbb{R}^n} I(x^n \le h) f(x) dx = \int\limits_{-1}^{h} g(q) dq.$$

The integral representation achieved by that way can be simplified significantly by two coordinate-transformations (cf. Borgwardt (1987),pp.134-137 and 145-147). In the first transformation we perform a rotation which puts all the basic points $a_1,\ldots,a_n$ into a hyperplane $\{x \mid x^n = h\}$, and we substitute the points a_i, $i = 1,\ldots m$ by vectors b_i, $i = 1,\ldots m$. After that we have $b_i^n = h \ge 0 \quad \forall i = 1,\ldots n$. Since we used a rotation, we have preserved all stochastic features because of RSM.

The same holds for our second transformation, which is a rotation in $\mathbb{R}^{n-1}$. Now we replace the vectors $b_1,\ldots,b_m$ by vectors $c_1,\ldots,c_m$ such that for $n-1$ of the basic vectors, namely for $c_1,\ldots,c_{n-1}$, we have coincidence even in the $(n-1)$th coordinate, i.e. $c_i^{n-1} = \Theta \ge 0 \quad \forall i = 1,\ldots n-1$. Still, of course $c_i^n = h \ge 0 \quad \forall i = 1,\ldots n$.

11

Lower indices in connection with vectors will distinguish different vectors, whereas upper indices are used to characterize their components. If such a component appears with an exponent, then the component is included in brackets, e.g. $(c^k)^2$ or $[c^k]^2$.

So $c_1, \ldots, c_n$ are (column-)vectors in $\mathbb{R}^n$, and $c = (c^1, \ldots, c^n)^T$.

$\bar{c} = (c^1, \ldots, c^{n-1})^T$ and $\bar{\bar{c}} = (c^1, \ldots, c^{n-2})^T$ give the corresponding truncated vectors.

The result of these two rotations is of the following form:

$$
\begin{aligned}
E_{m,n}(S) \;=\; & \binom{m}{n} \cdot n \cdot \{(n-2)!\}^2 \cdot \lambda_{n-1}\{\omega_n\} \cdot \lambda_{n-2}\{\omega_{n-1}\} \cdot \\
& \cdot \int_0^1 G(h)^{m-n} \int_{\mathbb{R}^{n-1}} \int_0^{\sqrt{1-h^2}} |\Theta - c_n^{n-1}| \int_{\mathbb{R}^{n-2}} \cdots \int_{\mathbb{R}^{n-2}} |\lambda_{n-2}\{\text{conv}(c_1, \ldots, c_{n-1})\}|^2 \cdot \\
& \cdot W(c_1, \ldots, c_{n-1})\, f(c_1) \cdots f(c_{n-1})\, d\bar{\bar{c}}_1 \ldots d\bar{\bar{c}}_{n-1}\, d\Theta\, f(c_n)\, d\bar{c}_n\, dh.
\end{aligned}
\tag{34}
$$

That integral describes the expected number of those facets of the polytope $\text{conv}(0, a_1, \ldots, a_m)$ which do not contain the origin and are intersected by $\text{lin}(u, v)$. But this integral does – up to now – not admit a direct evaluation of a satisfactory precision. Only for the asymptotic configuration of dimensions ($m \to \infty$, n fixed), such a direct evaluation led to reliable results as in Borgwardt (1982a).

The crucial idea to obtain a good estimation in the case of moderate dimensions was the following: compare the figure $E_{m,n}(S)$ with a closely related figure $E_{m,n}(Z)$, which gives the expected number of facets being intersected by the ray $\mathbb{R}^+ v$. The figure Z is at most 1 in all nondegenerate cases, because then a random ray will leave Y already at the origin or in the interior of one facet and those facets do not overlap. To derive an integral-formula for $E_{m,n}(Z)$, we introduce $V(a_1, \ldots, a_n)$ as the spherical angle of $\text{cone}(a_1, \ldots, a_n)$.

$$
V(a_1, \ldots, a_n) := \frac{\lambda_n \{\Omega_n \cap \text{cone}(a_1, \ldots, a_n)\}}{\lambda_n\{\Omega_n\}} = \frac{\lambda_{n-1} \{\omega_n \cap \text{cone}(a_1, \ldots, a_n)\}}{\lambda_{n-1}\{\omega_n\}}.
\tag{35}
$$

And, rotation symmetry (cf. Borgwardt (1987),p.143) provides that for a fixed $(n-1)$-dimensional simplex $\text{conv}(a_1, \ldots, a_n)$ we have

$$
P(\mathbb{R}^+ v \ \text{intersects} \ \text{conv}(a_1, \ldots, a_n)) \;=\; V(a_1, \ldots, a_n).
\tag{36}
$$

Corresponding to (33), the integral formula for Z is:

$$
E_{m,n}(Z) = \binom{m}{n} \int_{\mathbb{R}^n} \ldots \int_{\mathbb{R}^n} G(h[a_1, \ldots, a_n])^{m-n} V(a_1, \ldots, a_n) f(a_1) \cdots f(a_n)\, da_1 \ldots da_n.
\tag{37}
$$

12

And after the same coordinate-transformations as above, we have

$$E_{m,n}(Z) = \binom{m}{n} \cdot \{(n-2)!\}^2 \cdot \lambda_{n-1}\{\omega_n\} \cdot \lambda_{n-2}\{\omega_{n-1}\} \cdot$$

$$\cdot \int_0^1 G(h)^{m-n} \int_{\mathbb{R}^{n-1}} \int_0^{\sqrt{1-h^2}} |\Theta - c_n^{n-1}| \int_{\mathbb{R}^{n-2}} \cdots \int_{\mathbb{R}^{n-2}} |\lambda_{n-2}\{\mathrm{conv}(c_1,\ldots,c_{n-1})\}|^2 \cdot$$

$$\cdot V(c_1,\ldots,c_n) f(c_1) \cdots f(c_{n-1}) \, d\overline{\overline{c}}_1 \ldots d\overline{\overline{c}}_{n-1} \, d\Theta f(c_n) \, d\overline{c}_n \, dh. \tag{38}$$

This comparison turns out to be advantageous, since we know that

$$E_{m,n}(Z) \leq 1, \tag{39}$$

and therefore

$$E_{m,n}(S) \leq \frac{E_{m,n}(S)}{E_{m,n}(Z)}. \tag{40}$$

That means that as soon as we have a good upper bound for the quotient, then we are done.

The similarity of the integrals in numerator and denominator simplifies the evaluation of the quotient. Its estimation becomes much easier than direct calculation of (34).

A main trick in our derivation of an upper-bound for $\frac{E_{m,n}(S)}{E_{m,n}(Z)}$ (as desired in (40)) is the "Principle of Pointwise Comparison". For application of that principle, we dissect the set of all configurations in the common integration area of (34) and (38) in a number of cells, and compare the contributions of each such cell to the numerator- and denominator-integral in (40). Then we know that the quotient of (34) and (38) cannot exceed the maximal quotient observed at single cells. The dissection is chosen in such a way that the comparison becomes easy and that the maximum mentioned above stays small.

Precisely, we form cells by collecting the set of all configurations with identical values of $t := \sqrt{h^2 + \Theta^2}$. We can calculate such a t-contribution by integrating over the contributions of all pairs (Θ, h) with $\Theta = \sqrt{t^2 - h^2}$. For studying the contribution of one such pair we have to integrate over all potential configurations of the points $c_1, \ldots, c_{n-1}, c_n$, where $c_1^n = \ldots c_n^n = h$ and $c_1^n = \ldots c_{n-1}^{n-1} = \Theta$.

Note that the change from one such set with characteristic pair (Θ, h) to another with (Θ', h') and $t := \sqrt{h'^2 + \Theta'^2}$ simply results from a rotation in the space of the $(n-1)$-th and the n-th coordinate, and that such a rotation leaves the other coordinates unchanged. Here the assumption of rotation symmetry guarantees that every configuration of points for (Θ, h) has its analogous counterpart in the set with (Θ', h') and that many figures in the integrals
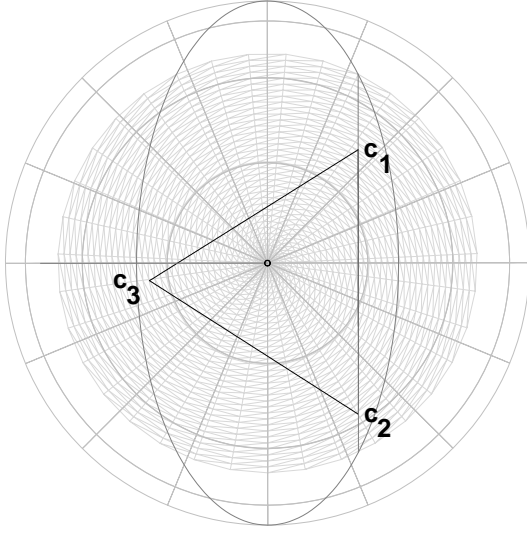
Figure 1: The configuration for $n = 3$ after the two coordinate-transformations. In the left figure we look from above (the position of e_3) down on the gray disk illustrating the region $\{x \in \Omega_3 \mid x^3 = h\}$. c_1 and c_2 have got an identical second coordinate, i.e. $c_1{}^2 = c_2{}^2 = T$. The point in the center of that region is $h \cdot e_3$. The line bearing c_1 and c_2 has distance T to $h \cdot e_3$ and distance t to the origin. The ellipse illustrates the intersection of ω_3 with the (hyper-)plane H, which contains the origin, c_1 and c_2. In the right figure we look from the side on the plane of the coordinates x^2, x^3. We see in height h the disk mentioned above, bearing c_1, c_2, c_3. The symbol c_{12} is used, because from our perspective c_1 is covered by c_2. In addition, we see a triangle explaining the role of the values t, h, T. The diagonal line shows $H \cap \omega_3$. We recognize its "highest" and its "lowest point" with respect to x^3. Finally, we see z, the normal vector to H, positively oriented towards e_3.

remain constant under such a rotation. This simplifies the integration over the dissection-subset, since constant factors will not cause complications. And it makes a significant reduction of the quotient possible, since these constant factors appear simultaneously in the numerator- and the denominator-integral. Some of these advantages are

- the internal distribution of the "random variables" $(\bar{\bar{c}}, \Theta, h)^T$ in $\{c \mid c^n = h, \ c^{n-1} = \Theta\}$ is identical for all pairs (h, Θ) with constant value of $t := \sqrt{h^2 + \Theta^2}$,

- the value of the integrals $\underset{\mathbb{R}^{n-2}}{\int} \cdots \underset{\mathbb{R}^{n-2}}{\int} f(c_1) \cdots f(c_{n-1}) \, d\bar{\bar{c}}_1 \ldots d\bar{\bar{c}}_{n-1}$ is independent of h as long as t is fixed,

- the spherical measures $W(c_1, \ldots, c_{n-1})$ do not vary, when $\bar{\bar{c}}_1, \ldots, \bar{\bar{c}}_{n-1}$ and t remain fixed, although h and Θ may vary simultaneously.

We would loose these advantages, if we tried to integrate over configurations with varying values of t.

With the substitution $t := \sqrt{h^2 + \Theta^2}$, $T := \sqrt{t^2 - h^2}$ we obtain

$$
\frac{E_{m,n}(S)}{E_{m,n}(Z)} = \frac{\binom{m}{n} \cdot n \cdot \{(n-2)!\}^2 \cdot \lambda_{n-1}(\omega_n) \cdot \lambda_{n-2}(\omega_{n-1}) \cdot}{\binom{m}{n} \cdot \{(n-2)!\}^2 \cdot \lambda_{n-1}(\omega_n) \cdot \lambda_{n-2}(\omega_{n-1}) \cdot}
$$

$$
\frac{\cdot \int_0^1 t \int_0^t G(h)^{m-n} T^{-1} \int_{\mathbb{R}^{n-1}} |T - c_n^{n-1}| \int_{\mathbb{R}^{n-2}} \cdots \int_{\mathbb{R}^{n-2}} |\lambda_{n-2}\{\mathrm{conv}(c_1, \ldots, c_{n-1})\}|^2 \cdot}{\cdot \int_0^1 t \int_0^t G(h)^{m-n} T^{-1} \int_{\mathbb{R}^{n-1}} |T - c_n^{n-1}| \int_{\mathbb{R}^{n-2}} \cdots \int_{\mathbb{R}^{n-2}} |\lambda_{n-2}\{\mathrm{conv}(c_1, \ldots, c_{n-1})\}|^2 \cdot}
$$

$$
\frac{\cdot W(c_1, \ldots, c_{n-1}) \, f(c_1) \cdots f(c_{n-1}) \, d\bar{\bar{c}}_1 \ldots d\bar{\bar{c}}_{n-1} \, f(c_n) \, d\bar{c}_n \, dh \, dt}{\cdot V(c_1, \ldots, c_n) \, f(c_1) \cdots f(c_{n-1}) \, d\bar{\bar{c}}_1 \ldots d\bar{\bar{c}}_{n-1} \, f(c_n) \, d\bar{c}_n \, dh \, dt}. \tag{41}
$$

Here we can exploit the "Principle of Pointwise Comparison," by writing

$$
\frac{E_{m,n}(S)}{E_{m,n}(Z)} \le \sup_{t \in (0,1]} PQ(t) \quad \text{and defining } PQ(t) \text{ as the pointwise quotient for } t. \tag{42}
$$

$$
PQ(t) := \frac{n \cdot \int_0^t G(h)^{m-n} T^{-1} \int_{\mathbb{R}^{n-1}} |T - c_n^{n-1}| \int_{\mathbb{R}^{n-2}} \cdots \int_{\mathbb{R}^{n-2}} |\lambda_{n-2}\{\mathrm{conv}(c_1, \ldots, c_{n-1})\}|^2 \cdot}{\cdot \int_0^t G(h)^{m-n} T^{-1} \int_{\mathbb{R}^{n-1}} |T - c_n^{n-1}| \int_{\mathbb{R}^{n-2}} \cdots \int_{\mathbb{R}^{n-2}} |\lambda_{n-2}\{\mathrm{conv}(c_1, \ldots, c_{n-1})\}|^2 \cdot}
$$

$$
\frac{\cdot W(c_1, \ldots, c_{n-1}) \, f(c_1) \cdots f(c_{n-1}) \, d\bar{\bar{c}}_1 \ldots d\bar{\bar{c}}_{n-1} \, f(c_n) \, d\bar{c}_n \, dh}{\cdot V(c_1, \ldots, c_n) \, f(c_1) \cdots f(c_{n-1}) \, d\bar{\bar{c}}_1 \ldots d\bar{\bar{c}}_{n-1} \, f(c_n) \, d\bar{c}_n \, dh} = \tag{43}
$$

$$
= \frac{n \cdot \int_0^t G(h)^{m-n} T^{-1} \int_{\mathbb{R}^{n-1}} |T - c_n^{n-1}| \int_{\mathbb{R}^{n-2}} \cdots \int_{\mathbb{R}^{n-2}} |\lambda_{n-2}\{\mathrm{conv}(c_1, \ldots, c_{n-1})\}|^2 \cdot}{\int_0^t G(h)^{m-n} T^{-1} \int_{\mathbb{R}^{n-1}} |T - c_n^{n-1}| \int_{\mathbb{R}^{n-2}} \cdots \int_{\mathbb{R}^{n-2}} |\lambda_{n-2}\{\mathrm{conv}(c_1, \ldots, c_{n-1})\}|^2 \cdot}
$$

$$
\frac{\cdot W(c_1, \ldots, c_{n-1}) \, f(c_1) \cdots f(c_{n-1}) \, d\bar{\bar{c}}_1 \ldots d\bar{\bar{c}}_{n-1} \, f(c_n) \, d\bar{c}_n \, dh}{\cdot \frac{V(c_1, \ldots, c_n)}{W(c_1, \ldots, c_{n-1})} \cdot W(c_1, \ldots, c_{n-1}) \, f(c_1) \cdots f(c_{n-1}) \, d\bar{\bar{c}}_1 \ldots d\bar{\bar{c}}_{n-1} \, f(c_n) \, d\bar{c}_n \, dh} \tag{44}
$$

From now on let us fix the value of t and try to find a bound on $PQ(t)$, independent of t.

3 Comparison of spherical angles

Formula (44) indicates that the main technical task is to clarify the relation between V and W. Note that both are spherical measures of cones, but of different dimensions. Whereas V measures a full dimensional cone, spanned by $(c_1, \ldots, c_n)$, W deals only with one side of this cone, namely cone$(c_1, \ldots, c_{n-1})$. Its spherical measure W is associated with ω_{n-1}, as it measures a part of a unit sphere in a hyperplane of dimension $n-1$. Thus, the position of c_n gains an extraordinary significance. Note that (44) describes the inverse of an expectation value of the random variable $\frac{V}{W}$ (under a rather strange distribution).

In this section, we explain two different strategies, which can be used for that comparison. The first was employed in Borgwardt (1987) and the second is the basis for our new result.

In the old approach we derived a relatively simple estimation of $\frac{V(c_1, \ldots, c_n)}{W(c_1, \ldots, c_{n-1})}$ and inserted this into (44). Unfortunately, this method caused a considerable underestimation.

Before explaining the new approach, let us generalize the use of conv and cone such that also complete sets (and not the elements of these sets) are treated as generators for forming the convex (conical) hull. This is done because we will have to deal with nonconvex sets, whose single elements are combined with an augmenting point by forming the convex hull of both points. But this operation does not form the convex hull of the mentioned nonconvex set. If M is a set, then $\widetilde{\mathrm{conv}}(x, M)$ $(\widetilde{\mathrm{cone}}(x, M))$ means the set of all convex (conical) combinations of x with single points $y \in M$. Formally:

$$\widetilde{\mathrm{conv}}(x, M) = \bigcup_{\gamma \in [0,1]} \gamma x + (1 - \gamma)M \qquad \text{and} \qquad \widetilde{\mathrm{cone}}(x, M) = \bigcup_{\rho_1, \rho_2 \in [0, \infty)} \rho_1 x + \rho_2 M.$$

Note that this may differ from conv(x, M), the convex hull of $\{x\} \cup M$. And also cone(x, M) may differ from $\widetilde{\mathrm{cone}}(x, M)$. If we apply these operators to one generator-set only, then we recognize that

$$\widetilde{\mathrm{conv}}(M) = M \qquad \text{and} \qquad \widetilde{\mathrm{cone}}(M) = \bigcup_{\rho \in [0, \infty)} \rho M.$$

The Cavalieri-Principle states the following:
The volume of an n-dimensional area (or body) enclosed between two parallel hyperplanes H_0 and H_1 can be calculated by integrating over all $(n-1)$-dimensional volumes resulting from intersecting the area (body) and hyperplanes H, which are parallel to H_0 and H_1.

This principle enables us to determine the volume of the body, if the size of all intersection areas is known and the resulting integral can be solved easily.
Both advantages are available, if our area (body) has the general structure $\widetilde{\mathrm{conv}}(c, M)$, where c is an arbitrary point in $\mathbb{R}^n$ and M is a measurable $(n-1)$-dimensional subset of a hyperplane H_0. $\widetilde{\mathrm{conv}}(c, M)$ is enclosed between H_0 and H_1, where the latter is the hyperplane, which is parallel to H_0 and contains c. Then all intersection areas with parallel hyperplanes are scaled copies of M. If we define τ as the quotient $\tau := \frac{\mathrm{dist}(c,H)}{\mathrm{dist}(c,H_0)}$ with H for the intersecting hyperplane, then the correponding intersection-volume is $\lambda_{n-1}\{M\} \cdot \tau^{n-1}$. Integration over all parallel hyperplanes yields the well-known formula for the volume of $\widetilde{\mathrm{conv}}(c, M)$

$$Volume = \frac{1}{n} \times \text{(height of } c) \times \text{ (size of ground area } M) = \frac{1}{n} \cdot \text{dist}(c, H) \cdot \lambda_{n-1}\{M\}. \quad (45)$$

This result can be applied to the following special case. Replace $c \notin H$ by c_n, H by $H(0, c_1, \ldots, c_{n-1})$, M by $\text{cone}(c_1, \ldots, c_{n-1}) \cap \Omega_n$. Then our formula yields

$$\lambda_n \left\{ \widetilde{\text{conv}} \left(\frac{c_n}{\|c_n\|}, \text{cone}(c_1, \ldots, c_{n-1}) \cap \Omega_n \right) \right\} =$$
$$= \frac{1}{n} \cdot \lambda_{n-1} \left\{ \text{cone}(c_1, \ldots, c_{n-1}) \cap \Omega_n \right\} \quad \cdot \quad \text{dist} \left(\frac{c_n}{\|c_n\|}, H(0, c_1, \ldots, c_{n-1}) \right). \quad (46)$$

On the other hand we know that

$$\lambda_n \left\{ \widetilde{\text{conv}} \left(\frac{c_n}{\|c_n\|}, \text{cone}(c_1, \ldots, c_{n-1}) \cap \Omega_n \right) \right\} \leq \lambda_n \left\{ \text{cone}(c_1, \ldots, c_n) \cap \Omega_n \right\}, \quad (47)$$

because the set on the right side is convex and contains $\frac{c_n}{\|c_n\|}$ and $\text{cone}(c_1, \ldots, c_{n-1}) \cap \Omega_n$. Recalling the definitions of $W(c_1, \ldots, c_{n-1})$ in (31) and of $V(c_1, \ldots, c_n)$ in (35) yields

$$\frac{V(c_1, \ldots, c_n)}{W(c_1, \ldots, c_{n-1})} \geq \frac{\lambda_n \left\{ \widetilde{\text{conv}} \left(\frac{c_n}{\|c_n\|}, \text{cone}(c_1, \ldots, c_{n-1}) \cap \Omega_n \right) \right\}}{\lambda_{n-1} \left\{ \text{cone}(c_1, \ldots, c_{n-1}) \cap \Omega_n \right\}} \cdot \frac{\lambda_{n-1}\{\Omega_{n-1}\}}{\lambda_n\{\Omega_n\}} =$$
$$= \frac{1}{n} \cdot \text{dist} \left(\frac{c_n}{\|c_n\|}, H(0, c_1, \ldots, c_{n-1}) \right) \cdot \frac{\lambda_{n-1}\{\Omega_{n-1}\}}{\lambda_n\{\Omega_n\}}. \quad (48)$$

The insertion of this formula into (44) was the basis for the successful overestimation in Borgwardt (1987).

Now it is time to present the new evaluation method. We avoid proceeding to (44) and consider formula (43). It will be our aim to reformulate $W(c_1, \ldots, c_{n-1})$ and $V(c_1, \ldots, c_n)$ such that their close relationship becomes evident and that the precision of our estimation remains high. Before going into details, we introduce some notation.

Definition 1
  1. Let $\angle(x, y)$ denote the angle between two vectors x and y. Correspondingly, we write $\angle(c_n, \text{lin}(c_1, \ldots, c_{n-1}))$ for the minimal angle between c_n and $\text{lin}(c_1, \ldots, c_{n-1})$, resp. the hyperplane $H(0, c_1, \ldots, c_{n-1})$.

  2. Let $z = z(c_1, \ldots, c_{n-1})$ be the normal vector on $\text{lin}(c_1, \ldots, c_{n-1})$. In order to make this correspondence unique, let that z be positively oriented towards e_n. Since after our two coordinate-transformations as in (43), z does only depend on the variables T and h, it is also useful and admitted to write $z = z[T, h]$.

As we have already fixed the value of t, we may now – for the moment – also fix the location of the hyperplane H containing $0, c_1, \ldots, c_{n-1}$ and of the augmenting point $\frac{c_n}{\|c_n\|}$ (in a nondegenerate configuration, since degenerate cases do not contribute to (43)). As long as we rely on that fixation, we shall abbreviate $H := H(0, c_1, \ldots, c_{n-1})$.
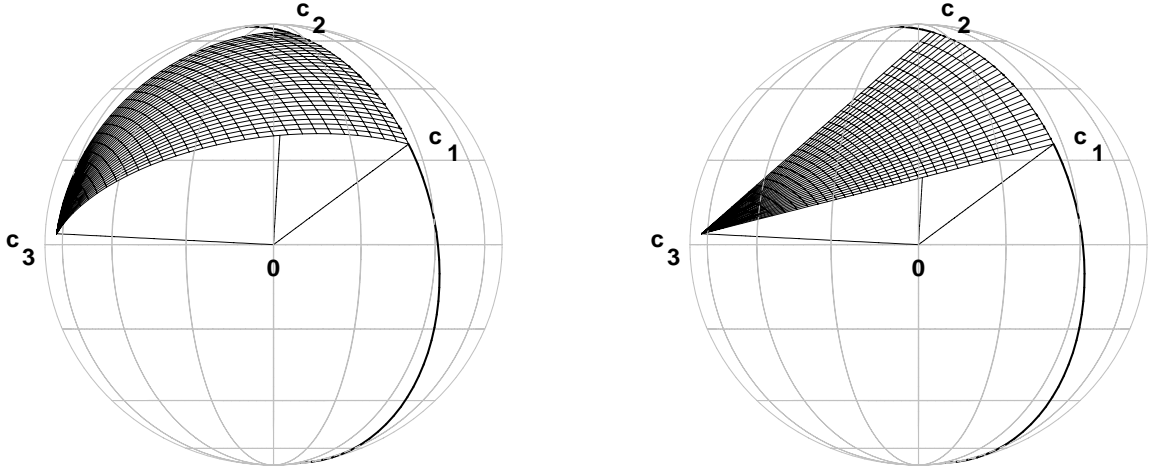
Figure 2: A spherical triangle generated by three (normalized) points c_1, c_2, c_3. Our task is to compare the size of the conical sector spanned by these points with that of the two-dimensional sector spanned by c_1, c_2 only. In the right figure we see a subset of the 3-dimensional sector, whose volume can easily be calculated using the Cavalieri-Principle. Here the surface is given by connecting c_3 with the points of the bow from c_1 to c_2 directly by straight lines. Note that this surface traverses the interior of the unit ball and that the size of the subset on the right underestimates the size of the full sector significantly. This right surface can in our generalized terminology be written as $\widetilde{\text{conv}}(c_3, \text{cone}(c_1, c_2) \cap \omega_n)$, and the calculated sector is $\widetilde{\text{conv}}(c_3, \text{cone}(c_1, c_2) \cap \Omega_n)$, whereas in the left figure we see $\text{cone}(c_1, c_2, c_3) \cap \omega_n$ and $\text{cone}(c_1, c_2, c_3) \cap \Omega_n$.

We are going to show that it is possible to formulate both spherical angles $W(c_1, \ldots, c_{n-1})$ and $V(c_1, \ldots, c_n)$ as integrals over an appropriate subset of $\omega_n \cap H$.

It is a tautology to see that

$$W(c_1, \ldots, c_{n-1})\lambda_{n-2}\{\omega_{n-1}\} = \lambda_{n-2}\left\{\text{cone}(c_1, \ldots, c_{n-1}) \cap \omega_n\right\} = \int\limits_{\text{cone}(c_1, \ldots, c_{n-1}) \cap \omega_n} \lambda_{n-2}\{d_{n-1}w\}, \quad (49)$$

and that correspondingly

$$V(c_1, \ldots, c_n)\lambda_{n-1}\{\omega_n\} = \lambda_{n-2}\left\{\text{cone}(c_1, \ldots, c_n) \cap \omega_n\right\} = \int\limits_{\text{cone}(c_1, \ldots, c_n) \cap \omega_n} \lambda_{n-1}\{d_n\breve{w}\}, \quad (50)$$

when we use w and $\breve{w}$ for the respective integration variables. $d_n\breve{w}$ stands for the surface-element over ω_n, and $d_{n-1}w$ works accordingly for ω_{n-1} or $\omega_n \cap H$.

It is very important and useful that even $V(c_1, \ldots, c_n) \cdot \lambda_{n-1}\{\omega_n\}$ can be written as an integral over $\text{cone}(c_1, \ldots, c_{n-1}) \cap \omega_n$, similarly to (49). This is proven in

Lemma 1

For fixed $z = z(c_1, \ldots, c_{n-1}) = z[T, h]$ and every nondegenerate set $\{c_1, \ldots, c_{n-1}\}$ inducing exactly that z, and as long as $c_n \notin H$, there exists a function $\Upsilon(\frac{c_n}{\|c_n\|}, z; w)$, defined for $w \in H \cap \omega_n$, such that

$$V(c_1, \ldots, c_n)\lambda_{n-1}\{\omega_n\} = \lambda_{n-1}\left\{\text{cone}(c_1, \ldots, c_n) \cap \omega_n\right\} = \int\limits_{\text{cone}(c_1, \ldots, c_{n-1}) \cap \omega_n} \Upsilon(\frac{c_n}{\|c_n\|}, z; w)\lambda_{n-2}\{d_{n-1}w\}.$$
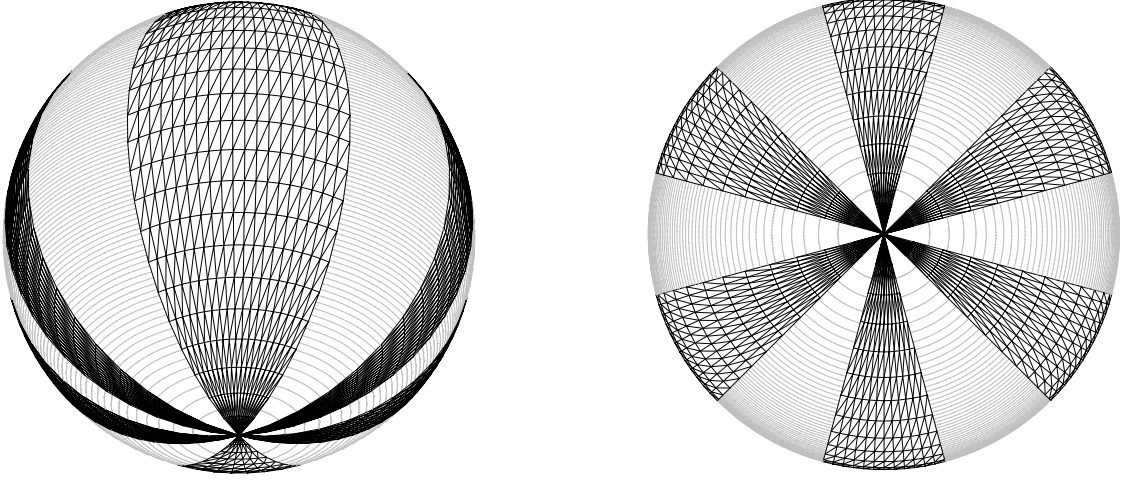
18

Figure 3: Imagine that we look at the hemisphere $\omega_3 \cap \{x|z^Tx \geq 0\}$ from the position of z (the "northpole"). Assume that the bounding circle $\omega_3 \cap \{x|z^Tx = 0\} = \omega_3 \cap H$ (the "equator") is dissected in 12 sections $M_1, \ldots, M_{12}$ of equal size. In the left figure we have put c_3 on a general place in the hemisphere and in the right figure exactly at z. Now we can observe the impact of the "projection along the sphere" $\breve{\Pi}_{c_3}$. The sectors given in different patterns on the hemisphere illustrate the respective preimages of the 12 circle-sections under that projection. In the right figure we obtain a uniform distribution, where the preimages all have equal sizes. But in the left figure the corresponding attribution is far from being uniform. Increasing the number of circle-sections and decreasing their size leads (in the limit) to an illustration for the density $\Upsilon(e_3, z; w)$.

Proof
Without loss of generality, let $c_n^T z > 0$. (For $c_n^T z < 0$, the proof would work analogously).
Note that every set of type $\text{cone}(c_1, \ldots, c_n) \cap \omega_n$ has the general structure $\widetilde{\text{cone}}(c, M) \cap \omega_n$, where c is a point on ω_n and M is a measurable set in $H \cap \omega_n$.
We set $H^+ := \{x|z^Tx \geq 0\}$ and define a finite measure $\lambda_{n-1}\{\breve{M}\}$ for measurable sets $\breve{M} \subset \omega_n \cap H^+$ (the hemisphere above $H(0, c_1, \ldots, c_{n-1})$).
Let us denote this measure by $\breve{\kappa}$.
On the other hand we have the finite Lebesgue-measure $\lambda_{n-2}\{M\}$ on $H \cap \omega_n$ for measurable sets M on that space. To avoid confusions, we call that measure ν and have

$$\nu\{M\} = \int\limits_M \lambda_{n-2}\{d_{n-1}w\}.$$

Now we introduce $\breve{\Pi}_c$, a "projection along the sphere", which maps the hemisphere $\omega_n \cap H^+$ onto the subdimensional sphere $H \cap \omega_n$. $\breve{\Pi}_c$ maps $\breve{x} \in \omega_n \cap \{x|z^Tx \geq 0\}$ as follows:

$$\breve{\Pi}_c(\breve{x}) = H \cap \text{cone}(c, \breve{x}, -c) \cap \omega_n = \frac{\breve{x} - \gamma \cdot z}{\|\breve{x} - \gamma \cdot z\|}, \text{ where } \gamma = \frac{z^T\breve{x}}{z^Tz} = z^T\breve{x}.$$

It is clear that $z^T(\frac{\breve{x}-\gamma\cdot z}{\|\breve{x}-\gamma\cdot z\|}) = 0$, hence $\breve{\Pi}_c(\breve{x}) \in H \cap \omega_n$, and that $\breve{\Pi}_c$ is a projection. This mapping is measurable and it induces a new measure κ on $H \cap \omega_n$ via the measure-transfer

19

rule $\kappa\{M\} := \breve{\kappa}\{\breve{\Pi}_c^{-1}(M)\}$ for every measurable set $M \subset H \cap \omega_n$, because $\breve{\Pi}_c^{-1}(M)$ is a measurable set on $\omega_n \cap H^+$. Note that as long as $c \notin H$, we have

$$\breve{\Pi}_c^{-1}(M) = \widetilde{\mathrm{cone}}(c, M) \cap \omega_n,$$

and that κ is proportional to ν only in the <u>very special case</u>, where $c = z$. Then

$$\kappa = \nu \cdot \frac{\lambda_{n-1}\{\omega_n\}}{2 \cdot \lambda_{n-2}(\omega_{n-1})}.$$

Moreover, the new (artificial) measure κ is ν-continuous, because for every measurable set with $\lambda_{n-2}\{M\} = \nu\{M\} = 0$ it is clear that as long as $c \notin H$

$$\kappa\{M\} = \breve{\kappa}\{\breve{\Pi}_c^{-1}(M)\} = \breve{\kappa}\{\widetilde{\mathrm{cone}}(c, M) \cap \omega_n\} = \lambda_{n-1}\{\widetilde{\mathrm{cone}}(c, M) \cap \omega_n\} = 0.$$

For this configuration, the Theorem of Radon-Nykodym confirms that κ posesses a density relative to ν. And this means for our situation, where $c = \frac{c_n}{\|c_n\|}$ and $z = z(c_1, \ldots, c_{n-1})$ are fixed, that there is a function of $w \in H \cap \omega_n$, denoted by $\Upsilon(\frac{c_n}{\|c_n\|}, z[T, h]; w)$, representing this ν-density of κ, which satisfies for every measurable set $M \subset H \cap \omega_n$:

$$
\begin{aligned}
\lambda_{n-1}\left\{\widetilde{\mathrm{cone}}\left(\frac{c_n}{\|c_n\|}, M\right) \cap \omega_n\right\} &= \breve{\kappa}\left\{\widetilde{\mathrm{cone}}\left(\frac{c_n}{\|c_n\|}, M\right) \cap \omega_n\right\} = \\
= \breve{\kappa}\{\breve{\Pi}_c^{-1}(M)\} = \kappa\{M\} &= \int_M \Upsilon(\frac{c_n}{\|c_n\|}, z[T, h]; w)\nu\{d_{n-1}w\} = \\
&= \int_M \Upsilon(\frac{c_n}{\|c_n\|}, z[T, h]; w)\lambda_{n-2}\{d_{n-1}w\}. \quad (51)
\end{aligned}
$$

(We write the semicolon, because the parameters left of it indicate that for different values of $\frac{c_n}{\|c_n\|}$ and $z[T, h]$ the density-functions will be different, which is meaningless for the present configuration, where c_n and z are fixed.)

But now we can replace M by $\mathrm{cone}(c_1, \ldots, c_{n-1}) \cap \omega_n$ and

$$\widetilde{\mathrm{cone}}\left(\frac{c_n}{\|c_n\|}, M\right) \quad \text{by} \quad \widetilde{\mathrm{cone}}\left(\frac{c_n}{\|c_n\|}, \mathrm{cone}(c_1, \ldots, c_{n-1})\right) = \mathrm{cone}(c_1, \ldots, c_n).$$

So we learn that

$$
\begin{aligned}
V(c_1, \ldots, c_n) \cdot \lambda_{n-1}\{\omega_n\} &= \lambda_{n-1}\{\mathrm{cone}(c_1, \ldots, c_n) \cap \omega_n\} = \quad (52) \\
&= \int_{\mathrm{cone}(c_1, \ldots, c_{n-1}) \cap \omega_n} \Upsilon(\frac{c_n}{\|c_n\|}, z[T, h]; w)\lambda_{n-2}\{d_{n-1}w\}. \quad \square
\end{aligned}
$$

Now measure theory gives the following additional information:

Remark 1

Any arbitrary density simulating the impact of κ can differ from $\Upsilon(\frac{c_n}{\|c_n\|}, z[T, h]; w)$ only on a nullset with respect to $\nu = \lambda_{n-2}$. $\qquad\qquad\square$

These integral-formulas for W and V give us an alternative to the Cavalieri-underestimation. In the next section, we try to calculate and to quantify $\Upsilon(\frac{c_n}{\|c_n\|}, z[T, h]; w)$.

4 A refined analysis of the spherical angle

In order to install a more precise estimation for $V(c_1, \ldots, c_n)$, we must know $\Upsilon(\frac{c_n}{\|c_n\|}, z[T, h]; w)$, as introduced in Lemma 1, resp. in (52).
We have to generalize the use of V and W – as we did it with $\widetilde{\text{conv}}, \widetilde{\text{cone}}$ – and study

$$\frac{\widetilde{V}\left(\frac{c_n}{\|c_n\|}, M(w)\right)}{\widetilde{W}(M(w))} = \frac{\lambda_{n-1}\left\{\widetilde{\text{cone}}\left(\frac{c_n}{\|c_n\|}, M(w)\right) \cap \omega_n\right\}}{\lambda_{n-2}\{M(w) \cap \omega_n\}} \cdot \frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}}$$

for points w in $\omega_n \cap H$, and for very small areas $M(w)$ about w in $\omega_n \cap H$. We are interested in the spherical angle (with respect to ω_n) induced by $M(w)$ in cooperation with an arbitrary point $\frac{c_n}{\|c_n\|} \in \omega_n$, i.e.

$$\widetilde{V}\left(\frac{c_n}{\|c_n\|}, M(w)\right) = \frac{\lambda_{n-1}\left\{\widetilde{\text{cone}}\left(\frac{c_n}{\|c_n\|}, M(w)\right) \cap \omega_n\right\}}{\lambda_{n-1}\{\omega_n\}}.$$

As before, z shall denote the normal vector to $H(0, c_1, \ldots, c_{n-1})$ (normalized and positively oriented towards e_n). Throughout this section $c_1, \ldots, c_{n-1}$ and $H = H(0, c_1, \ldots, c_{n-1})$ will be kept fixed. e_n is the vector $(0, \ldots, 0, 1)^T$, which will be regarded as the "north pole of ω_n." W. l. o. g. we may assume that $\frac{1}{\|c_n\|}c_n$ and e_n belong to the same halfspace induced by H. The other case, where c_n and e_n belong to opposite halfspaces, can be treated symmetrically. It is even allowed to identify $\frac{1}{\|c_n\|}c_n$ with e_n, and to assume that $z^T e_n > 0$ in order to learn how to calculate the spherical angles. This configuration is much simpler to imagine, to illustrate and to describe than the general case. We are going to make use of that simplification for local purposes, as in some proofs (where indicated). Claims are stated for the general configuration.

This section has the task to give the confirmation of the following auxiliary result.

Proposition 1
If $\frac{c_n}{\|c_n\|}$ and $z[T, h]$ are fixed, and if $\frac{c_n}{\|c_n\|} \notin H$, then the density function is of the form

$$\Upsilon(\frac{c_n}{\|c_n\|}, z[T, h]; w) := \frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \cdot \text{dist}\left(\frac{c_n}{\|c_n\|}, H\right) \cdot$$

$$\cdot \frac{1}{(1 - (\frac{c_n}{\|c_n\|}^T w)^2)^{\frac{n-1}{2}}} \int_{\frac{c_n}{\|c_n\|}^T w}^{1} \sqrt{1 - h^2}^{n-3} dh, \qquad (53)$$

and the spherical angle generated by $\text{cone}(c_1, \ldots, c_n)$ admits the formulation

$$V(c_1, \ldots, c_n) = \frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \cdot \text{dist}\left(\frac{c_n}{\|c_n\|}, H\right) \cdot \qquad (54)$$

$$\cdot \int_{\text{cone}(c_1,\ldots,c_{n-1})\cap\omega_n} \frac{1}{(1 - (\frac{c_n}{\|c_n\|}^T w)^2)^{\frac{n-1}{2}}} \int_{\frac{c_n}{\|c_n\|}^T w}^{1} \sqrt{1 - h^2}^{n-3} dh \ \lambda_{n-2}\{d_{n-1}w\}. \qquad \square$$

The following measure-method for spherical angles of a simple structure is elementary and well-known from standard calculus. We cite it without proof and use it as a basic tool for the calculation of more complicated angles.

Remark 2

Consider a measurable set $K \subset \{x \mid \|x\| = 1, \ x^n = 0\}$ and $\widetilde{\mathrm{cone}}(K)$, and the spherical measure of the (general) cone generated by the set K and the point e_n. Then the spherical measure of the spherical set $\widetilde{\mathrm{cone}}(e_n, K, -e_n) \cap \omega_n \cap \{x | x^n \geq \overline{h}\}$, where $\overline{h} \in (-1, +1)$, is

$$\frac{\lambda_{n-1}\left\{\widetilde{\mathrm{cone}}(e_n, K, -e_n) \cap \omega_n \cap \{x | x^n \geq \overline{h}\}\right\}}{\lambda_{n-1}\{\omega_n\}} =$$

$$= \frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \int\limits_{\overline{h}}^{1} \sqrt{1 - h^2}^{n-3} dh \frac{\lambda_{n-2}\{K \cap \omega_{n-1}\}}{\lambda_{n-2}\{\omega_{n-1}\}} =$$

$$= \frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \int\limits_{\overline{h}}^{1} \sqrt{1 - h^2}^{n-3} dh \cdot \frac{\lambda_{n-1}\{\widetilde{\mathrm{cone}}(K) \cap \Omega_{n-1}\}}{\lambda_{n-1}\{\Omega_{n-1}\}}. \tag{55}$$

Based on (55), we can determine a function $\hat{\Upsilon}(\frac{c_n}{\|c_n\|}, z[T, h], w)$ with the following property.

Lemma 2

Let z and c_n with $z^T c_n > 0$ be fixed. And let w be a point of $\omega_n \cap H$. Consider a sequence of positive δ-values such that $\delta \to 0_+$. Let $M^\delta(w)$ be a corresponding sequence of sets in $\omega_n \cap H$ with $M^\delta(w) \subset \{x | \angle(x, w) < \delta\}$, and $\widetilde{W}(M^\delta(w)) > 0$. Then the approximation-error in the formula

$$\hat{\Upsilon}(\frac{c_n}{\|c_n\|}, z[T, h], w) = \frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \int\limits_{\frac{c_n}{\|c_n\|}^T w}^{1} \sqrt{1 - h^2}^{n-3} dh \cdot \frac{|\cos\left(\angle(\frac{c_n}{\|c_n\|}, z)\right)|}{(1 - (\frac{c_n}{\|c_n\|}^T w)^2)^{\frac{n-1}{2}}} \tag{56}$$

$$\approx \frac{\widetilde{V}\left(\frac{c_n}{\|c_n\|}, M^\delta(w)\right)}{\widetilde{W}(M^\delta(w))} = \frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \cdot \frac{\lambda_{n-1}\left\{\widetilde{\mathrm{cone}}\left(\frac{c_n}{\|c_n\|}, M^\delta(w)\right) \cap \omega_n\right\}}{\lambda_{n-2}\left\{\widetilde{\mathrm{cone}}(M^\delta(w)) \cap \omega_n\right\}} \tag{57}$$

converges uniformly to 0 for $\delta \to 0_+$.

That means that for every $\epsilon > 0$ we can find a value $\delta(\epsilon, z)$, independent of w, such that this error is smaller than ϵ for all $w \in \omega_n \cap H$. $\qquad \square$

<u>Proof</u>

In this proof we identify $\frac{c_n}{\|c_n\|}$ with e_n . Then $-1 < \frac{c_n}{\|c_n\|}^T w = w^n < 1$.

For the beginning, we may abbreviate $M(w) := M^\delta(w)$ for a (so far) fixed δ.

Let $\Pi : \mathrm{I\!R}^n \to \mathrm{I\!R}^{n-1} \times \{0\}$ describe the orthogonal projection on $\mathrm{I\!R}^{n-1} \times \{0\}$, which sends $x = (x^1, \ldots, x^n)^T$ to $x = (x^1, \ldots, x^{n-1}, 0)^T$.

Then we discover an "equatorial counterpart" $K[M(w)]$ of $M(w)$ defined as

$$K[M(w)] := \Pi(\widetilde{\mathrm{cone}}(M(w))) \cap \{x \mid \|x\| = 1, x^n = 0\} \qquad \text{and} \quad \widetilde{\mathrm{cone}}(K[M(w)]). \tag{58}$$

Throughout this proof we use positive reals $\underline{\tau}$ and $\overline{\tau}$ to describe

$$\inf\{x^n | x \in M(w)\} := w^n - \underline{\tau} \ \text{and} \ \sup\{x^n | x \in M(w)\} := w^n + \overline{\tau}.$$

22

Hence $\forall\, x \in M(w):\ w^n - x^n \leq \underline{\tau}$ and $x^n - w^n \leq \overline{\tau}$. Without loss of generality we may assume that either $M(w) \subset \{x \mid \|x\| = 1, x^n \geq 0\}$ or $M(w) \subset \{x \mid \|x\| = 1, x^n \leq 0\}$. (If both signs exist in $M(w)$, we divide it in one part "above" and one part "below" the "equator" $\{x \mid \|x\| = 1, x^n = 0\}$ and apply the proof to both parts separately).

So, we start with the <u>positive case</u>. Here, we observe a set-inclusion for $\widetilde{\mathrm{cone}}(e_n, M(w)) \cap \omega_n$:

$$\widetilde{\mathrm{cone}}(e_n, K[M(w)]) \cap \omega_n \cap \{x | x^n \geq w^n + \overline{\tau}\} \quad \subset \quad \widetilde{\mathrm{cone}}(e_n, M(w)) \cap \omega_n$$

$$\subset \quad \widetilde{\mathrm{cone}}(e_n, K[M(w)]) \cap \omega_n \cap \{x | x^n \geq w^n - \underline{\tau}\}. \tag{59}$$

The consequence is

$$\lambda_{n-1}\left\{\widetilde{\mathrm{cone}}(e_n, K[M(w)]) \cap \omega_n \cap \{x | x^n \geq w^n + \overline{\tau}\}\right\} \quad \leq \quad \lambda_{n-1}\left\{\widetilde{\mathrm{cone}}(e_n, M(w)) \cap \omega_n\right\}$$

$$\leq \quad \lambda_{n-1}\left\{\widetilde{\mathrm{cone}}(e_n, K[M(w)]) \cap \omega_n \cap \{x | x^n \geq w^n - \underline{\tau}\}\right\}. \tag{60}$$

Application of (55) leads to

$$\frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \int_{w^n + \overline{\tau}}^{1} \sqrt{1 - h^2}^{\,n-3} dh \cdot \frac{\lambda_{n-1}\left\{\widetilde{\mathrm{cone}}(K[M(w)]) \cap \Omega_{n-1}\right\}}{\lambda_{n-1}\{\Omega_{n-1}\}} \tag{61}$$

$$\leq \quad \frac{\lambda_{n-1}\left\{\widetilde{\mathrm{cone}}(e_n, M(w)) \cap \omega_n\right\}}{\lambda_{n-1}\{\omega_n\}} = \frac{\lambda_n\left\{\widetilde{\mathrm{cone}}(e_n, M(w)) \cap \Omega_n\right\}}{\lambda_n\{\Omega_n\}} = \widetilde{V}(e_n, M(w)) \tag{62}$$

$$\leq \quad \frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \int_{w^n - \underline{\tau}}^{1} \sqrt{1 - h^2}^{\,n-3} dh \cdot \frac{\lambda_{n-1}\left\{\widetilde{\mathrm{cone}}(K[M(w)]) \cap \Omega_{n-1}\right\}}{\lambda_{n-1}\{\Omega_{n-1}\}}. \tag{63}$$

This inclusion result for $\widetilde{V}(e_n, M(w))$ from (62) should be compared with

$$\widetilde{W}(M(w)) = \frac{\lambda_{n-2}\{M(w) \cap \omega_n\}}{\lambda_{n-2}\{\omega_{n-1}\}} = \frac{\lambda_{n-1}\left\{\widetilde{\mathrm{cone}}(M(w) \cap \Omega_n)\right\}}{\lambda_{n-1}\{\Omega_{n-1}\}}.$$

For this quantity we make use of

$$\lambda_{n-1}\left\{\Pi(\widetilde{\mathrm{cone}}(M(w)) \cap \Omega_n)\right\} = \cos(\angle(e_n, z)) \cdot \lambda_{n-1}\left\{\widetilde{\mathrm{cone}}(M(w)) \cap \Omega_n\right\}. \tag{64}$$

But our set-inclusion delivers

$$\widetilde{\mathrm{cone}}(K[M(w)]) \cap \Omega_{n-1}\left(\sqrt{1 - [w^n + \overline{\tau}]^2}\right) \quad \subset \quad \Pi(\widetilde{\mathrm{cone}}(M(w)) \cap \Omega_n) \tag{65}$$

$$\subset \quad \widetilde{\mathrm{cone}}(K[M(w)]) \cap \Omega_{n-1}\left(\sqrt{1 - [w^n - \underline{\tau}]^2}\right),$$

and hence

$$\lambda_{n-1}\left\{\widetilde{\mathrm{cone}}(K[M(w)]) \cap \Omega_{n-1}\right\} \cdot \left(\sqrt{1 - [w^n + \overline{\tau}]^2}\right)^{n-1} =$$
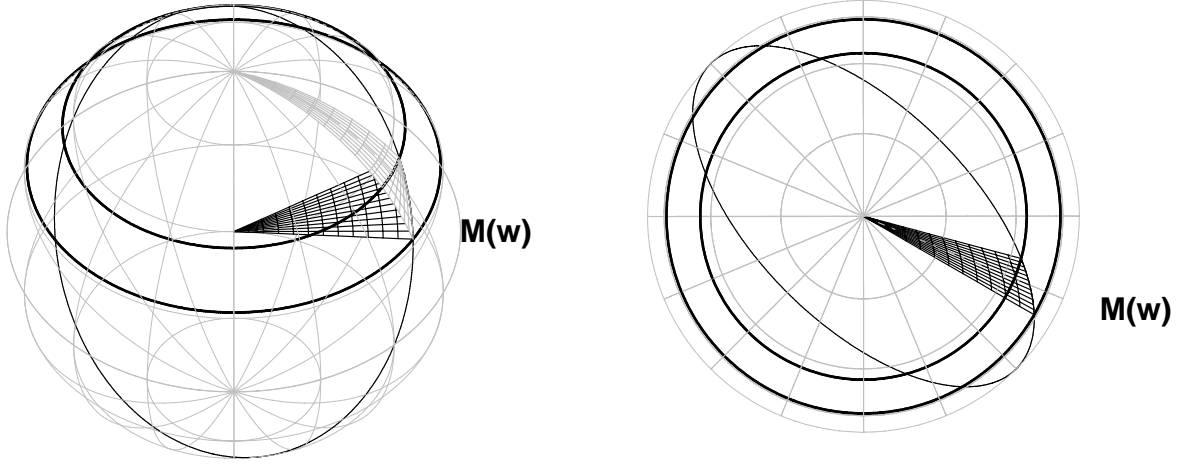
23

Figure 4: The left figure shows a spherical triangle generated by $e_3 = c_3$ and a small area $M(w) \subset H \cap \omega_n$. $M(w)$ has w as its central point and it is the intersection set of the spherical triangle with the ellipse illustrating the location of H. Also we see the conical sector generated by $M(w)$. Its size is proportional to $\widetilde{W}(M(w))$. In the right figure we look down from e_3 and recognize, that both areas have the same image when orthogonally projected on the equatorial plane $\mathrm{lin}(e_1, e_2)$. It gives a hint how to bound $\widetilde{W}(M(w))$. The projection image of $M(w)$ lies between two circles. Reduction of the cone to the inner circle resp. extension to the outer circle will provide a lower resp. an upper bound for the size of the image. If we take into account that the projection hits the conical sector under a certain angle, which causes a known reduction of its size under our projection, we can derive bounds for the size of $M(w)$.

$$= \lambda_{n-1}\left\{\widetilde{\mathrm{cone}}(K[M(w)]) \cap \Omega_{n-1}(\sqrt{1 - [w^n + \overline{\tau}]^2}\,)\right\}$$

$$\leq \lambda_{n-1}\left\{\Pi(\widetilde{\mathrm{cone}}(M(w)) \cap \Omega_n)\right\} = \cos(\angle(e_n, z)) \cdot \lambda_{n-1}\left\{\widetilde{\mathrm{cone}}(M(w)) \cap \Omega_n\right\} = \quad (66)$$

$$= \cos(\angle(e_n, z)) \cdot \lambda_{n-1}\{\Omega_{n-1}\} \cdot \widetilde{W}(M(w)) \leq \quad (67)$$

$$\leq \lambda_{n-1}\left\{\widetilde{\mathrm{cone}}(K[M(w)]) \cap \Omega_{n-1}(\sqrt{1 - [w^n - \underline{\tau}]^2}\,)\right\}$$

$$= \lambda_{n-1}\left\{\widetilde{\mathrm{cone}}(K[M(w)]) \cap \Omega_{n-1}\right\} \cdot \left(\sqrt{1 - [w^n - \underline{\tau}]^2}\,\right)^{n-1}.$$

Since $\cos(\angle(e_n, z)) > 0$, we can divide (62) by (67), and we can conclude from that result:

$$\frac{\frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \int\limits_{w^n + \overline{\tau}}^{1} \sqrt{1 - h^2}^{n-3} dh \cdot \cos(\angle(e_n, z))}{\left(\sqrt{1 - [w^n - \underline{\tau}]^2}\,\right)^{n-1}} = \quad (68)$$

$$= \frac{\frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \int\limits_{w^n + \overline{\tau}}^{1} \sqrt{1 - h^2}^{n-3} dh \cdot \lambda_{n-1}\{\widetilde{\mathrm{cone}}(K[M(w)]) \cap \Omega_{n-1}\} \cdot \cos(\angle(e_n, z))}{\lambda_{n-1}\{\widetilde{\mathrm{cone}}(K[M(w)]) \cap \Omega_{n-1}\} \cdot \left(\sqrt{1 - [w^n - \underline{\tau}]^2}\,\right)^{n-1}} \quad (69)$$

24

$$\leq \frac{\lambda_n \left\{ \widetilde{\mathrm{cone}}(e_n, M(w)) \cap \Omega_n \right\} \cdot \lambda_{n-1}\{\Omega_{n-1}\}}{\lambda_{n-1}\left\{ \widetilde{\mathrm{cone}}(M(w)) \cap \Omega_n \right\} \cdot \lambda_n\{\Omega_n\}} = \frac{\widetilde{V}(e_n, M(w))}{\widetilde{W}(M(w))} \tag{70}$$

$$\leq \frac{\frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \int\limits_{w^n - \underline{\tau}}^{1} \sqrt{1-h^2}^{\,n-3} dh \cdot \lambda_{n-1}\left\{ \widetilde{\mathrm{cone}}(K[M(w)]) \cap \Omega_{n-1} \right\} \cdot \cos(\angle(e_n, z))}{\lambda_{n-1}\left\{ \widetilde{\mathrm{cone}}(K[M(w)]) \cap \Omega_{n-1} \right\} \left(\sqrt{1 - [w^n + \overline{\tau}]^2} \right)^{n-1}} \tag{71}$$

$$= \frac{\frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \int\limits_{w^n - \underline{\tau}}^{1} \sqrt{1-h^2}^{\,n-3} dh \cdot \cos(\angle(e_n, z))}{\left(\sqrt{1 - [w^n + \overline{\tau}]^2} \right)^{n-1}}. \tag{72}$$

Note that not only $\frac{\widetilde{V}(e_n, M(w))}{\widetilde{W}(M(w))}$ from (70) but also $\hat{\Upsilon}(e_n, z, w)$ from (56) is bounded between (72) and (68). To study the difference of both, just have a short look on the geometry of $\omega_n \cap H$.
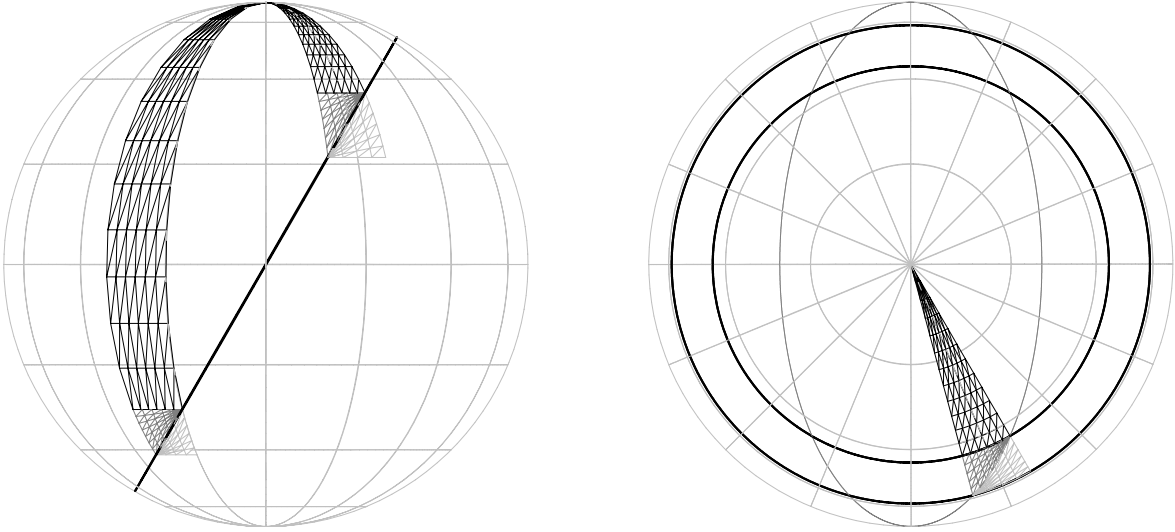

Figure 5: In the left figure we see the foreground of the unit ball and two spherical triangles generated by $e_3 = c_3$ and two examples of small areas $M(w) \subset H \cap \omega_3$. H is illustrated by the diagonal line. $M(w)$ is the respective border betweeen the small triangles at the bottom . The true size of $\widetilde{V}(e_3, M(w))$ results from adding the stripe with horizontal bottom line above H and the small triangle above H. We can underestimate the size of $V(e_3, M(w))$ by measuring only that stripe and ignoring the triangle. We can overestimate it by adding the stripe and both triangles, even that below H. This overestimation-method again delivers a (this time larger) stripe with horizontal bottom line. In the right figure we observe the conical sector generated by $M(w)$ of the right example from above (position e_3). Bounding the size of $\widetilde{W}(M(w))$ can be done analogously, using the fact that $M(w)$ completely lies in the ring between two level-circles, and proceeding as mentioned above.

Our hyperplane H is determined by the vector $z = (0, \ldots, 0, z^{n-1}, z^n)^T$. Taking the dependence $z = z[T, h]$ into regard, shows that $z^{n-1} = -\frac{h}{t}$ and $z^n = \frac{T}{t}$. We recall that

$z^n > 0$ and observe that the "highest point" (with maximal n-th coordinate) in $\omega_n \cap H$ is $z^\perp = (0, \ldots, 0, z^n, -z^{n-1})^T = (0, \ldots, 0, \frac{T}{t}, \frac{h}{t})^T$, where $0 \leq -z^{n-1} = \sqrt{1 - [z^n]^2} = \frac{h}{t} < 1$. It is clear that for any $y \in \omega_n \cap H$ we have

$$\sqrt{1 - [y^n]^2} \geq \sqrt{1 - \left[\frac{h}{t}\right]^2}, \quad \text{and this explains}$$

$$\sqrt{1 - [w^n - \underline{\tau}]^2} \geq \sqrt{1 - \left[\frac{h}{t}\right]^2} \quad \text{and} \quad \sqrt{1 - [w^n + \overline{\tau}]^2} \geq \sqrt{1 - \left[\frac{h}{t}\right]^2}.$$

Now let y be any point of $M(w)$ such that $\angle(y, w) < \delta$. Then we see that

$$w^n - \delta \cdot \frac{h}{t} \leq y^n \leq w^n + \delta \cdot \frac{h}{t} \quad \text{and} \quad \underline{\tau} \leq \delta \cdot \frac{h}{t} \quad \text{as well as} \quad \overline{\tau} \leq \delta \cdot \frac{h}{t}.$$

For showing the uniform convergence of the error mentioned in the claim to 0 on $\omega_n \cap H$ under a sequence of δ's decreasing to 0_+ , we simply have to study the difference between (72) and (68), more precisely the difference

$$\frac{\int_{w^n - \underline{\tau}}^{1} \sqrt{1 - h^2}^{n-3} dh}{\left(\sqrt{1 - [w^n + \overline{\tau}]^2}\right)^{n-1}} - \frac{\int_{w^n + \overline{\tau}}^{1} \sqrt{1 - h^2}^{n-3} dh}{\left(\sqrt{1 - [w^n - \underline{\tau}]^2}\right)^{n-1}}. \tag{73}$$

This difference converges uniformly to 0 (as a function of δ), as (for $w^n > 0$ and sufficiently small δ) it is bounded from above by

$$\frac{\int_{w^n - \delta \cdot \frac{h}{t}}^{1} \sqrt{1 - h^2}^{n-3} dh}{\left(\sqrt{1 - [w^n + \delta \cdot \frac{h}{t}]^2}\right)^{n-1}} - \frac{\int_{w^n + \delta \cdot \frac{h}{t}}^{1} \sqrt{1 - h^2}^{n-3} dh}{\left(\sqrt{1 - [w^n - \delta \cdot \frac{h}{t}]^2}\right)^{n-1}}. \tag{74}$$

Both quotients in (74) converge uniformly to

$$\frac{\int_{w^n}^{1} \sqrt{1 - h^2}^{n-3} dh}{\left(\sqrt{1 - [w^n]^2}\right)^{n-1}}. \tag{75}$$

In addition, this limit function (75) itself is a uniformly continuous function of w^n on the compact interval $[-\frac{h}{t}, +\frac{h}{t}]$. Hence also the difference in (73) tends to 0 uniformly for $\delta \to 0_+$ and for every value $w^n \in [-\frac{h}{t}, +\frac{h}{t}]$.

Similar convergence results can be obtained for the negative case, where $w^n < 0$ and $M(w)$ is completely in the lower hemisphere. Here, we see that

$$\frac{\frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \int_{w^n + \overline{\tau}}^{1} \sqrt{1 - h^2}^{n-3} dh \cdot \cos(\angle(e_n, z))}{\left(\sqrt{1 - [w^n + \overline{\tau}]^2}\right)^{n-1}} \leq \frac{\lambda_n \left\{\widetilde{\text{cone}}(e_n, M(w)) \cap \Omega_n\right\} \lambda_{n-1}\{\Omega_{n-1}\}}{\lambda_{n-1} \left\{\widetilde{\text{cone}}(M(w)) \cap \Omega_n\right\} \cdot \lambda_n\{\Omega_n\}} = \tag{76}$$

26

$$= \frac{\widetilde{V}(e_n, M(w))}{\widetilde{W}(M(w))} \leq \frac{\frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \int\limits_{w^n - \underline{\tau}}^{1} \sqrt{1 - h^2}^{n-3} dh \cdot \cos(\angle(e_n, z))}{\left(\sqrt{1 - [w^n - \underline{\tau}]^2}\right)^{n-1}} . \tag{77}$$

This case, where $M(w)$ is completely below the equator, leads to the difference

$$\frac{\int\limits_{w^n - \underline{\tau}}^{1} \sqrt{1 - h^2}^{n-3} dh}{\left(\sqrt{1 - [w^n - \underline{\tau}]^2}\right)^{n-1}} - \frac{\int\limits_{w^n + \overline{\tau}}^{1} \sqrt{1 - h^2}^{n-3} dh}{\left(\sqrt{1 - [w^n + \overline{\tau}]^2}\right)^{n-1}} .$$

Here, the claim results from the uniform continuity on the compact interval $w^n \in [-\frac{h}{t}, \frac{h}{t}]$ of

$$\frac{\int\limits_{w^n}^{1} \sqrt{1 - h^2}^{n-3} dh}{\left(\sqrt{1 - [w^n]^2}\right)^{n-1}} . \qquad \square$$

To show that $\Upsilon = \hat{\Upsilon}$ almost everywhere, is – in the view of Lemma 1 and Lemma 2 – a typical and pure task of measure theory. Therefore, we give the proof in the Appendix.

Lemma 3
Let c_n and z be fixed such that $\frac{c_n}{\|c_n\|}^T z \neq 0$. Consider the orthogonal hyperplane H to z.
Then with the exception of a λ_{n-2}-nullset, $\hat{\Upsilon}(\frac{c_n}{\|c_n\|}, z[T, h]; w)$ from (56) in Lemma 2 and $\Upsilon(\frac{c_n}{\|c_n\|}, z[T, h]; w)$ from Lemma 1, are identical functions on $\omega_n \cap H$. $\qquad \square$

At this point, Proposition 1 is proven.

Before proceeding to technical estimations, we want to tell something more about Υ, to illustrate the geometrical significance of that figure. Again, we simplify the notation by setting $e_n = \frac{c_n}{\|c_n\|}$.
Consider a small set $M(w) \subset \omega_n \cap H$. Since the spherical angle of $\widetilde{\text{cone}}(e_n, M(w))$ represents a certain share of the surface of ω_n, it also admits another interpretation. It can be described by the following three formulas

$$\text{spherical angle} = \text{horizontal extension} \times \text{depth-extension}, \tag{78}$$

$$\text{horizontal extension of } e_n \text{ on } M(w) := \frac{\lambda_{n-2}\{K[M(w)]\}}{\lambda_{n-2}\{x \mid \|x\| = 1, \ x^n = 0\}}, \tag{79}$$

$$\text{depth-extension of } e_n \text{ on } M(w) := \frac{\lambda_{n-1}\{x \mid \|x\| = 1, x^n \geq w^n\}}{\lambda_{n-1}\{x \mid \|x\| = 1\}} . \tag{80}$$

It is known that $\frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \cdot \sqrt{1 - h^2}^{n-3}$ is the marginal density of the share of ω_n below the level $x^n = h$.
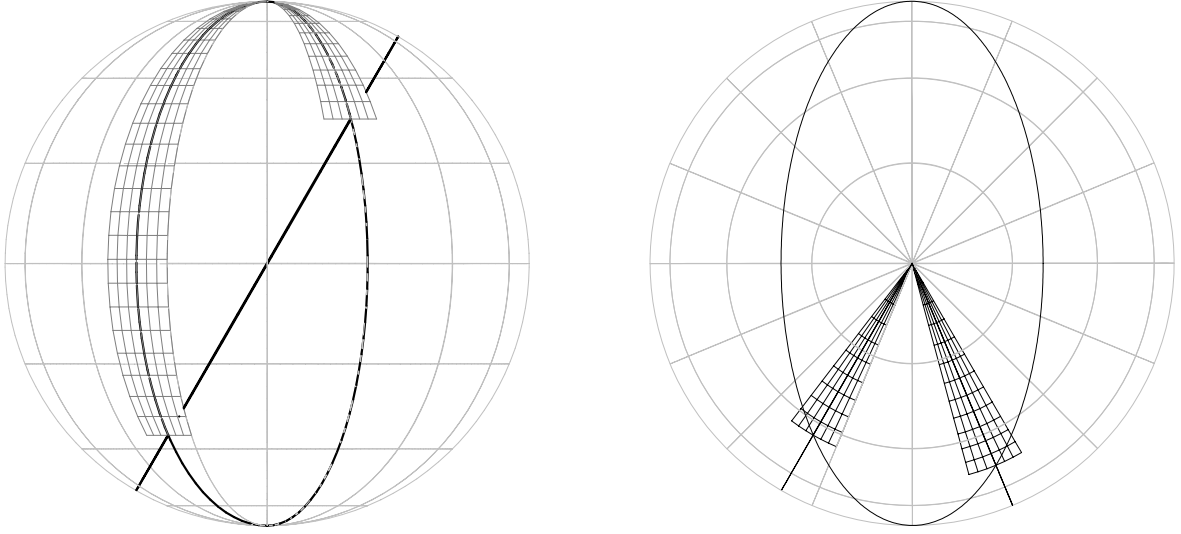
Figure 6: A chance for approximation of $\widetilde{V}(e_3, M(w))$ and $\widetilde{W}(M(w))$ comes from rectification of the sectors. We replace the spherical triangle by another one, which uniformly ends at level w^n. Note that a circle-movement from e_3 to $-e_3$ over w hits H (illustrated by the diagonal line) under a certain angle. The right figure describes the parallel approximation method for $\widetilde{W}(M(w))$.

If our position is in $y = e_n$ and if we look at $M(w)$ "along the sphere," then from our perspective $M(w)$ covers $K[M(w)]$, the "equatorial counterpart of $M(w)$" or the "shadow of $M(w)$ on the equator $\{x \mid x^n = 0\}$ under that projection."

If we want to reach all points of $M(w)$ from e_n, we may start movements on all half-circles $\omega_n \cap \mathrm{cone}(e_n, x, -e_n)$, for which $x \in K[M(w)]$, in direction to $\{x \mid x^n = 0\}$ until we arrive at $M(w)$. During such a movement the value y^n of our current point y is reduced. To reach $M(w)$, it may even be necessary to cross the equator. If $M(w)$ has a very small diameter, then we have to run approximately until $y^n = w^n$.

To understand the formula for the horizontal extension, imagine that we observe $M(w)$ from e_n. Since we are unable to measure the distance exactly, we may attribute this set to a certain horizon $\{x \mid \|x\| = 1, x^n = w^n\}$ of ω_n. The projection of $M(w)$ on that horizon can closely be approximated by $\{x \mid \|x\| = 1, x^n = w^n, \mathrm{cone}(e_n, x, -e_n) \cap M(w) \neq \emptyset\}$. We compare this set with the complete size of this horizon. Relying on that observation, we can predict, which share of the equator will be covered by the "equator-shadow of $M(w)$." That share depends on the size of $M(w)$, and on our perspective on $M(w)$. The perspective causes a reduction-factor $|\cos(\angle(z, \hat{w}^\perp))|$. z is the normal vector to H and $-\hat{w}^\perp$ is the tangential vector to our movement-circle $\omega_n \cap \mathrm{cone}(e_n, x, -e_n)$ at w, where we reach and traverse H. Our reduction factor reflects the divergence from an orthogonal movement across H, taking into regard that such a divergence causes the size of the shadow to shrink.

Remark 3
For the horizontal extension the following formula holds

$$\frac{\lambda_{n-2}\{K[M(w)]\}}{\lambda_{n-2}\{x \mid \|x\| = 1, \ x^n = 0\}} = \frac{\lambda_{n-2}\{M(w)\} \cdot |\cos(\angle(z, \hat{w}^\perp))|}{(1 - (w^n)^2)^{\frac{n-2}{2}} \cdot \lambda_{n-2}\{\omega_{n-1}\}}. \tag{81}$$

28

The depth-extension is

$$\frac{\lambda_{n-1}\{x \mid \|x\| = 1, x^n \geq w^n\}}{\lambda_{n-1}\{x \mid \|x\| = 1\}} = \frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \cdot \int\limits_{w^n}^{1} \sqrt{1 - h^2}^{\,n-3} dh. \qquad (82)$$

The total spherical angle is calculated by

$$\frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \cdot \int\limits_{w^n}^{1} \sqrt{1 - h^2}^{\,n-3} dh \cdot \frac{\lambda_{n-2}\{M(w)\} \cdot |\cos(\angle(z, \hat{w}^\perp))|}{(1 - (w^n)^2)^{\frac{n-2}{2}} \cdot \lambda_{n-2}\{\omega_{n-1}\}}. \qquad \square \qquad (83)$$

Of course, this formula holds for $w^n < 0$ as well. In that case, the horizontal extension is the same (symmetry between e_n and $-e_n$). Different is the behaviour of the depth-extension, where the traversed angle between e_n and w now exceeds $\frac{\pi}{2}$.

Remark 4
The expression $(1 - (w^n)^2)^{\frac{1}{2}} \cdot |\cos(\angle(z, \hat{w}^\perp))|$ tells the distance of the point e_n to the hyperplane H (whose normal vector is z). So it equals $|\cos(\angle(z, e_n))|$. $\square$

Proof
We exploit that $z, e_n, w, \hat{w}^\perp$ are normalized and that w is orthogonal to z and $\hat{w}^\perp$ as well.

$$e_n = (e_n{}^T \hat{w}^\perp) \hat{w}^\perp + (e_n{}^T w) w = (e_n{}^T \hat{w}^\perp) \hat{w}^\perp + w^n w = (1 - (w^n)^2)^{\frac{1}{2}} \hat{w}^\perp + w^n w$$

$$\implies z^T e_n = (1 - (w^n)^2)^{\frac{1}{2}} (z^T \hat{w}^\perp) \qquad \implies z^T e_n = (1 - (w^n)^2)^{\frac{1}{2}} |\cos(\angle(z, \hat{w}^\perp))|. \quad \square$$

Obviously, this purely geometrical interpretation yields the same extension-formula as before, namely Υ. Let us study this function in detail.

Lemma 4
The formula $\Upsilon(\frac{c_n}{\|c_n\|}, z; w)$ admits the following transformations.

$$\Upsilon(\frac{c_n}{\|c_n\|}, z; w) = \frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \cdot \int\limits_{\frac{c_n}{\|c_n\|}^T w}^{1} \sqrt{1 - h^2}^{\,n-3} dh \cdot \frac{|\cos(\angle(\frac{c_n}{\|c_n\|}, z))|}{(1 - (\frac{c_n}{\|c_n\|}^T w)^2)^{\frac{n-1}{2}}} =$$

$$= \frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \cdot \int\limits_{|\frac{c_n}{\|c_n\|}^T w|}^{1} \sqrt{1 - h^2}^{\,n-3} h \, dh \cdot \frac{|\cos(\angle(\frac{c_n}{\|c_n\|}, z))|}{(1 - (\frac{c_n}{\|c_n\|}^T w)^2)^{\frac{n-1}{2}}} \cdot \frac{\int\limits_{\frac{c_n}{\|c_n\|}^T w}^{1} \sqrt{1 - h^2}^{\,n-3} dh}{\int\limits_{|\frac{c_n}{\|c_n\|}^T w|}^{1} \sqrt{1 - h^2}^{\,n-3} h \, dh} =$$

$$= \left\langle \frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \cdot \frac{1}{n-1} \cdot |\cos(\angle(\frac{c_n}{\|c_n\|}, z))| \right\rangle \frac{\int\limits_{\frac{c_n}{\|c_n\|}^T w}^{1} \sqrt{1 - h^2}^{\,n-3} dh}{\int\limits_{|\frac{c_n}{\|c_n\|}^T w|}^{1} \sqrt{1 - h^2}^{\,n-3} h \, dh} =$$

$$= \left\langle \frac{\lambda_{n-1}\{\Omega_{n-1}\}}{n \cdot \lambda_n\{\Omega_n\}} \cdot \text{dist}\left(\frac{c_n}{\|c_n\|}, H\right) \right\rangle \frac{\int\limits_{\frac{c_n}{\|c_n\|}^T w}^{1} \sqrt{1 - h^2}^{\,n-3} dh}{\int\limits_{|\frac{c_n}{\|c_n\|}^T w|}^{1} \sqrt{1 - h^2}^{\,n-3} h \, dh}. \qquad \square \qquad (84)$$

29

The term in $\langle\ \rangle$ gives just the Cavalieri-estimation for the quotient of the spherical angles from Borgwardt (1987), when we consider $M(w)$ at the place of $\omega_n \cap \mathrm{cone}(c_1, \ldots, c_{n-1})$. Our chance of improving old bounds now lies in a careful analysis of the factor outside $\langle\ \rangle$. We regard this factor as a function of $\frac{c_n}{\|c_n\|}$ and w, briefly as a factor of $\frac{c_n}{\|c_n\|}^T w$.

Definition 2

$$We\ define: \quad \eta := \eta\left(\frac{c_n}{\|c_n\|}, w\right) := \frac{c_n}{\|c_n\|}^T w \qquad and\ the\ function \tag{85}$$

$$\Phi : (-1, +1) \to \mathbb{R}^+ \quad with \quad \Phi(\eta) := \frac{\int\limits_{\eta}^{1} \sqrt{1 - h^2}^{\,n-3} \, dh}{\int\limits_{|\eta|}^{1} \sqrt{1 - h^2}^{\,n-3} h \, dh}. \qquad \square \tag{86}$$

The function Φ tends to ∞ for $\frac{c_n}{\|c_n\|}^T w \to -1$. Already at $\frac{c_n}{\|c_n\|}^T w = 0$ it yields an enlargement of the denominator-integral in (43) by a factor $\frac{1}{\mu_n}$. In the following we use the notation

Definition 3

$$\mu_n := \frac{2 \cdot \lambda_{n-2}\{\omega_{n-1}\}}{(n-1)\lambda_{n-1}\{\omega_n\}} = \frac{2 \cdot \lambda_{n-1}\{\Omega_{n-1}\}}{n \cdot \lambda_n\{\Omega_n\}} \quad \forall n \geq 2. \qquad \square \tag{87}$$

From the Appendix of Borgwardt (1987) we know that the following relation holds

$$\sqrt{\frac{(n-2)}{2\pi}} \leq \frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \leq \sqrt{\frac{(n-1)}{2\pi}}. \tag{88}$$

This gives us information on the size of μ_n, which will be applied in the following forms:

Remark 5

$$\sqrt{\frac{2(n-2)}{(n-1)^2\pi}} \leq \mu_n = \frac{2 \cdot \lambda_{n-2}\{\omega_{n-1}\}}{(n-1)\lambda_{n-1}\{\omega_n\}} \leq \sqrt{\frac{2}{(n-1)\pi}}, \tag{89}$$

$$\mu_2 = \frac{2}{\pi} \quad and \quad \mu_3 = \frac{1}{2} \quad and \quad \mu_4 = \frac{4}{3\pi}, \tag{90}$$

$$for\ n > 3 : \mu_{n-1}^2 \leq \frac{2}{(n-2)\pi} \leq \frac{1}{n-1}, \qquad for\ n = 3 : \mu_{n-1}^2 = \frac{4}{\pi^2} \leq \frac{1}{2} = \frac{1}{n-1}, \tag{91}$$

$$for\ n > 3 : \mu_n \geq \frac{1}{n-1}, \qquad for\ n = 3 : \mu_n = \frac{1}{2} = \frac{1}{n-1}, \tag{92}$$

$$\frac{2}{\mu_n \cdot \mu_{n-1}} = \frac{\lambda_n\{\Omega_n\}n}{\lambda_{n-1}\{\Omega_{n-1}\}\mu_{n-1}} = \pi \cdot (n-1). \qquad \square \tag{93}$$

Our comparison of dealing with $\frac{V}{W}$ in the old and new approach shows, that we can still use the factor in $\langle\ \rangle$, which is uniform for all points $w \in H \cap \omega_n$. But for a precise estimation we have to correct this with a w-dependent factor $\Phi(\eta) = \Phi\left(\frac{c_n}{\|c_n\|}^T w\right)$ to manage the complication of the curvature. This dependence on w introduces an additional level of integration (the innermost integration), which makes the new approach more complicated.

5 Insertion into the integral-quotient

The purpose of this section is to bound the pointwise quotient $PQ(t)$ from (43) above by another integral-quotient, which is much simpler and easier to evaluate. We will prove

Proposition 2

$$PQ(t) \leq \frac{\lambda_n\{\Omega_n\}}{\lambda_{n-1}\{\Omega_{n-1}\}} \frac{n^2 \cdot \int\limits_0^t G(h)^{m-n} T^{-1} \int\limits_{\mathbb{R}^{n-1}} |T - c_n^{n-1}| f(c_n) \, d\overline{c}_n \, dh}{\int\limits_0^t G(h)^{m-n} T^{-1} \frac{h}{t} \int\limits_{\mathbb{R}^{n-1}} |T - c_n^{n-1}|^2 \frac{1}{\|c_n\|} \Psi(h, t, \|c_n\|) f(c_n) \, d\overline{c}_n \, dh} \tag{94}$$

$$where \quad \Psi(h, t, \|c_n\|) \quad is \ defined \ as \quad \frac{1}{\frac{h \cdot h}{t \cdot \|c_n\|} + (1 - \frac{h \cdot h}{t \cdot \|c_n\|})\mu_n}. \qquad \Box \tag{95}$$

Remember that we have fixed t. But still we are going to vary h (and implicitly T) over all possible values. Once again, z shall denote the normal vector to $H(0, c_1, \ldots, c_{n-1})$ with orientation towards e_n. Then z is uniquely characterized by the values T and h. Variation of h means that we have to vary z, too. So we should now release z and replace the ";" in the parameter list of Υ back to a ",". Inserting this into (43) yields

$$PQ(t) = \frac{n \cdot \int\limits_0^t G(h)^{m-n} T^{-1} \int\limits_{\mathbb{R}^{n-1}} |T - c_n^{n-1}| \int\limits_{\mathbb{R}^{n-2}} \cdots \int\limits_{\mathbb{R}^{n-2}} |\lambda_{n-2}\{\text{conv}(c_1, \ldots, c_{n-1})\}|^2}{\int\limits_0^t G(h)^{m-n} T^{-1} \int\limits_{\mathbb{R}^{n-1}} |T - c_n^{n-1}| \int\limits_{\mathbb{R}^{n-2}} \cdots \int\limits_{\mathbb{R}^{n-2}} |\lambda_{n-2}\{\text{conv}(c_1, \ldots, c_{n-1})\}|^2}$$
$$\frac{\int\limits_{\omega_n \cap \text{cone}(c_1, \ldots, c_{n-1})} d_{n-1}w \quad \cdot f(c_1) \cdots f(c_{n-1}) \, d\overline{c}_1 \cdots d\overline{c}_{n-1} f(c_n) \, d\overline{c}_n \, dh}{\int\limits_{\omega_n \cap \text{cone}(c_1, \ldots, c_{n-1})} \Upsilon(\frac{c_n}{\|c_n\|}, z, w) d_{n-1}w \cdot f(c_1) \cdots f(c_{n-1}) \, d\overline{c}_1 \cdots d\overline{c}_{n-1} f(c_n) \, d\overline{c}_n \, dh}. \tag{96}$$

It will pay to apply Fubini's Theorem to the integrals given in (96) and to interchange the order of integrations. We can place the integration over $\overline{c}_n$ on $\mathbb{R}^{n-1}$ on the innermost position and regard this as the first integration to be performed. Note that the interchange with the integrations over $\overline{c}_1 \cdots \overline{c}_{n-1}$ is absolutely harmless, since $f(c_1) \cdots f(c_{n-1})$ are density-functions and the remaining inner terms are all bounded. Note that the innermost integrals deliver

$$\int\limits_{\omega_n \cap \text{cone}(c_1, \ldots, c_{n-1})} 1 \cdot d_{n-1}w = W(c_1, \ldots, c_n) \cdot \lambda_{n-2}\{\omega_{n-1}\} \leq \lambda_{n-2}\{\omega_{n-1}\} \quad \text{and}$$

$$\int\limits_{\omega_n \cap \text{cone}(c_1, \ldots, c_{n-1})} \Upsilon(\frac{c_n}{\|c_n\|}, z[T, h], w) d_{n-1}w = V(c_1, \ldots, c_n) \cdot \lambda_{n-1}\{\omega_n\} \leq \lambda_{n-1}\{\omega_n\}.$$

This helps even in the last interchange with the integration over w on $\omega_n \cap H$ (for fixed values of $c_1, \ldots, c_{n-1}$), although there the unbounded function $\Upsilon(\frac{c_n}{\|c_n\|}, z[T, h], w)$ is used.

After that change in the order of integrations, we obtain

$$
PQ(t) = \frac{n\int\limits_0^t G(h)^{m-n}\cdot T^{-1}\int\limits_{\mathbb{R}^{n-2}}\cdots\int\limits_{\mathbb{R}^{n-2}}\int\limits_{\omega_n\cap\mathrm{cone}(c_1,\ldots,c_{n-1})}|\lambda_{n-2}\{\mathrm{conv}(c_1,\ldots,c_{n-1})\}|^2}{\int\limits_0^t G(h)^{m-n}T^{-1}\int\limits_{\mathbb{R}^{n-2}}\cdots\int\limits_{\mathbb{R}^{n-2}}\int\limits_{\omega_n\cap\mathrm{cone}(c_1,\ldots,c_{n-1})}|\lambda_{n-2}\{\mathrm{conv}(c_1,\ldots,c_{n-1})\}|^2}
$$

$$
\frac{\cdot\int\limits_{\mathbb{R}^{n-1}}|T-c_n^{n-1}|f(c_n)\,d\overline{c}_n\cdot d_{n-1}w\,f(c_1)\cdots f(c_{n-1})\,d\overline{\overline{c}}_1\cdots d\overline{\overline{c}}_{n-1}\,dh}{\cdot\int\limits_{\mathbb{R}^{n-1}}|T-c_n^{n-1}|\cdot\Upsilon(\frac{c_n}{\|c_n\|},z[T,h],w)f(c_n)\,d\overline{c}_n\cdot d_{n-1}w\,f(c_1)\cdots f(c_{n-1})\,d\overline{\overline{c}}_1\cdots d\overline{\overline{c}}_{n-1}\,dh}. \tag{97}
$$

It would be very helpful to know more about the critical function $\Upsilon(\frac{c_n}{\|c_n\|},z[T,h],w)$ and its inherent function $\Phi(\eta)$, as defined in (86). It is easy to see that the characteristic values of Φ are:

$$
\Phi(\eta) = \begin{cases} 1 & \text{for } \eta = 1 \\ \frac{1}{\mu_n} & \text{for } \eta = 0 \\ \infty & \text{for } \eta = -1. \end{cases} \tag{98}
$$

The following Lemma, describing the behaviour of Φ, will be proved in the Appendix.

Lemma 5

The term

$$
\Phi(\eta) = \frac{\int\limits_\eta^1 \sqrt{1-h^2}^{\,n-3}dh}{\int\limits_{|\eta|}^1 \sqrt{1-h^2}^{\,n-3}hdh} = \frac{\int\limits_\eta^1 \sqrt{1-h^2}^{\,n-3}dh}{(1-\eta^2)^{\frac{n-1}{2}}}\cdot(n-1) \tag{99}
$$

represents a monotonously decreasing, convex function of η in the interval $(-1,1)$. $\qquad\square$

$\Phi(\eta)$ is our "correction function" with respect to the old approach in Borgwardt (1987). More difficult is the treatment of $\int\limits_\eta^1 \sqrt{1-h^2}^{\,n-3}dh$. In the case $\eta \gg 0$ a simple approximation is given by

$$
\int\limits_\eta^1 \sqrt{1-h^2}^{\,n-3}dh \sim \int\limits_\eta^1 \sqrt{1-h^2}^{\,n-3}\cdot h\,dh = \frac{1}{n-1}\cdot(1-(\eta)^2)^{\frac{n-1}{2}}. \tag{100}
$$

If $1 > h > \eta > Const.$, then we have underestimated the integral at most by a constant factor. Now we derive an estimation for all $\eta > 0$. This technical result combined with Theorem 1 was the essential tool for saving a first factor of $\sqrt{n}$ in the main result of Borgwardt (1994a). Its proof is given in the Appendix.

Lemma 6

$$
\eta + (1-\eta)\frac{2}{(n+1)} \leq \frac{\int\limits_\eta^1 \sqrt{1-h^2}^{\,n-3}\cdot h\,dh}{\int\limits_\eta^1 \sqrt{1-h^2}^{\,n-3}dh} \leq \eta + (1-\eta)\frac{2\cdot\lambda_{n-2}\{\omega_{n-1}\}}{(n-1)\lambda_{n-1}\{\omega_n\}} \quad \forall\,\eta > 0. \tag{101}
$$

32

And for the correction factor Φ it is known that for all $\eta > 0$

$$\Phi(\eta) \geq \frac{1}{\eta + (1-\eta)\mu_n}. \qquad \square \qquad (102)$$

$\Phi(\eta)$ will contribute to a decrement of the total expression (97) in particular, when $\eta << 1$. Recall the definition of our extension–function Υ from (84).

$$\Upsilon(\frac{c_n}{\|c_n\|}, z[T,h], w) = \left\langle \frac{1}{n} \frac{\lambda_{n-1}\{\Omega_{n-1}\}}{\lambda_n\{\Omega_n\}} \left[\text{dist}\left(\frac{c_n}{\|c_n\|}, H(0, c_1, \ldots, c_{n-1}) \right) \right] \right\rangle \Phi(w^T \frac{c_n}{\|c_n\|}) \quad (103)$$

From elementary trigonometry we know the following two facts

Remark 6

$$\text{dist}\,(c_n, H(0, c_1, \ldots, c_{n-1})) = |T - c_n^{n-1}| \cdot \frac{h}{t}, \qquad (104)$$

$$\text{dist}\left(\frac{c_n}{\|c_n\|}, H(0, c_1, \ldots, c_{n-1}) \right) = |T - c_n^{n-1}| \cdot \frac{h}{t} \cdot \frac{1}{\|c_n\|}. \qquad \square \qquad (105)$$

Φ is a convex and monotonously decreasing function. Let us insert (103) and (105) into (97).

$$PQ(t) = \frac{\lambda_n\{\Omega_n\}}{\lambda_{n-1}\{\Omega_{n-1}\}} \frac{n^2 \cdot \int_0^t G(h)^{m-n} T^{-1} \int_{\mathbb{R}^{n-2}} \cdots \int_{\mathbb{R}^{n-2}} \int_{\omega_n \cap \text{cone}(c_1,\ldots,c_{n-1})}}{\int_0^t G(h)^{m-n} T^{-1} \frac{h}{t} \cdot \int_{\mathbb{R}^{n-2}} \cdots \int_{\mathbb{R}^{n-2}} \int_{\omega_n \cap \text{cone}(c_1,\ldots,c_{n-1})}}$$

$$\frac{\left[\int_{\mathbb{R}^{n-1}} |T - c_n^{n-1}| f(c_n)\, d\overline{c}_n \cdot \right]}{\left[\int_{\mathbb{R}^{n-1}} |T - c_n^{n-1}|^2 \cdot \frac{1}{\|c_n\|} \Phi(w^T \frac{c_n}{\|c_n\|}) f(c_n)\, d\overline{c}_n \cdot \right]}$$

$$\frac{\cdot |\lambda_{n-2}\{\text{conv}(c_1, \ldots, c_{n-1})\}|^2 d_{n-1} w\, f(c_1) \cdots f(c_{n-1})\, d\overline{\overline{c}}_1 \cdots d\overline{\overline{c}}_{n-1}\, dh}{\cdot |\lambda_{n-2}\{\text{conv}(c_1, \ldots, c_{n-1})\}|^2 d_{n-1} w\, f(c_1) \cdots f(c_{n-1})\, d\overline{\overline{c}}_1 \cdots d\overline{\overline{c}}_{n-1}\, dh}. \quad (106)$$

Let us think about the quotient of integrals in [ ], particularly about the term in the denominator. For simplification we try to exploit the convexity of Φ, which had been confirmed in Lemma 5 .

Lemma 7
We achieve a decrement of the denominator of (106), if we consistently use the point $\frac{h}{\|c_n\|} e_n$ instead of $\frac{c_n}{\|c_n\|}$ in the argument of Φ. That means

$$\int_{\mathbb{R}^{n-1}} |T - c_n^{n-1}|^2 \frac{1}{\|c_n\|} \Phi(w^T \frac{c_n}{\|c_n\|}) f(c_n) d\overline{c}_n \geq \int_{\mathbb{R}^{n-1}} |T - c_n^{n-1}|^2 \frac{1}{\|c_n\|} \Phi(\frac{h}{\|c_n\|} w^T e_n) f(c_n) d\overline{c}_n. \quad (107)$$

<u>Proof</u>

We consider a c_n and cumulate each time over a quadrupel of corresponding points c_{ni}, $i = 1, 2, 3, 4$, resp. over their normalized counterparts $\xi_i := \frac{1}{\|c_{ni}\|}c_{ni}$. Let $c_{n1} := c_n$ induce the following four points:

$$\xi_1 := \frac{1}{\|c_n\|}c_{n1} := \frac{1}{\|c_n\|}\begin{pmatrix} \bar{\bar{c}}_n \\ c_n^{n-1} \\ h \end{pmatrix}, \quad \xi_2 := \frac{1}{\|c_n\|}c_{n2} := \frac{1}{\|c_n\|}\begin{pmatrix} -\bar{\bar{c}}_n \\ c_n^{n-1} \\ h \end{pmatrix},$$

$$\xi_3 := \frac{1}{\|c_n\|}c_{n3} := \frac{1}{\|c_n\|}\begin{pmatrix} \bar{\bar{c}}_n \\ -c_n^{n-1} \\ h \end{pmatrix}, \quad \xi_4 := \frac{1}{\|c_n\|}c_{n4} := \frac{1}{\|c_n\|}\begin{pmatrix} -\bar{\bar{c}}_n \\ -c_n^{n-1} \\ h \end{pmatrix}. \quad (108)$$

Without loss of generality let $c_n^{n-1} \leq 0$. The barycenter of the four points c_{ni} lies in $\begin{pmatrix} \bar{\bar{0}} \\ 0 \\ h \end{pmatrix}$, the barycenter of the ξ_i is located in $\frac{h}{\|c_n\|}e_n = \frac{1}{\|c_n\|}\begin{pmatrix} \bar{\bar{0}} \\ 0 \\ h \end{pmatrix}$. In (107) all points c_{ni} have the same density f. c_{n1} and c_{n2} (the two of the points with identical negative $(n-1)$th coordinate) are even identically weighted (with exception of Φ), because in $|T - c_n^{n-1}|^2$ only this coordinate is relevant. The same holds for the two other points c_{n3} and c_{n4}.

w belongs to $\omega_n \cap \text{cone}(c_1, \ldots, c_{n-1})$, so it is clear that $w^n > 0$ and that $w^{n-1} > 0$.

The weight of the two points ξ_1 and ξ_2 will be greater than that of the pair ξ_3 and ξ_4, since $|T - c_n^{n-1}| > |T + c_n^{n-1}|$. Now the convexity of Φ yields:

$$\frac{|T - c_n^{n-1}|^2 \cdot (\Phi(w^T\xi_1) + \Phi(w^T\xi_2)) + |T + c_n^{n-1}|^2 \cdot (\Phi(w^T\xi_3) + \Phi(w^T\xi_4))}{|T - c_n^{n-1}|^2 \cdot 2 + |T + c_n^{n-1}|^2 \cdot 2} \geq$$

$$\geq \Phi\left(w^T \frac{|T - c_n^{n-1}|^2 \cdot (\xi_1 + \xi_2) + |T + c_n^{n-1}|^2 \cdot (\xi_3 + \xi_4)}{|T - c_n^{n-1}|^2 \cdot 2 + |T + c_n^{n-1}|^2 \cdot 2}\right) =: \Phi(w^T\xi). \quad (109)$$

Here ξ is a vector with the properties $\bar{\bar{\xi}} = 0$, $\xi^{n-1} \leq 0$, $\xi^n = \frac{h}{\|c_n\|}$. Since Φ increases while $\eta = w^T x$ decreases, and because

$$w^T\xi = w^n \cdot \xi^n + w^{n-1} \cdot \xi^{n-1} \leq w^n \cdot \xi^n = w^n \cdot \frac{h}{\|c_n\|} = \frac{h}{\|c_n\|}w^T e_n, \quad (110)$$

the replacement mentioned above yields a smaller value of the denominator in (106). $\quad\square$

Lemma 8

$$\Phi(\frac{h}{\|c_n\|}w^T e_n) \geq \frac{1}{\frac{w^n h}{\|c_n\|} + (1 - \frac{w^n h}{\|c_n\|})\mu_n} \geq \frac{1}{\frac{h \cdot h}{t \cdot \|c_n\|} + (1 - \frac{h \cdot h}{t \cdot \|c_n\|})\mu_n} =: \Psi(h, t, \|c_n\|). \quad (111)$$

That means that a replacement of $\Phi(w^T \frac{c_n}{\|c_n\|})$ by $\Psi(h, t, \|c_n\|)$ in the denominator of (106) increases the quotient from (106) even more than the replacement suggested by (107). $\quad\square$

<u>Proof</u>

Since in any case

$$\frac{h}{\|c_n\|}w^n = \frac{h}{\|c_n\|}w^T e_n = \eta \geq 0,$$

we may make use of $(102): \Phi(\eta) \geq \frac{1}{\eta+(1-\eta)\mu_n}$ $\forall \eta > 0$ in Lemma 6. But the figure

$$\frac{1}{\frac{w^n h}{\|c_n\|} + (1 - \frac{w^n h}{\|c_n\|})\mu_n}$$

in (111) is monotonously decreasing for growing w^n. And the greatest value of w^n in $H \cap \omega_n$ simply is $\frac{h}{t}$. $\qquad\qquad\square$

So we obtain

$$PQ(t) \leq \frac{\lambda_n\{\Omega_n\}}{\lambda_{n-1}\{\Omega_{n-1}\}} \frac{n^2 \cdot \int\limits_0^t G(h)^{m-n}T^{-1}\left[\int\limits_{\mathbb{R}^{n-1}}|T-c_n^{n-1}|f(c_n)\,d\overline{c}_n\right]}{\int\limits_0^t G(h)^{m-n}T^{-1}\frac{h}{t}\left[\int\limits_{\mathbb{R}^{n-1}}|T-c_n^{n-1}|^2 \cdot \frac{1}{\|c_n\|}\Psi(h,t,\|c_n\|)\,d\overline{c}_n\right]} \tag{112}$$

$$\frac{\left\langle \int\limits_{\mathbb{R}^{n-2}} ... \int\limits_{\mathbb{R}^{n-2}} \int\limits_{\omega_n\cap\mathrm{cone}(c_1,...,c_{n-1})} |\lambda_{n-2}\{\mathrm{conv}(c_1,...,c_{n-1})\}|^2 d_{n-1}w\, f(c_1)\cdots f(c_{n-1})\,d\overline{\overline{c}}_1...d\overline{\overline{c}}_{n-1}\right\rangle dh}{\left\langle \int\limits_{\mathbb{R}^{n-2}} ... \int\limits_{\mathbb{R}^{n-2}} \int\limits_{\omega_n\cap\mathrm{cone}(c_1,...,c_{n-1})} |\lambda_{n-2}\{\mathrm{conv}(c_1,...,c_{n-1})\}|^2 d_{n-1}w\, f(c_1)\cdots f(c_{n-1})\,d\overline{\overline{c}}_1...d\overline{\overline{c}}_{n-1}\right\rangle dh}$$

$$= \frac{\lambda_n\{\Omega_n\}}{\lambda_{n-1}\{\Omega_{n-1}\}} \frac{n^2 \cdot \int\limits_0^t G(h)^{m-n}T^{-1}\int\limits_{\mathbb{R}^{n-1}}|T-c_n^{n-1}|f(c_n)\,d\overline{c}_n\, dh}{\int\limits_0^t G(h)^{m-n}T^{-1}\frac{h}{t} \cdot \int\limits_{\mathbb{R}^{n-1}}|T-c_n^{n-1}|^2 \cdot \frac{1}{\|c_n\|}\Psi(h,t,\|c_n\|)f(c_n)\,d\overline{c}_n\, dh}. \tag{113}$$

Note that the integral in [ ] in the denominator has lost all dependence upon $c_1,\ldots,c_{n-1}$, while its counterpart in the numerator had been independent from $c_1,\ldots,c_{n-1}$ all the time. So we could change the order of integrations .

The integrals in $\langle\,\rangle$ depend exclusively on $(c_1,\ldots,c_{n-1})$, (neither on h nor on c_n) and they are identical in numerator and denominator. Hence, reduction is admitted.

Now Proposition 2 is proved and the possibility of an evaluation, based on (113), seems much more likely.

6   Partitioning the set of possible configurations

This section is devoted to preparing detailed and extensive case–studies on the remaining integral-quotient in the next section. For this purpose, we change to polar coordinates, and use $r := r(c_n) = \|c_n\|, h, \gamma(c_n) \in \omega_{n-1}$, such that

$$c_n = \begin{pmatrix} \sqrt{r^2 - h^2}\,\gamma(c_n) \\ h \end{pmatrix} \text{ and } \overline{c}_n = \sqrt{r^2 - h^2}\,\gamma(c_n), \text{ and we set } R := R(r,h) = \sqrt{r^2 - h^2}.$$

Now we can uniquely characterize our distribution by its "radial distribution function."

$$\text{Let } \mathcal{F} \text{ stand for the probability } \mathcal{F}(r) := P(\|x\| \le r) \text{ for } r \in [0, \infty]. \tag{114}$$

Analyzing (113), we exploit the fact, that for fixed $r(c_n)$ we have a sphere of radius R in height h, where the points $\overline{c}_n$ will vary. $d\gamma_R(\overline{c}_n)$ denotes the corresponding integration element. As a result of symmetry (compare Borgwardt (1987), page 157-158) we get:

$$\frac{1}{r(c_n)}\left[\int\limits_{\omega_{n-1}(R)} |T - c_n^{n-1}|^2 d\gamma_R(\overline{c}_n)\right] = \frac{1}{r(c_n)}\left[T^2 + \frac{1}{n-1}R^2\right] R^{n-2}\lambda_{n-2}\{\omega_{n-1}\}. \tag{115}$$

For the numerator we analogously obtain

$$
\begin{aligned}
\int\limits_{\omega_{n-1}(R)} |T - c_n^{n-1}|d\gamma_R(\overline{c}_n) &= \int\limits_{c_n^{n-1}\le -T\le 0} (T + |c_n^{n-1}|)d\gamma_R(\overline{c}_n) + \int\limits_{0\le T\le c_n^{n-1}} (|c_n^{n-1}| - T)d\gamma_R(\overline{c}_n) \\
&\quad + \int\limits_{-T\le c_n^{n-1}\le 0} (T + |c_n^{n-1}|)d\gamma_R(\overline{c}_n) + \int\limits_{0\le c_n^{n-1}\le T} (T - |c_n^{n-1}|)d\gamma_R(\overline{c}_n) \\
&= \int\limits_{-T\le c_n^{n-1}\le T} T\,d\gamma_R(\overline{c}_n) + \int\limits_{|T|\le |c_n^{n-1}|} |c_n^{n-1}|\,d\gamma_R(\overline{c}_n).
\end{aligned}
$$

So we have for $R < T$:

$$\le T \cdot \int\limits_{\omega_{n-1}(R)} d\gamma_R(\overline{c}_n) = T \cdot \lambda_{n-2}\{\omega_{n-1}\} \cdot R^{n-2}. \tag{116}$$

For $R > T$ we obtain:

$$
\begin{aligned}
&\le \int\limits_{\omega_{n-1}(R)} |c_n^{n-1}|d\gamma_R(\overline{c}_n) + \int\limits_{-T\le c_n^{n-1}\le T} (T - |c_n^{n-1}|)d\gamma_R(\overline{c}_n) \\
&\le \frac{2 \cdot \lambda_{n-3}\{\omega_{n-2}\}}{(n-2)\lambda_{n-2}\{\omega_{n-1}\}} \cdot R \cdot \lambda_{n-2}\{\omega_{n-1}\}R^{n-2} + T \int\limits_{\omega_{n-1}(R)} d\gamma_R(\overline{c}_n) \\
&= \frac{2 \cdot \lambda_{n-3}\{\omega_{n-2}\}}{(n-2)\lambda_{n-2}\{\omega_{n-1}\}}R^{n-1}\lambda_{n-2}\{\omega_{n-1}\} + T\lambda_{n-2}\{\omega_{n-1}\}R^{n-2} \\
&= \mu_{n-1} \cdot R \cdot \lambda_{n-2}\{\omega_{n-1}\}R^{n-2} + T \cdot \lambda_{n-2}\{\omega_{n-1}\} \cdot R^{n-2}. \tag{117}
\end{aligned}
$$

Translating the definition (95) resp. (111) of Ψ into the language of polar coordinates yields

$$\Psi(h, t, r) = \frac{1}{\frac{h \cdot h}{r \cdot t} + (1 - \frac{h \cdot h}{r \cdot t})\mu_n}. \tag{118}$$

(115), (116), (117) and our transformation deliver a new upper bound for (113).

$$PQ(t) \leq \frac{\lambda_n\{\Omega_n\}n^2 \int\limits_0^t G(h)^{m-n} T^{-1} \int\limits_h^1 \frac{R^{n-3}}{r^{n-2}}[T + I_{\{r>t\}}(r) \cdot \mu_{n-1}R] \, d\mathcal{F}(r) \, dh}{\lambda_{n-1}\{\Omega_{n-1}\} \int\limits_0^t G(h)^{m-n} T^{-1} \frac{h}{t} \int\limits_h^1 \frac{R^{n-3}}{r^{n-2}}(T^2 + \frac{1}{n-1}R^2)\frac{1}{r}\Psi(h,t,r) \, d\mathcal{F}(r) \, dh}. \tag{119}$$

Now we partition the area of integration $(r,h) \in [h,1] \times [0,t]$ into different subareas. For each part we estimate the corresponding integral-quotient. The very worst item of those upper bounds gives us – according to the principle of pointwise comparison – an upper bound for the complete integral-quotient. After an appropriate and feasible permutation in the order of integrations, we have the following result, which we will use as the basis for the evaluation on subsets in the following sections.

Proposition 3

$$PQ(t) \leq \frac{\lambda_n\{\Omega_n\}n^2}{\lambda_{n-1}\{\Omega_{n-1}\}} \cdot \tag{120}$$

$$\cdot \left\{ \frac{\int\limits_0^t \int\limits_0^r G(h)^{m-n} T^{-1} R^{n-3} r^{-n+2} \cdot T \, dh \, d\mathcal{F}(r) +}{\int\limits_0^t \int\limits_0^r G(h)^{m-n} T^{-1} h t^{-1} r^{-1} R^{n-3} r^{-n+2}[T^2 + \frac{1}{n-1}R^2]\Psi(h,t,r) \, dh \, d\mathcal{F}(r) +} \right.$$

$$\left. \frac{+ \int\limits_t^1 \int\limits_0^t G(h)^{m-n} T^{-1} R^{n-3} r^{-n+2}(T + \mu_{n-1}R) \, dh \, d\mathcal{F}(r)}{+ \int\limits_t^1 \int\limits_0^t G(h)^{m-n} T^{-1} h t^{-1} r^{-1} R^{n-3} r^{-n+2}[T^2 + \frac{1}{n-1}R^2]\Psi(h,t,r) \, dh \, d\mathcal{F}(r)} \right\} . \qquad \square$$

The partition-subsets will be:

$$\begin{align} B_1 &:= B_{11} \cup B_{12}, \text{where} \tag{121}\\ B_{11} &:= \{(r,h)|0 \leq h \leq \mu_n r \leq r \leq t\}\\ B_{12} &:= \{(r,h)|0 \leq h \leq \mu_n t \leq t < r\} \end{align}$$

$$B_2 := \{(r,h)|0 \leq r \leq t \wedge \mu_n r \leq h \leq \mu_n t\} \tag{122}$$

$$\begin{align} B_3 &:= B_{31} \cup B_{32}, \text{where} \tag{123}\\ B_{31} &:= \{(r,h)|\mu_n t \leq h \leq r \leq t\}\\ B_{32} &:= \{(r,h)|\mu_n t \leq h \leq t < r\} . \end{align}$$

Note that $h \leq r$ and $h \leq t$ are tautological in our construction.
For each subarea we are able to derive bounds on $\Psi(h,t,r)$.

Lemma 9

The function $\Psi(h, t, r)$ can be bounded from below on the dissection subsets as follows

$$\Psi(h, t, r) := \frac{1}{\frac{h \cdot h}{r \cdot t} + (1 - \frac{h \cdot h}{r \cdot t})\mu_n} \quad \geq \quad \begin{cases} \frac{1}{2\mu_n} & in\ B_1 \\[2mm] \frac{1}{2\frac{h}{r}} & in\ B_2 \\[2mm] \frac{1}{2\frac{h}{r}} & in\ B_{31} \\[2mm] \frac{1}{2\frac{h}{t}} & in\ B_{32} \end{cases} \qquad \Box \qquad (124)$$

Proof:

We use $\mu_n < 1$ and the fact that a replacement of $\frac{h \cdot h}{r \cdot t}$ by the same upper bound at both places in the denominator will increase the denominator and decrease the quotient.

In B_1 we know $\frac{h \cdot h}{r \cdot t} \leq \mu_n$ and conclude

$$\Psi(h, t, r) = \frac{1}{\frac{h \cdot h}{r \cdot t} + (1 - \frac{h \cdot h}{r \cdot t})\mu_n} \geq \frac{1}{\mu_n + (1 - \mu_n)\mu_n} \geq \frac{1}{2\mu_n}. \qquad (125)$$

In B_2 we have $\frac{h \cdot h}{r \cdot t} \leq \frac{h}{r}$ and $\mu_n \leq \frac{h}{r}$. So

$$\Psi(h, t, r) = \frac{1}{\frac{hh}{rt} + (1 - \frac{hh}{rt})\mu_n} \geq \frac{1}{\frac{h}{r} + \mu_n} \geq \frac{1}{2\frac{h}{r}}. \qquad (126)$$

In B_{31} we replace $\frac{h \cdot h}{r \cdot t}$ by $\frac{h}{r}$ and obtain with $\mu_n \leq \frac{h}{t} \leq \frac{h}{r}$:

$$\Psi(h, t, r) = \frac{1}{\frac{h \cdot h}{r \cdot t} + (1 - \frac{h \cdot h}{r \cdot t})\mu_n} \geq \frac{1}{\frac{h}{r} + (1 - \frac{h}{r})\mu_n} \geq \frac{1}{2\frac{h}{r}}. \qquad (127)$$

And in B_{32} the inequality $\frac{h \cdot h}{r \cdot t} \leq \frac{h}{t}$ yields with $\mu_n \leq \frac{h}{t}$:

$$\Psi(h, t, r) = \frac{1}{\frac{h \cdot h}{r \cdot t} + (1 - \frac{h \cdot h}{r \cdot t})\mu_n} \geq \frac{1}{\frac{h}{t} + (1 - \frac{h}{t})\mu_n} \geq \frac{1}{2\frac{h}{t}}. \qquad \Box \qquad (128)$$

The purpose of the partition is the following:

In (119) or (120) the denominator contains some factors not appearing in the numerator, e.g.

$$\frac{h}{tr} \quad or \quad \Psi(h, t, r) \quad or \quad \frac{[T^2 + \frac{1}{n-1}R^2]}{(T + \mu_{n-1}R)}.$$

Each of these factors will become (relatively) small somewhere in the integration area, which causes the quotient to get large. The impact of that fact would be dramatic, if this happened simultaneously, i.e. if two or more of these will be small at the same configuration of integration variables.

Now the given partition enables us to show that in each of the subsets from (121), (122), (123) only one of these factors becomes critical. So, it will be possible to bound the "inverse expectation" of the mentioned "product" from above on each subset. But this has to be done carefully, using different methods and estimations, which are adapted to the local
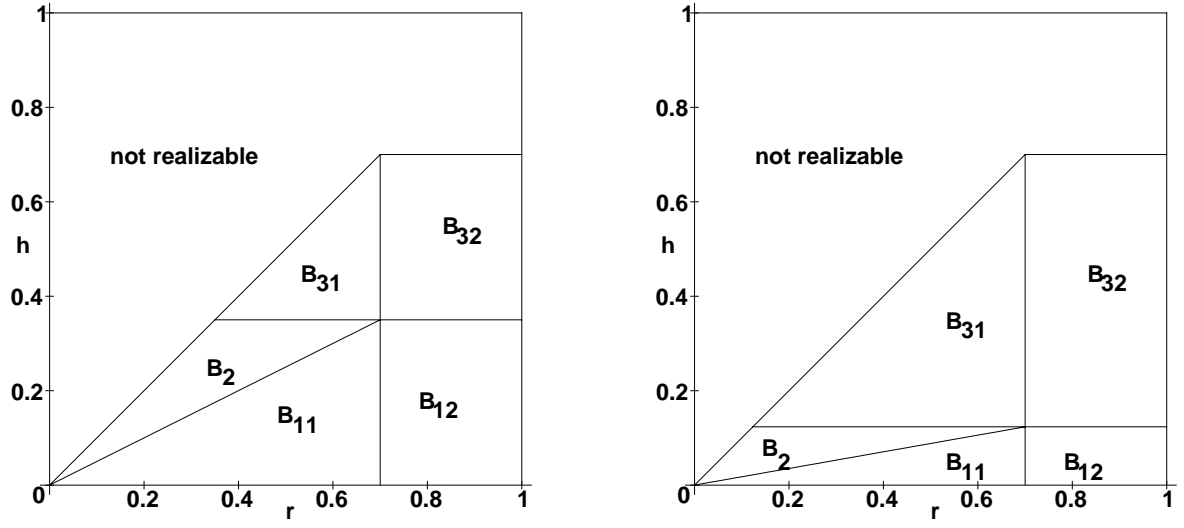
Figure 7: The partition into subsets for $t = 0.7$ and $n = 3$ (left) and $n = 21$ (right). The vertical line is located at $r = t$, and the horizontal line at $\mu_n t$. The diagonal line illustrates $h = \mu_n r$. Note that for growing n the value of μ_n decreases.

configuration of the respective subset.

The following consequence of a Binomial Theorem and of (91) lets us handle the third factor and will be applied several times. Its use was the crucial initial step in finding a chain of simplifying underestimations for the denominator, such that comparisons between numerator and denominator (e.g. with the help of Jensen's inequality) became possible.

Remark 7

$$[T + \mu_{n-1}R]^2 = T^2 + 2\mu_{n-1}RT + \mu_{n-1}{}^2R^2 \leq 2[T^2 + \mu_{n-1}{}^2R^2] \leq 2[T^2 + \frac{1}{n-1}R^2]. \quad (129)$$

7 Upper bounds for partial pointwise quotients

This section deals with deriving upper bounds for (120) on the different subareas. The corresponding integral-quotients will be denoted by $Q_1, Q_{11}, Q_{12}, Q_2, Q_3, Q_{31}, Q_{32}$.

Proposition 4
In $B_1 = \{(r,h)|0 \leq h \leq \mu_n r \leq r \leq t\} \cup \{(r,h)|0 \leq h \leq \mu_n t \leq t < r\}$ we have

$$Q_1 \leq \frac{\lambda_n\{\Omega_n\}n^2 8e^{\frac{1}{\pi}}}{\lambda_{n-1}\{\Omega_{n-1}\}\mu_{n-1}\sqrt{1-\mu_n{}^2}}. \tag{130}$$

<u>Proof</u>

$$
\begin{aligned}
Q_{11} \quad &\leq \quad \frac{\lambda_n\{\Omega_n\}n^2}{\lambda_{n-1}\{\Omega_{n-1}\}} \cdot \left\{ \frac{\int\limits_0^t \int\limits_0^{\mu_n r} G(h)^{m-n} T^{-1} R^{n-3} r^{-n+2} T \, dh \, d\mathcal{F}(r)}{\int\limits_0^t \int\limits_0^{\mu_n r} G(h)^{m-n} T^{-1} \frac{h}{tr} R^{n-3} r^{-n+2}(T^2 + \frac{1}{n-1}R^2) \cdot \frac{1}{2\mu_n} \, dh \, d\mathcal{F}(r)} \right\} \\[2mm]
&\leq \quad \frac{\lambda_n\{\Omega_n\}n^2 2\mu_n}{\lambda_{n-1}\{\Omega_{n-1}\}} \cdot \left\{ \frac{\int\limits_0^t \int\limits_0^{\mu_n r} G(h)^{m-n} R^{n-3} r^{-n+2} \, dh \, d\mathcal{F}(r)}{\int\limits_0^t \int\limits_0^{\mu_n r} G(h)^{m-n} R^{n-3} r^{-n+2}\frac{hT}{rt} \, dh \, d\mathcal{F}(r)} \right\} \\[2mm]
&\leq \quad \frac{\lambda_n\{\Omega_n\}n^2 2\mu_n}{\lambda_{n-1}\{\Omega_{n-1}\}} \cdot \sup_{0 \leq r \leq t} \left\{ \frac{\int\limits_0^{\mu_n r} G(h)^{m-n} R^{n-3} r^{-n+2} \, dh}{\int\limits_0^{\mu_n r} G(h)^{m-n} R^{n-3} r^{-n+2}\frac{hT}{rt} \, dh} \right\}. \tag{131}
\end{aligned}
$$

In B_{11}, we have $\frac{T}{t} \geq \sqrt{1-\mu_n{}^2}$, since $\frac{h}{t} \leq \frac{h}{r} \leq \mu_n$. The remaining objective in the denominator $\frac{h}{r}$ increases with h, as $G(h)$ does. So, dropping $G(h)^{m-n}$ increases the quotient.

$$
\begin{aligned}
Q_{11} \quad &\leq \quad \frac{\lambda_n\{\Omega_n\}n^2 2\mu_n}{\lambda_{n-1}\{\Omega_{n-1}\}\sqrt{1-\mu_n{}^2}} \cdot \max_{0 \leq r \leq t} \left\{ \frac{\int\limits_0^{\mu_n r} R^{n-3} r^{-n+2} \, dh}{\int\limits_0^{\mu_n r} \frac{h}{r} R^{n-3} r^{-n+2} \, dh} \right\} \\[2mm]
&= \quad \frac{\lambda_n\{\Omega_n\}n^2 2\mu_n}{\lambda_{n-1}\{\Omega_{n-1}\}\sqrt{1-\mu_n{}^2}} \cdot \left\{ \frac{\int\limits_0^{\mu_n t} T^{n-3} t^{-n+2} \, dh}{\int\limits_0^{\mu_n t} \frac{h}{t} T^{n-3} t^{-n+2} \, dh} \right\}. \tag{132}
\end{aligned}
$$

Analogously, we get for the set of radii with $t < r$ (using t_+ as lower bound for the integration area in order to exclude t):

$$
Q_{12} \quad \leq \quad \frac{\lambda_n\{\Omega_n\}n^2}{\lambda_{n-1}\{\Omega_{n-1}\}} \cdot \frac{\int\limits_{t_+}^1 \int\limits_0^{\mu_n t} G(h)^{m-n} T^{-1} R^{n-3} r^{-n+2}(T + \mu_{n-1}R) \, dh \, d\mathcal{F}(r)}{\int\limits_{t_+}^1 \int\limits_0^{\mu_n t} G(h)^{m-n} T^{-1} \frac{h}{tr} R^{n-3} r^{-n+2}(T^2 + \frac{1}{n-1}R^2) \cdot \frac{1}{2\mu_n} \, dh \, d\mathcal{F}(r)}
$$

$$\leq \quad \frac{\lambda_n\{\Omega_n\}n^2 2\mu_n}{\lambda_{n-1}\{\Omega_{n-1}\}} \cdot \sup_{t<r\leq 1} \left\{ \frac{\int_0^{\mu_n t} G(h)^{m-n}\, T^{-1}\, R^{n-3}\, r^{-n+2}\, (T + \mu_{n-1}R)\, dh}{\int_0^{\mu_n t} G(h)^{m-n} T^{-1} \frac{h}{tr} R^{n-3} r^{-n+2}(T^2 + \frac{1}{n-1}R^2)\, dh} \right\}. \quad (133)$$

Here, we apply Remark 7 resp. (129).

$$Q_{12} \leq \frac{\lambda_n\{\Omega_n\}n^2 2\mu_n \cdot 2}{\lambda_{n-1}\{\Omega_{n-1}\}} \cdot \sup_{t<r\leq 1} \left\{ \frac{\int_0^{\mu_n t} G(h)^{m-n}\, T^{-1}\, R^{n-3}\, r^{-n+2}\, (T + \mu_{n-1}R)\, dh}{\int_0^{\mu_n t} G(h)^{m-n} T^{-1} \frac{h}{tr} R^{n-3} r^{-n+2}(T + \mu_{n-1}R)^2\, dh} \right\}. \quad (134)$$

The true objective in the denominator is $\frac{h}{tr}(T + \mu_{n-1}R)$. Replacing it by a smaller objective will deliver a higher upper bound for Q_{12} in general. So, we make two replacements of this kind.

$$\frac{h}{tr}(T + \mu_{n-1}R) \geq \frac{h}{tr}(\mu_{n-1}R) = \frac{h}{t} \cdot \mu_{n-1} \cdot \frac{R}{r} \geq \frac{h}{t} \cdot \mu_{n-1} \cdot \sqrt{1 - \mu_n^2}. \quad (135)$$

Here, we have exploited that $\frac{R}{r} \geq \sqrt{1 - \mu_n^2}$ on B_{12}.

This figure (the reduced objective) increases proportionally with h. G supports larger values of h. Hence we are allowed to drop the G-terms once more. The same applies to the product $T^{-1} \cdot (T + \mu_{n-1}R)$, because $T^{-1}R$ increases with h when $r > t$, which holds in B_{12}.

$$\begin{aligned}
Q_{12} &\leq \frac{\lambda_n\{\Omega_n\}n^2\mu_n \cdot 4}{\lambda_{n-1}\{\Omega_{n-1}\}\mu_{n-1}\sqrt{1 - \mu_n^2}} \cdot \sup_{t<r\leq 1} \left\{ \frac{\int_0^{\mu_n t} G(h)^{m-n}\, T^{-1}\, R^{n-3}\, r^{-n+2}\, (T + \mu_{n-1}R)\, dh}{\int_0^{\mu_n t} G(h)^{m-n} T^{-1} \frac{h}{t} R^{n-3} r^{-n+2}(T + \mu_{n-1}R)\, dh} \right\} \\[2ex]
&\leq \frac{\lambda_n\{\Omega_n\}n^2\mu_n \cdot 4}{\lambda_{n-1}\{\Omega_{n-1}\}\mu_{n-1}\sqrt{1 - \mu_n^2}} \cdot \sup_{t<r\leq 1} \left\{ \frac{\int_0^{\mu_n t} R^{n-3}\, r^{-n+2}\, dh}{\int_0^{\mu_n t} \frac{h}{t} R^{n-3} r^{-n+2}\, dh} \right\}. \quad (136)
\end{aligned}$$

Without any effect, we may replace r^{-n+2} by t^{-n+2} (in numerator and denominator). Substituting T^{n-3} for R^{n-3} will support the lower values of $\frac{h}{t}$. This means that this substitution will deteriorate our bound. Hence we know

$$Q_{12} \leq \frac{\lambda_n\{\Omega_n\}n^2\mu_n \cdot 4}{\lambda_{n-1}\{\Omega_{n-1}\}\mu_{n-1}\sqrt{1 - \mu_n^2}} \cdot \frac{\int_0^{\mu_n t} T^{n-3}\, t^{-n+2}\, dh}{\int_0^{\mu_n t} \frac{h}{t} T^{n-3} t^{-n+2}\, dh}. \quad (137)$$

So the remaining integral-quotient gets a common pattern for Q_{11} and Q_{12}. Since the upper bound of (137) exceeds that of (132), it suffices to evaluate (137) further.

On $[0, \mu_n t]$ the term $\frac{T^{n-3}}{t^{n-3}} = \sqrt{1 - \frac{h^2}{t^2}}^{n-3}$ is almost constant, because it is 1 at $h = 0$, it is monotonously decreasing in the interior of the interval, and at $h = \mu_n t$ we have (compare (89))

$$\sqrt{1 - \mu_n^2}^{n-3} \geq \sqrt{1 - \frac{2}{(n-1)\pi}}^{n-3} = \left(1 - \frac{2}{\pi}\frac{1}{n-1}\right)^{\frac{1}{2}(n-3)} > e^{-\frac{1}{\pi}}.$$

Enlarging $(\frac{T}{t})^{n-3}$ in the numerator to 1 and reducing it to $e^{-\frac{1}{\pi}}$ in the denominator delivers

$$Q_1 \leq \frac{\lambda_n\{\Omega_n\}n^2\mu_n \cdot 4e^{\frac{1}{\pi}}}{\lambda_{n-1}\{\Omega_{n-1}\}\mu_{n-1}\sqrt{1-\mu_n^2}} \cdot \frac{\int_0^{\mu_n t} \frac{1}{t}dh}{\int_0^{\mu_n t} \frac{h}{t^2}\,dh} = \frac{\lambda_n\{\Omega_n\} \cdot n^2 \cdot 8 \cdot e^{\frac{1}{\pi}}}{\lambda_{n-1}\{\Omega_{n-1}\}\mu_{n-1}\sqrt{1-\mu_n^2}}. \ \square \qquad (138)$$

Proposition 5
For $B_2 = \{(r,h)|0 \leq r \leq t \wedge \mu_n r \leq h \leq \mu_n t\}$ we obtain

$$Q_2 \leq \frac{2 \cdot \lambda_n\{\Omega_n\} \cdot n^2}{\lambda_{n-1}\{\Omega_{n-1}\}\sqrt{1-\mu_n^2}}. \qquad (139)$$

<u>Proof</u>

$$Q_2 \ \leq \ \frac{\lambda_n\{\Omega_n\}n^2}{\lambda_{n-1}\{\Omega_{n-1}\}} \cdot \frac{\int_0^t \int_{\mu_n r}^{\mu_n t} G(h)^{m-n} R^{n-3} r^{-n+2}\,dh\,d\mathcal{F}(r)}{\int_0^t \int_{\mu_n r}^{\mu_n t} G(h)^{m-n}\frac{h}{trT}[T^2 + \frac{1}{n-1}R^2]R^{n-3}r^{-n+2}\frac{r}{2h}\,dh\,d\mathcal{F}(r)}$$

$$= \ \frac{\lambda_n\{\Omega_n\}n^2 \cdot 2}{\lambda_{n-1}\{\Omega_{n-1}\}} \cdot \frac{\int_0^t \int_{\mu_n r}^{\mu_n t} G(h)^{m-n} R^{n-3} r^{-n+2}\,dh\,d\mathcal{F}(r)}{\int_0^t \int_{\mu_n r}^{\mu_n t} G(h)^{m-n}\frac{1}{tT}[T^2 + \frac{1}{n-1}R^2]R^{n-3}r^{-n+2}\,dh\,d\mathcal{F}(r)}$$

$$\leq \ \frac{\lambda_n\{\Omega_n\}n^2 \cdot 2}{\lambda_{n-1}\{\Omega_{n-1}\}} \cdot \frac{\int_0^t \int_{\mu_n r}^{\mu_n t} G(h)^{m-n} R^{n-3} r^{-n+2}\,dh\,d\mathcal{F}(r)}{\int_0^t \int_{\mu_n r}^{\mu_n t} G(h)^{m-n}\frac{T}{t}R^{n-3}r^{-n+2}\,dh\,d\mathcal{F}(r)}$$

$$\leq \ \frac{\lambda_n\{\Omega_n\}n^2 \cdot 2}{\lambda_{n-1}\{\Omega_{n-1}\}\sqrt{1-\mu_n^2}} \qquad \text{because of } \frac{T}{t} \geq \sqrt{1-\mu_n^2} \text{ in } B_2. \qquad \square$$

Now only the rest of the integral $(h > \mu_n t)$ remains for evaluation. Here we cannot ignore the influence of the monotonously increasing function $G(h)$, since it forces $\frac{R}{r}$ to become very small (it pushes h to the top, i.e. to t). Therefore we handle the growth of $G(h)$ in another way this time.

Proposition 6
In $B_3 = B_{31} \cup B_{32} = \{(r,h)|\mu_n t \leq h \leq r \leq t\} \cup \{(r,h)|\mu_n t \leq h \leq t < r\}$ it is true that

$$Q_3 \leq \pi(n-1)n4e \cdot \left[\frac{n(1 + \mu_{n-1})}{\mu_{n-1}}\right]^{\frac{1}{n-1}} \max\left\{\left(\frac{m-n+1}{2}\right)^{\frac{1}{n-1}}, \frac{1}{\sqrt{1-(\mu_n)^2}}\right\}. \qquad (140)$$

Consider the term (120) for the third area B_3

$$Q_3 \leq \frac{\lambda_n\{\Omega_n\}n^2}{\lambda_{n-1}\{\Omega_{n-1}\}} \cdot \left\{ \frac{\int\limits_{\mu_n t}^{t}\int\limits_{\mu_n t}^{r} G(h)^{m-n} R^{n-3} r^{-n+2}\, dh\, d\mathcal{F}(r) + }{\int\limits_{\mu_n t}^{t}\int\limits_{\mu_n t}^{r} G(h)^{m-n}\frac{h}{t\,r\,T} R^{n-3} r^{-n+2}[T^2 + \frac{1}{n-1}R^2]\frac{r}{2h}\, dh\, d\mathcal{F}(r) + } \right.$$

$$\left. \frac{+\int\limits_{t_+}^{1}\int\limits_{\mu_n t}^{t} G(h)^{m-n}[1 + \mu_{n-1}\frac{R}{T}]R^{n-3} r^{-n+2}\, dh\, d\mathcal{F}(r)}{+\int\limits_{t_+}^{1}\int\limits_{\mu_n t}^{t} G(h)^{m-n}\frac{h}{t\,r\,T} R^{n-3} r^{-n+2}[T^2 + \frac{1}{n-1}R^2]\frac{t}{2h}\, dh\, d\mathcal{F}(r)} \right\} \quad (141)$$

$$= \frac{\lambda_n\{\Omega_n\}n^2 2}{\lambda_{n-1}\{\Omega_{n-1}\}} \cdot \quad (142)$$

$$\cdot \frac{\int\limits_{\mu_n t}^{t}\int\limits_{\mu_n t}^{r} G(h)^{m-n}\frac{R^{n-3}}{r^{n-2}}\, dh\, d\mathcal{F}(r) + \int\limits_{t_+}^{1}\int\limits_{\mu_n t}^{t} G(h)^{m-n}[1 + \mu_{n-1}\frac{R}{T}]\frac{R^{n-3}}{r^{n-2}}\, dh\, d\mathcal{F}(r)}{\int\limits_{\mu_n t}^{t}\int\limits_{\mu_n t}^{r} G(h)^{m-n}\frac{T}{t}[1 + \frac{1}{n-1}\frac{R^2}{T^2}]\frac{R^{n-3}}{r^{n-2}}\, dh\, d\mathcal{F}(r) + \int\limits_{t_+}^{1}\int\limits_{\mu_n t}^{t} G(h)^{m-n}\frac{T}{r}\frac{R^{n-3}}{r^{n-2}}[1 + \frac{1}{n-1}\frac{R^2}{T^2}]\, dh\, d\mathcal{F}(r)}$$

The following Lemma 10 will be proved in the Appendix – due to the lengthy and technical character of the proof. It clarifies that it is possible to work (in our analysis) with a distribution concentrating all its weight on the radii in the interval $[t, 1]$.

Lemma 10
There is a radial distribution function $\hat{\mathcal{F}}^{(t)}$ with $\hat{\mathcal{F}}^{(t)}(r) = 0\ \forall\ r < t$ and $\hat{\mathcal{F}}^{(t)}(1) = 1$, such that the quotient corresponding to (142) when based on the distribution function $\hat{\mathcal{F}}^{(t)}$ becomes worse (larger) than (142), which is based on $\mathcal{F}$. One instance of such a distribution function is of the form

$$\hat{\mathcal{F}}^{(t)}(r) = \begin{cases} 0 & \text{for } r < t \\ \gamma_t \cdot \mathcal{F}(t) & \text{for } r = t \\ \gamma_t \cdot \mathcal{F}(t) + [1 - \gamma_t \cdot \mathcal{F}(t)] \cdot \frac{\mathcal{F}(r) - \mathcal{F}(t)}{1 - \mathcal{F}(t)} & \text{for } r > t, \end{cases} \quad (143)$$

with an appropriate (t-dependent) value of $\gamma_t \leq 1$. $\qquad\square$

Relying on the statement in Lemma 10 and using it in (142) we achieve

$$Q_3 \leq \frac{\lambda_n\{\Omega_n\}n^2 2}{\lambda_{n-1}\{\Omega_{n-1}\}} \cdot$$

$$\cdot \frac{\int\limits_{\mu_n t}^{t} G(h)^{m-n}[1 + \mu_{n-1}\frac{T}{T}]\frac{T^{n-3}}{t^{n-2}}\, dh \cdot \hat{\mathcal{F}}^{(t)}(t) + \int\limits_{t_+}^{1}\int\limits_{\mu_n t}^{t} G(h)^{m-n}[1 + \mu_{n-1}\frac{R}{T}]\frac{R^{n-3}}{r^{n-2}}\, dh\, d\hat{\mathcal{F}}^{(t)}(r)}{\int\limits_{\mu_n t}^{t} G(h)^{m-n}\frac{T}{t}[1 + \frac{1}{n-1}\frac{T^2}{T^2}]\frac{T^{n-3}}{t^{n-2}}\, dh \cdot \hat{\mathcal{F}}^{(t)}(t) + \int\limits_{t_+}^{1}\int\limits_{\mu_n t}^{t} G(h)^{m-n}\frac{T}{r}[1 + \frac{1}{n-1}\frac{R^2}{T^2}]\frac{R^{n-3}}{r^{n-2}}\, dh\, d\hat{\mathcal{F}}^{(t)}(r)}$$

$$= \quad \frac{\lambda_n\{\Omega_n\}n^2 2}{\lambda_{n-1}\{\Omega_{n-1}\}} \cdot \frac{\int\limits_t^1 \int\limits_{\mu_n t}^t G(h)^{m-n}\frac{R^{n-3}}{r^{n-2}}[1 + \mu_{n-1}\frac{R}{T}]\,dh\,d\hat{\mathcal{F}}^{(t)}(r)}{\int\limits_t^1 \int\limits_{\mu_n t}^t G(h)^{m-n}\frac{R^{n-3}}{r^{n-2}}\frac{1}{rT}[T^2 + \frac{1}{n-1}R^2]\,dh\,d\hat{\mathcal{F}}^{(t)}(r)} =: \hat{Q}_3. \tag{144}$$

From now on we can ignore the radii with $r < t$.

Remark 8

Since $\hat{\mathcal{F}}^{(t)}$ is based on $\mathcal{F}$ and two modifications, which move distribution mass (monotonously) towards higher radii, we know that

$$\hat{\mathcal{F}}^{(t)}(h) \leq \mathcal{F}(h) \; \forall \, h \in [0,1]. \tag{145}$$

Remark 9

The same holds for the two marginal distribution functions

$$G_{\mathcal{F}}(h) \quad = \quad \frac{1}{2} + \frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \int\limits_0^1 \int\limits_0^{\frac{h}{r}} (1 - \sigma^2)^{\frac{n-3}{2}}\,d\sigma d\mathcal{F}(r), \tag{146}$$

$$G_{\hat{\mathcal{F}}^{(t)}}(h) \quad = \quad \frac{1}{2} + \frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \int\limits_0^1 \int\limits_0^{\frac{h}{r}} (1 - \sigma^2)^{\frac{n-3}{2}}\,d\sigma d\hat{\mathcal{F}}^{(t)}(r). \tag{147}$$

From the remark above it is clear that

$$G_{\hat{\mathcal{F}}^{(t)}}(h) \leq G_{\mathcal{F}}(h) = G(h) \quad \forall \, h \in [0,1]. \tag{148}$$

Here we <u>interrupt</u> the proof for a short comment. This is the point where we lost precision in Borgwardt (1987) caused by a crude, but seemingly inevitable overestimation.
Consider the quotient

$$\frac{(T + \mu_{n-1}R)}{[T^2 + \frac{1}{n-1}R^2]}, \tag{149}$$

which is the essential part of the quotient of integrands in (144). Its numerator and denominator are embedded in integrals over r and h . The variation of both variables makes handling the quotient very troublesome. One attempt to overcome that difficulty lies in considering only the dominating term in numerator and denominator respectively. That means that we focus on

$$\frac{\max\{T, \mu_{n-1}R\}}{\max\{T^2, \frac{1}{n-1}R^2\}}.$$

But now the question arises, for which (r,h)-configurations which term dominates the other. Even for the relevant case $r > t$ and different r's, the "turnover" happens at different values of h. This would force us even to introduce flexible bounds on the integration area, what would make a sharp estimation overcomplicated.

In Borgwardt (1982b), Borgwardt (1987), Borgwardt (1994a) and Borgwardt (1994b) these problems had been avoided by use of the simplifying overestimation

$$\frac{(T + \mu_{n-1} R)}{[T^2 + \frac{1}{n-1} R^2]} \le \frac{\max\{T, R\}}{\frac{1}{n-1} \max\{T^2, R^2\}}, \tag{150}$$

because here for $r > t$ the R-term consistently dominates for every h. But the price we had to pay for that simplification was the big loss of a factor μ_{n-1}, in particular for $R \gg T$.

Now we continue the new proof. This time, we avoid such an overestimation of (144) as in (150) and use refined methods. In the following we shall deal with $\hat{Q}_3$ only. Again, we make use of Remark 7 resp. (129):

$$[T + \mu_{n-1} R]^2 \le 2[T^2 + \frac{1}{n-1} R^2].$$

We conclude

$$\hat{Q}_3 \le \frac{\lambda_n\{\Omega_n\} n^2 4}{\lambda_{n-1}\{\Omega_{n-1}\}} \cdot \frac{\int_t^1 \int_{\mu_n t}^t G(h)^{m-n} R^{n-3} r^{-n+2} \frac{R}{T}[\frac{T}{R} + \mu_{n-1}] \, dh \, d\hat{\mathcal{F}}^{(t)}(r)}{\int_t^1 \int_{\mu_n t}^t G(h)^{m-n} R^{n-3} r^{-n+2} \frac{1}{Tr}[T + \mu_{n-1} R]^2 \, dh \, d\hat{\mathcal{F}}^{(t)}(r)} =$$

$$= \frac{\lambda_n\{\Omega_n\} n^2 4}{\lambda_{n-1}\{\Omega_{n-1}\}} \cdot \frac{\int_t^1 \int_{\mu_n t}^t G(h)^{m-n} R^{n-3} r^{-n+2} \frac{R}{T}[\frac{T}{R} + \mu_{n-1}] \, dh \, d\hat{\mathcal{F}}^{(t)}(r)}{\int_t^1 \int_{\mu_n t}^t G(h)^{m-n} R^{n-3} r^{-n+2} \frac{R^2}{Tr}[\frac{T}{R} + \mu_{n-1}]^2 \, dh \, d\hat{\mathcal{F}}^{(t)}(r)} \le$$

$$\le \frac{\lambda_n\{\Omega_n\} n^2 4}{\lambda_{n-1}\{\Omega_{n-1}\} \mu_{n-1}} \cdot \frac{\int_t^1 \int_{\mu_n t}^t G(h)^{m-n} R^{n-3} r^{-n+2} \frac{R}{T}[\frac{T}{R} + \mu_{n-1}] \, dh \, d\hat{\mathcal{F}}^{(t)}(r)}{\int_t^1 \int_{\mu_n t}^t G(h)^{m-n} R^{n-3} r^{-n+2} \frac{R^2}{Tr}[\frac{T}{R} + \mu_{n-1}] \, dh \, d\hat{\mathcal{F}}^{(t)}(r)} =: \overline{Q}_3. \tag{151}$$

It will pay to interpret the numerator and the denominator of (151) as integrals over functions of h. Therefore, we permute the order of integrations and apply (93) to simplify the first factor

$$\overline{Q}_3 = 4 \cdot \pi \cdot (n-1) \cdot n \cdot \frac{\int_{\mu_n t}^t G(h)^{m-n} \int_t^1 R^{n-3} r^{-n+2} \frac{R}{T}[\frac{T}{R} + \mu_{n-1}] \, d\hat{\mathcal{F}}^{(t)}(r) \, dh}{\int_{\mu_n t}^t G(h)^{m-n} \int_t^1 R^{n-3} r^{-n+2} \frac{R^2}{Tr}[\frac{T}{R} + \mu_{n-1}] \, d\hat{\mathcal{F}}^{(t)}(r) \, dh}. \tag{152}$$

We analyze that integral-quotient more precisely for a single value of h. Then T is a constant.

$$\frac{\int_t^1 R^{n-3} r^{-n+2} \frac{R}{T}[\frac{T}{R} + \mu_{n-1}] \, d\hat{\mathcal{F}}^{(t)}(r)}{\int_t^1 R^{n-3} r^{-n+2} \frac{R^2}{rT}[\frac{T}{R} + \mu_{n-1}] \, d\hat{\mathcal{F}}^{(t)}(r)} = \frac{\int_t^1 R^{n-2} r^{-n+2}[\frac{T}{R} + \mu_{n-1}] \, d\hat{\mathcal{F}}^{(t)}(r)}{\int_t^1 R^{n-1} r^{-n+1}[\frac{T}{R} + \mu_{n-1}] \, d\hat{\mathcal{F}}^{(t)}(r)}.$$

We keep in mind that

$$\int\limits_t^1 d\hat{\mathcal{F}}^{(t)}(r) = 1, \ \ T \le R, \ \ 0 \le \frac{T}{R} \le 1, \ \ \mu_{n-1} \le [\frac{T}{R} + \mu_{n-1}] \le 1 + \mu_{n-1},$$

and apply Jensen's inequality

$$\left[\frac{\int\limits_t^1 \left(\frac{R}{r}\right)^{n-2} [\frac{T}{R} + \mu_{n-1}] d\hat{\mathcal{F}}^{(t)}(r)}{\int\limits_t^1 [\frac{T}{R} + \mu_{n-1}] d\hat{\mathcal{F}}^{(t)}(r)} \right] \le \left[\frac{\int\limits_t^1 \left(\frac{R}{r}\right)^{n-1} [\frac{T}{R} + \mu_{n-1}] d\hat{\mathcal{F}}^{(t)}(r)}{\int\limits_t^1 [\frac{T}{R} + \mu_{n-1}] d\hat{\mathcal{F}}^{(t)}(r)} \right]^{\frac{n-2}{n-1}}. \tag{153}$$

Hence it is clear that

$$\left[\frac{\int\limits_t^1 \left(\frac{R}{r}\right)^{n-2} [\frac{T}{R} + \mu_{n-1}] d\hat{\mathcal{F}}^{(t)}(r)}{\int\limits_t^1 \left(\frac{R}{r}\right)^{n-1} [\frac{T}{R} + \mu_{n-1}] d\hat{\mathcal{F}}^{(t)}(r)} \right] \le \left[\frac{\int\limits_t^1 [\frac{T}{R} + \mu_{n-1}] d\hat{\mathcal{F}}^{(t)}(r)}{\int\limits_t^1 \left(\frac{R}{r}\right)^{n-1} [\frac{T}{R} + \mu_{n-1}] d\hat{\mathcal{F}}^{(t)}(r)} \right]^{\frac{1}{n-1}} \le$$

$$\le \left[\frac{\int\limits_t^1 [1 + \mu_{n-1}] d\hat{\mathcal{F}}^{(t)}(r)}{\int\limits_t^1 \left(\frac{R}{r}\right)^{n-1} [0 + \mu_{n-1}] d\hat{\mathcal{F}}^{(t)}(r)} \right]^{\frac{1}{n-1}} \le \left(\frac{1 + \mu_{n-1}}{\mu_{n-1}} \right)^{\frac{1}{n-1}} \left[\int\limits_t^1 \left(\frac{R}{r}\right)^{n-1} d\hat{\mathcal{F}}^{(t)}(r) \right]^{-\frac{1}{n-1}}$$

$$\le \left(\frac{1 + \mu_{n-1}}{\mu_{n-1}} \right)^{\frac{1}{n-1}} 2^{-\frac{1}{n-1}} [1 - G_{\hat{\mathcal{F}}^{(t)}}(h)]^{-\frac{1}{n-1}}. \tag{154}$$

This results from the definition of G and the relation (compare Borgwardt (1987), page 170-171):

$$\int\limits_h^1 \frac{\sqrt{r^2 - h^2}^{n-1}}{r^{n-1}} d\hat{\mathcal{F}}^{(t)}(r) = \int\limits_h^1 R^{n-1} r^{-n+1} d\hat{\mathcal{F}}^{(t)}(r) = (n-1) \int\limits_h^1 \int\limits_{\frac{h}{r}}^1 (1 - \sigma^2)^{\frac{n-3}{2}} \sigma d\sigma d\hat{\mathcal{F}}^{(t)}(r) \ge$$

$$\ge 2 \cdot \frac{\lambda_{n-2}\{\omega_{n-1}\}}{\lambda_{n-1}\{\omega_n\}} \int\limits_h^1 \int\limits_{\frac{h}{r}}^1 (1 - \sigma^2)^{\frac{n-3}{2}} d\sigma d\hat{\mathcal{F}}^{(t)}(r) = 2[1 - G_{\hat{\mathcal{F}}^{(t)}}(h)], \tag{155}$$

which also guarantees that

$$\left[\frac{\int\limits_t^1 \left(\frac{R}{r}\right)^{n-2} d\hat{\mathcal{F}}^{(t)}(r)}{\int\limits_t^1 \left(\frac{R}{r}\right)^{n-1} d\hat{\mathcal{F}}^{(t)}(r)} \right] \le 2^{-\frac{1}{n-1}} [1 - G_{\hat{\mathcal{F}}^{(t)}}(h)]^{-\frac{1}{n-1}}. \tag{156}$$

As we already know from Remark 9 and (148), the original distribution induced by $\mathcal{F}$ and the auxiliary distribution induced by $\hat{\mathcal{F}}^{(t)}$ make sure that

$$G_{\hat{\mathcal{F}}^{(t)}}(h) \le G_{\mathcal{F}}(h) = G(h) \ \ \forall h \in [0, 1]. \tag{157}$$

Particularly, for a value $\tilde{h}$ such that

$$G_{\mathcal{F}}(\tilde{h}) = 1 - \frac{1}{m-n+1}$$

it is sure that

$$G_{\hat{\mathcal{F}}^{(t)}}(\tilde{h}) \leq 1 - \frac{1}{m-n+1} = G_{\mathcal{F}}(\tilde{h}). \tag{158}$$

So we have $\forall h \in [\mu_n t, t]$ such that $h < \tilde{h}: \; G_{\hat{\mathcal{F}}^{(t)}}(h) \leq G_{\mathcal{F}}(h) < 1 - \frac{1}{m-n+1}. \tag{159}$

In the following we shall make some case-studies on $\overline{Q}_3$, compare (152) and (151), based on the respective position of $\tilde{h}$. In (brackets) the corresponding case or the configuration will be described. For this purpose we introduce the notation:

$$\zeta := \max\{\mu_n t, \tilde{h}\}. \tag{160}$$

Suppose, that $\zeta > \mu_n t$, i.e. $\tilde{h} > \mu_n t$. If we then restrict our integral-quotient $\overline{Q}_3$ from (152) to the region where $\mu_n t \leq h \leq \min\{\tilde{h}, t\}$, and if we denote the corresponding quotient by $\overline{Q}_3(\mu_n t \leq h \leq \min\{\tilde{h}, t\})$, then the evaluation based on (154) and (158) delivers

$$\overline{Q}_3(\mu_n t \leq h \leq \min\{\tilde{h}, t\}) \;\leq\; \pi(n-1)n4\left[1 + \frac{1}{\mu_{n-1}}\right]^{\frac{1}{n-1}} 2^{-\frac{1}{n-1}}(1 - G_{\hat{\mathcal{F}}^{(t)}}(\tilde{h}))^{-\frac{1}{n-1}}$$

$$\leq\; \pi(n-1)n4\left[1 + \frac{1}{\mu_{n-1}}\right]^{\frac{1}{n-1}} 2^{-\frac{1}{n-1}}(m-n+1)^{\frac{1}{n-1}}. \tag{161}$$

If $\tilde{h} \leq t$, then still $[\max\{\mu_n t, \tilde{h}\}, t] = [\zeta, t]$ remains to be analyzed.
In the area under consideration we have

$$G(h) \geq 1 - \frac{1}{m-n+1} \quad \text{and hence} \quad [G(h)]^{m-n} \geq e^{-1}. \tag{162}$$

Consider the region $\mu_n t \leq \zeta \leq h \leq t$. Let us denote the corresponding quotient of (152) by $\overline{Q}_3(\mu_n t \leq \zeta \leq h \leq t)$. We make use of (162) and estimate $G(h) \leq 1$ in the numerator and $G(h)^{m-n} \geq e^{-1}$ in the denominator. So we achieve

$$\overline{Q}_3(\mu_n t \leq \zeta \leq h \leq t) \leq \pi \cdot (n-1) \cdot n \cdot 4 \cdot e \cdot \frac{\int_{\zeta}^{t}\int_{t}^{1} \frac{R^{n-2}}{r^{n-2}}\frac{1}{T}[\frac{T}{R} + \mu_{n-1}] \, d\hat{\mathcal{F}}^{(t)}(r)\, dh}{\int_{\zeta}^{t}\int_{t}^{1} \frac{R^{n-1}}{r^{n-1}}\frac{1}{T}[\frac{T}{R} + \mu_{n-1}] \, d\hat{\mathcal{F}}^{(t)}(r)\, dh}. \tag{163}$$

For this most important and most critical area we use the following <u>essential</u> estimation.

Lemma 11
If $r \geq t$, then we know that

$$\frac{\int_{\zeta}^{t}\int_{t}^{1} \frac{1}{T}\frac{R^{n-1}}{r^{n-1}}[\frac{T}{R} + \mu_{n-1}] \, d\hat{\mathcal{F}}^{(t)}(r)\, dh}{\int_{\zeta}^{t}\int_{t}^{1} \frac{1}{T}\frac{R^{n-2}}{r^{n-2}}[\frac{T}{R} + \mu_{n-1}] \, d\hat{\mathcal{F}}^{(t)}(r)\, dh} \geq \left[\frac{\mu_{n-1}}{n(1+\mu_{n-1})}\right]^{\frac{1}{n-1}}\left[\int_{t}^{1} \frac{\sqrt{r^2 - \zeta^2}^{n-1}}{r^{n-1}} \, d\hat{\mathcal{F}}^{(t)}(r)\right]^{\frac{1}{n-1}} \tag{164}$$

$$\frac{\int\limits_{\zeta}^{t}\int\limits_{t}^{1}\frac{1}{\sqrt{t^2-h^2}}\frac{\sqrt{r^2-h^2}^{n-1}}{r^{n-1}}\left[\frac{T}{R}+\mu_{n-1}\right]d\hat{\mathcal{F}}^{(t)}(r)\,dh}{\int\limits_{\zeta}^{t}\int\limits_{t}^{1}\frac{1}{\sqrt{t^2-h^2}}\frac{\sqrt{r^2-h^2}^{n-2}}{r^{n-2}}\left[\frac{T}{R}+\mu_{n-1}\right]d\hat{\mathcal{F}}^{(t)}(r)\,dh} \tag{165}$$

$$\geq\left[\frac{1}{1+\mu_{n-1}}\right]^{\frac{1}{n-1}}\frac{\int\limits_{\zeta}^{t}\int\limits_{t}^{1}\frac{1}{\sqrt{t^2-h^2}}\frac{\sqrt{r^2-h^2}^{n-1}}{r^{n-1}}\left[\frac{T}{R}+\mu_{n-1}\right]^{\frac{n-1}{n-1}}d\hat{\mathcal{F}}^{(t)}(r)\,dh}{\int\limits_{\zeta}^{t}\int\limits_{t}^{1}\frac{1}{\sqrt{t^2-h^2}}\frac{\sqrt{r^2-h^2}^{n-2}}{r^{n-2}}\left[\frac{T}{R}+\mu_{n-1}\right]^{\frac{n-2}{n-1}}d\hat{\mathcal{F}}^{(t)}(r)\,dh} \tag{166}$$

$$\geq\left[\frac{1}{1+\mu_{n-1}}\right]^{\frac{1}{n-1}}\left[\frac{\int\limits_{\zeta}^{t}\int\limits_{t}^{1}\frac{1}{\sqrt{t^2-h^2}}\frac{\sqrt{r^2-h^2}^{n-1}}{r^{n-1}}\left[\frac{T}{R}+\mu_{n-1}\right]^{\frac{n-1}{n-1}}d\hat{\mathcal{F}}^{(t)}(r)\,dh}{\int\limits_{\zeta}^{t}\int\limits_{t}^{1}\frac{1}{\sqrt{t^2-h^2}}\,d\hat{\mathcal{F}}^{(t)}(r)\,dh}\right]^{\frac{1}{n-1}} \tag{167}$$

$$\geq\left[\frac{\mu_{n-1}}{1+\mu_{n-1}}\right]^{\frac{1}{n-1}}\left[\frac{\int\limits_{\zeta}^{t}\int\limits_{t}^{1}\frac{1}{\sqrt{t^2-h^2}}\frac{\sqrt{r^2-h^2}^{n-1}}{r^{n-1}}\,d\hat{\mathcal{F}}^{(t)}(r)\,dh}{\int\limits_{\zeta}^{t}\int\limits_{t}^{1}\frac{1}{\sqrt{t^2-h^2}}\,d\hat{\mathcal{F}}^{(t)}(r)\,dh}\right]^{\frac{1}{n-1}} \tag{168}$$

$$=\left[\frac{\mu_{n-1}}{1+\mu_{n-1}}\right]^{\frac{1}{n-1}}\left[\frac{\int\limits_{\zeta}^{t}\int\limits_{t}^{1}\frac{1}{\sqrt{t^2-h^2}}\frac{\sqrt{r^2-h^2}^{n-1}\sqrt{t^2-h^2}^{n-1}}{r^{n-1}\sqrt{t^2-h^2}^{n-1}}\,d\hat{\mathcal{F}}^{(t)}(r)\,dh}{\int\limits_{\zeta}^{t}\int\limits_{t}^{1}\frac{1}{\sqrt{t^2-h^2}}\,d\hat{\mathcal{F}}^{(t)}(r)\,dh}\right]^{\frac{1}{n-1}} \tag{169}$$

$$\geq\left[\frac{\mu_{n-1}}{1+\mu_{n-1}}\right]^{\frac{1}{n-1}}\left[\frac{\int\limits_{\zeta}^{t}\int\limits_{t}^{1}\frac{1}{\sqrt{t^2-h^2}}\frac{\sqrt{r^2-\zeta^2}^{n-1}\sqrt{t^2-h^2}^{n-1}}{r^{n-1}\sqrt{t^2-\zeta^2}^{n-1}}\,d\hat{\mathcal{F}}^{(t)}(r)\,dh}{\int\limits_{\zeta}^{t}\int\limits_{t}^{1}\frac{1}{\sqrt{t^2-h^2}}\,d\hat{\mathcal{F}}^{(t)}(r)\,dh}\right]^{\frac{1}{n-1}} \tag{170}$$

$$=\left[\frac{\mu_{n-1}}{1+\mu_{n-1}}\right]^{\frac{1}{n-1}}\left[\frac{\int\limits_{t}^{1}\frac{\sqrt{r^2-\zeta^2}^{n-1}}{r^{n-1}}\,d\hat{\mathcal{F}}^{(t)}(r)}{\int\limits_{t}^{1}d\hat{\mathcal{F}}^{(t)}(r)}\cdot\frac{1}{\sqrt{t^2-\zeta^2}^{n-1}}\cdot\frac{\int\limits_{\zeta}^{t}\frac{1}{\sqrt{t^2-h^2}}\sqrt{t^2-h^2}^{n-1}\,dh}{\int\limits_{\zeta}^{t}\frac{1}{\sqrt{t^2-h^2}}\,dh}\right]^{\frac{1}{n-1}} \tag{171}$$

$$\geq\left[\frac{\mu_{n-1}}{1+\mu_{n-1}}\right]^{\frac{1}{n-1}}\left[\frac{\int\limits_{t}^{1}\frac{\sqrt{r^2-\zeta^2}^{n-1}}{r^{n-1}}\,d\hat{\mathcal{F}}^{(t)}(r)}{\int\limits_{t}^{1}d\hat{\mathcal{F}}^{(t)}(r)}\cdot\frac{1}{\sqrt{t^2-\zeta^2}^{n-1}}\cdot\frac{\int\limits_{\zeta}^{t}h\sqrt{t^2-h^2}^{n-2}\,dh}{\int\limits_{\zeta}^{t}\frac{h}{\sqrt{t^2-h^2}}\,dh}\right]^{\frac{1}{n-1}} \tag{172}$$

$$=\left[\frac{\mu_{n-1}}{1+\mu_{n-1}}\right]^{\frac{1}{n-1}}\left[\frac{\int\limits_{t}^{1}\frac{\sqrt{r^2-\zeta^2}^{n-1}}{r^{n-1}}\,d\hat{\mathcal{F}}^{(t)}(r)}{\int\limits_{t}^{1}d\hat{\mathcal{F}}^{(t)}(r)}\cdot\frac{1}{\sqrt{t^2-\zeta^2}^{n-1}}\cdot\frac{\frac{1}{n}\sqrt{t^2-\zeta^2}^{n}}{\sqrt{t^2-\zeta^2}}\right]^{\frac{1}{n-1}} \tag{173}$$

48

$$
= \left[\frac{\mu_{n-1}}{n(1+\mu_{n-1})}\right]^{\frac{1}{n-1}} \left[\frac{\int_t^1 \frac{\sqrt{r^2-\zeta^2}^{\,n-1}}{r^{n-1}} d\hat{\mathcal{F}}^{(t)}(r)}{\int_t^1 d\hat{\mathcal{F}}^{(t)}(r)}\right]^{\frac{1}{n-1}} \tag{174}
$$

$$
= \left[\frac{\mu_{n-1}}{n(1+\mu_{n-1})}\right]^{\frac{1}{n-1}} \left[\int_t^1 \frac{\sqrt{r^2-\zeta^2}^{\,n-1}}{r^{n-1}} d\hat{\mathcal{F}}^{(t)}(r)\right]^{\frac{1}{n-1}}. \tag{175}
$$

In (167) it is exploited, that the expectation value of an objective-variable increases, when the density is multiplied with a function, which is monotonously increasing with the objective variable. This amounts to an application of the Jensen-inequality. (170) relies on the fact that $\frac{\sqrt{r^2-h^2}}{\sqrt{t^2-h^2}} \geq \frac{\sqrt{r^2-\zeta^2}}{\sqrt{t^2-\zeta^2}}$ for $r > t > h > \zeta$. In (170) we observe that the integrands consist of factors which either depend on h only, or on r only, or are independent of both. So we can separate the integrations as in (171). Multiplication of the density with h in (171) in the rightmost quotient of integrals decreases this quotient, since its objective is a decreasing function of h. In the very last equation we apply $\int_t^1 d\hat{\mathcal{F}}^{(t)}(r) = 1$. $\square$

Exploitation of Lemma 11 for proving Proposition 6

Insertion into (163) and application of (155) and (157) yields

$$
\overline{Q}_3(\mu_n t \leq \zeta \leq h \leq t) \leq \pi(n-1)n4e \left[\frac{n(1+\mu_{n-1})}{\mu_{n-1}}\right]^{\frac{1}{n-1}} \left[\int_t^1 \frac{\sqrt{r^2-\zeta^2}^{\,n-1}}{r^{n-1}} d\hat{\mathcal{F}}^{(t)}(r)\right]^{-\frac{1}{n-1}} \tag{176}
$$

$$
\leq \pi(n-1)n4e \left[\frac{n(1+\mu_{n-1})}{2\mu_{n-1}}\right]^{\frac{1}{n-1}} [1-G_{\hat{\mathcal{F}}^{(t)}}(\zeta)]^{-\frac{1}{n-1}} \tag{177}
$$

$$
\leq \pi(n-1)n4e \left[\frac{n(1+\mu_{n-1})}{2\mu_{n-1}}\right]^{\frac{1}{n-1}} [1-G_{\mathcal{F}}(\zeta)]^{-\frac{1}{n-1}}. \tag{178}
$$

This means for the case that $\zeta > \mu_n t$

$$
\overline{Q}_3(\mu_n t < \zeta \leq h \leq t) \leq \pi \cdot (n-1) \cdot n \cdot 4 \cdot e \cdot \left[\frac{n(1+\mu_{n-1})}{2\mu_{n-1}}\right]^{\frac{1}{n-1}} \left\{(m-n+1)^{\frac{1}{n-1}}\right\}, \tag{179}
$$

in accordance to (157), (158), (159) and (160).

Finally, consider the potential case that $\zeta = \mu_n t$, i.e. $\tilde{h} \leq \mu_n t$. Denote the corresponding quotient by $\overline{Q}_3(\mu_n t = \zeta \leq h \leq t)$. From (176) we learn that

$$
\overline{Q}_3(\mu_n t = \zeta \leq h \leq t) \leq \pi(n-1)n4e \left[\frac{n(1+\mu_{n-1})}{\mu_{n-1}}\right]^{\frac{1}{n-1}} \left[\int_t^1 \frac{\sqrt{r^2-\zeta^2}^{\,n-1}}{r^{n-1}} d\hat{\mathcal{F}}^{(t)}(r)\right]^{-\frac{1}{n-1}} \tag{180}
$$

For any single radius r in our special configuration $\tilde{h} \leq \mu_n t$ we have

$$
\frac{\sqrt{r^2-(\zeta)^2}}{r} = \frac{\sqrt{r^2-(\mu_n t)^2}}{r} \geq \frac{\sqrt{r^2-(\mu_n r)^2}}{r} \geq \sqrt{1-(\mu_n)^2}. \tag{181}
$$

Insertion in (180) delivers

$$\overline{Q}_3(\mu_n t = \zeta \le h \le t) \le \pi \cdot (n-1) \cdot n \cdot 4 \cdot e \cdot \frac{1}{\sqrt{1-(\mu_n)^2}} \left[\frac{n(1+\mu_{n-1})}{\mu_{n-1}}\right]^{\frac{1}{n-1}} . \tag{182}$$

Having analyzed all the possible cases of $\overline{Q}_3$ in B_3, namely (161), (179), (182), we observe that

$$Q_3 \le \pi(n-1)n4e \cdot \left[\frac{n(1+\mu_{n-1})}{\mu_{n-1}}\right]^{\frac{1}{n-1}} \max\left\{\left(\frac{m-n+1}{2}\right)^{\frac{1}{n-1}}, \frac{1}{\sqrt{1-(\mu_n)^2}}\right\} . \tag{183}$$

This finishes the Proof for Proposition 6. □

Now we can formulate our main result, recalling from (89) that $\mu_n \approx \sqrt{\frac{2}{(n-1)\pi}}$:

Theorem 3

$$E_{m,n}(S) \le \pi \cdot (n-1) \cdot n \cdot 4 \cdot \tag{184}$$
$$\cdot \max\left\{\frac{2e^{\frac{1}{\pi}}}{\sqrt{1-\mu_n{}^2}}, e\left[\frac{n(1+\mu_{n-1})}{2\mu_{n-1}}\{m-n+1\}\right]^{\frac{1}{n-1}}, e\left[\frac{n(1+\mu_{n-1})}{\mu_{n-1}}\right]^{\frac{1}{n-1}}\frac{1}{\sqrt{1-(\mu_n)^2}}\right\} .$$

Proof
Comparison of (140) with the bounds derived earlier in (130) resp. (139)

$$Q_1 \le \frac{\lambda_n\{\Omega_n\} \cdot n^2 \cdot 8 \cdot e^{\frac{1}{\pi}}}{\lambda_{n-1}\{\Omega_{n-1}\}\mu_{n-1}\sqrt{1-\mu_n{}^2}} = \pi(n-1) \cdot n \cdot 8 \cdot e^{\frac{1}{\pi}}\frac{1}{\sqrt{1-\mu_n{}^2}},$$

$$Q_2 \le \frac{2 \cdot \lambda_n\{\Omega_n\} \cdot n^2}{\lambda_{n-1}\{\Omega_{n-1}\}\sqrt{1-\mu_n{}^2}},$$

$$Q_3 \le \pi(n-1)n4e \cdot \left[\frac{n(1+\mu_{n-1})}{\mu_{n-1}}\right]^{\frac{1}{n-1}} \max\left\{\left(\frac{m-n+1}{2}\right)^{\frac{1}{n-1}}, \frac{1}{\sqrt{1-(\mu_n)^2}}\right\}$$

leads to the result that the bound given in the Theorem is the worst. □

8 Final Result, Consequences and Conclusions

Finally, let us give a brief overwiew over the main results in the RSM-investigation. As before, C or $Const.$ denote constants, independent of m and of n.

In a concise form our main result reads:

Theorem 4

$$E_{m,n}(S) \leq m^{\frac{1}{n-1}} \cdot n^2 \cdot Const. \quad \text{for all pairs } (m,n) \text{ with } m \geq n. \tag{185}$$

One should compare this upper bound for general distributions with a lower bound for a special distribution which was first derived in Borgwardt (1987).

For uniform distribution on ω_n it is known that

$$m^{\frac{1}{n-1}} \cdot n^2 \cdot 2 \cdot \overline{C}_n \cdot [1 + \epsilon(m,n)] \geq E_{m,n}(S) \geq m^{\frac{1}{n-1}} \cdot n^2 \cdot 2 \cdot \underline{C}_n \cdot [1 - \epsilon(m,n)] \tag{186}$$

with $\epsilon(m,n) \to 0$ for $m \to \infty$, n fixed   and with $\overline{C}_n, \underline{C}_n \to 1$ for $n \to \infty$.

The asymptotic result (186) for a special distribution shows that our result in Theorem 4 comes close to the lower bound, and that it differs from any sharpest uniform upper bound, which is valid for all RSM-distributions, only by a moderate constant factor.

As explained in section 1, our main result helps to improve the bound for the average effort for the so-called dimension-by-dimension method from Borgwardt (1987), our complete solution method for LPs of type (1). The new bound on its average number of pivot steps is

$$E_{m,n}(s_t) \leq m^{\frac{1}{n-1}} \cdot n^3 \cdot Const. \tag{187}$$

These insights and average-case bounds can also be transferred to problems of the type

$$\text{maximize } v^T x \quad \text{s.t. } a_1^T x \leq 1, \ldots, a_m^T x \leq 1 \text{ and } x \geq 0. \tag{188}$$

Polynomial upper bounds on the average number of pivot steps have also been derived under the so-called sign invariance model in Adler, Karp and Shamir (1987), Adler and Megiddo (1985), Todd (1986). There the stochastic assumptions are extremely different and it makes no sense to start a comparison here. For RSM the study presented here finally gives a quantitatively precise information on the behaviour in single stages, in particular in the last stage (Phase II), as had been available for SIM from Haimovich (1983). The analysis of the expected number of pivot steps under RSM is still developing and in progress, as the following examples may demonstrate.

Höfner (1995) could show that the dimension-by-dimension method has – under distributions with bounded support Ω_n – a better asymptotic bound of

$$E_{m,n}(s_t) \leq m^{\frac{1}{n-1}} \cdot n^{\frac{5}{2}} \cdot Const. \quad \text{for } m \to \infty, \text{ n fixed.}$$

The reason for that improvement lies in the observation, that in the asymptotic case, where $\text{conv}(0, a_1, \ldots, a_m)$ approximates the unit ball from inside very closely, there is not much left to do in higher stages ($k \gg 1$). This holds, because stages $1, \ldots, k-1$ have provided us with an entrance vertex for stage k, which is very close to the optimal vertex in stage k. This confirmation could – so far – be given only for the asymptotic case.

And, we can combine the present result with Höfner (1995) in order to construct a complete solution-algorithm, which performs two stages of the shadow vertex algorithm only, but which is explicitly designed for the asymptotic configuration, so that its success-probability is only there extremely close to 1. This algorithm would reflect the common belief in the equal complexity of the two Phases, since it does not only the work of Phase II, but also that of Phase I for problems of type (1) or (188) in an average number of steps less than

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

Küfer (1995) proved that (for a certain subclass of RSM-distributions with bounded support) – and again for the asymptotic case – the quotient of variance and squared expected value of $s_{u,v}$ tends to behave as $\frac{1}{n}$, and the corresponding quotient for S behaves (asymptotically seen as a function of m for fixed n) better than $m^{\frac{-1}{n-1}}$.

Finally, it should be mentioned that Huhn (1997) derived lower bounds on $E_{m,n}(s)$ for a family of RSM-distributions (the same subclass as mentioned above). For uniform distribution on ω_n, she calculated a constant $C > 0$, such that **any** Phase II-Simplex-Variant requires more than

$$m^{\frac{1}{n-1}} \cdot C \cdot n^0 \qquad \text{steps on the average.}$$

Although this result holds for every pair (m, n), it is significant and meaningful only when $m \gg n$, because the constant is small and m has to become large until the bound becomes nontautological. This insight exploits the fact that edges of X (stochastically) are not long for arbitrary dimensions and tend to be very short for $m \gg n$ in RSM. This shows that under RSM, no variant can avoid the factor $m^{\frac{1}{n-1}}$, which had been observed with the shadow vertex algorithm.

For these three types of results, it is highly desirable to study the question, whether they hold for moderate dimensions (m, n), too. Here the situation is similar as it was with the question on $E_{m,n}(S)$, before clarification in this paper: The asymptotic behavior is proven to be quite good, but still we do not know whether this quality is a general effect, or whether it is only a consequence of the asymptotic configuration.

So far not very satisfactory is the state of the art with problems (of the general type) with arbitrary capacities b^i (not necessarily positive)

$$\text{maximize } v^T x \text{ s.t. } a_1^T x \leq b^1, \ldots, a_m^T x \leq b^m \quad \text{where} \quad b \in \mathbb{R}^m. \tag{189}$$

We developed a complete algorithm, which augments one additional stage to the dimension-by-dimension method, in order to translate the general problem into a problem of our type (1). Then it solves the auxiliary problem and implicitly the original problem (189). Good upper bounds for the average number of pivot steps for the extended method like

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

could so far be derived only for two special distribution-cases, namely

- when all entries a_i^j and $1 - b^i$ follow a Gaussian distribution, or when
- the augmented row-vectors $(a_i^1, \ldots, a_i^m, 1 - b^i)$ follow an RSM-distribution.

So, there is still enough to do.

APPENDIX

This Appendix delivers some results and their rather technical proofs, which have already been mentioned in the main text, but not been proved there.

Lemma 3

Let c_n and z be fixed such that $\frac{c_n}{\|c_n\|}^T z \neq 0$. Consider the orthogonal hyperplane H to z.

Then with the exception of a λ_{n-2}-nullset, $\hat{\Upsilon}(\frac{c_n}{\|c_n\|}, z[T, h]; w)$ from (56) in Lemma 2 and $\Upsilon(\frac{c_n}{\|c_n\|}, z[T, h]; w)$ from Lemma 1, are identical functions on $\omega_n \cap H$. $\qquad\square$

Proof

Again, we set $e_n = \frac{c_n}{\|c_n\|}$ in this proof.

Fix a point $\tilde{w} \in \omega_n \cap H$ and a value $\epsilon > 0$. We recall from Lemma 2 that there is a value $\delta(\epsilon, z)$ such that any set $M(\tilde{w}) \subset \omega_n \cap H$ with positive measure and $M(\tilde{w}) \subset \{x | \angle(x, \tilde{w}) < \delta(\epsilon, z)\}$ satisfies

$$|\hat{\Upsilon}(e_n, z; \tilde{w}) - \frac{\widetilde{V}(e_n, M(\tilde{w}))}{\widetilde{W}(M(\tilde{w}))}| \leq \epsilon. \tag{190}$$

We may as well set $\delta_2 := \frac{1}{2} \cdot \delta$ and consider a set $N := N^{\delta_2} \subset \omega_n \cap H$ about $\tilde{w}$ with positive measure and $N^{\delta_2} \subset \{x | \angle(x, \tilde{w}) < \delta_2(\epsilon, z)\}$.

Now consider any point $x \in N^{\delta_2}$. Of course, for x the set N^{δ_2} is in a δ-neighbourhood of x. We learn from (190) that

$$|\hat{\Upsilon}(e_n, z; x) - \frac{\widetilde{V}(e_n, N)}{\widetilde{W}(N)}| \leq \epsilon \quad \forall x \in N. \tag{191}$$

From Lemma 1 we know the density function $\Upsilon(e_n, z; x)$, which is of course also defined on N. It is unique outside a nullset with respect to the measure λ_{n-2}. Let us study one such version and call this Υ_0.

The comparison between Υ_0 and $\hat{\Upsilon}$ induces a (disjoint) dissection of N into three subsets, namely

$$\begin{align}
N_\approx &:= N \cap \{x | \ |\Upsilon_0(e_n, z; x) - \hat{\Upsilon}(e_n, z; x)| \leq 2\epsilon\} \tag{192} \\
N_> &:= N \cap \{x | \ \Upsilon_0(e_n, z; x) - \hat{\Upsilon}(e_n, z; x) > 2\epsilon\} \tag{193} \\
N_< &:= N \cap \{x | \ \Upsilon_0(e_n, z; x) - \hat{\Upsilon}(e_n, z; x) < 2\epsilon\}. \tag{194}
\end{align}$$

Neither $N_>$ nor $N_<$ can contain a (measurable) subset of positive λ_{n-2}-measure. Otherwise, this subset would contradict the statement in Lemma 1. To explain this, assume that $N_>$ has positive measure. Since $N_>$ is a subset of N, it should satisfy according to (190)

$$|\frac{V(e_n, N_>)}{W(N_>)} - \hat{\Upsilon}(e_n, z; x)| \leq \epsilon \quad \forall x \in N_>. \tag{195}$$

But in this set we find according to (193) and (190) for every element $x \in N_>$

$$\Upsilon_0(e_n, z; x) > \{\hat{\Upsilon}(e_n, z; x) + \epsilon\} + \epsilon \geq \frac{\widetilde{V}(e_n, N_>)}{\widetilde{W}(N_>)} + \epsilon. \tag{196}$$

Since Υ_0 is a density for $\frac{V}{W}$ on $N_>$, this is the contradiction.

For $N_<$ we can argue analogously. That means that $\lambda_{n-2}\{N_> \cup N_<\} = 0$.

We generalize that argument to a finite dissection of the full region $\omega_n \cap H$ into sets of type N_i of maximal "angle-radius" δ_2. Since this angle-radius can be chosen independently of w (compare Lemma 2), a finite number of such neighbourhood-subsets suffices to cover the full region. This confirms that globally only on a nullset the two functions $\Upsilon_0(e_n, z[T, h]; w)$ and $\hat{\Upsilon}(e_n, z[T, h]; w)$ may differ by more than 2ϵ. Finishing that argument with arbitrary small ϵ yields the λ_{n-2}-almost-everywhere coincidence of both functions. $\qquad\square$

Lemma 5

The term

$$\frac{\int\limits_{\eta}^{1} \sqrt{1-h^2}^{n-3} dh}{(n-1)\int\limits_{|\eta|}^{1} \sqrt{1-h^2}^{n-3} h dh} = \frac{\int\limits_{\eta}^{1} \sqrt{1-h^2}^{n-3} dh}{(1-\eta^2)^{\frac{n-1}{2}}} \tag{197}$$

represents a monotonously decreasing, convex function of η in the interval $(-1, 1)$.

Proof

The first derivative of the function (197) is

$$\frac{-(1-\eta^2)^{\frac{n-3}{2}}(1-\eta^2)^{\frac{n-1}{2}} + \int\limits_{\eta}^{1} \sqrt{1-h^2}^{n-3} dh \cdot \eta(1-\eta^2)^{\frac{n-3}{2}}(n-1)}{(1-\eta^2)^{n-1}} =$$

$$= \frac{-(1-\eta^2)^{\frac{n-1}{2}} + \int\limits_{\eta}^{1} \sqrt{1-h^2}^{n-3} dh \cdot \eta(n-1)}{(1-\eta^2)^{\frac{n-1}{2}+1}}. \tag{198}$$

This shows that the value of the derivative (198) at $\eta = 0$ is just -1.

For $\eta < 0$ the convexity of (197) is obvious. Have a look at (198). The numerator of (198) is negative then. It has its own derivative of just $\int\limits_{\eta}^{1} \sqrt{1-h^2}^{n-3} dh \cdot (n-1) > 0$.

So, the numerator of (198) is negative and monotonously increasing.

The denominator of (198) is positive and increases while η grows.

The combination of both observations yields that (198) is negative and increases on the interval $\eta \in (-1, 0]$ up to the value -1.

So, (197) has a negative, but increasing derivative on $(-1, 0]$. The value of (197) at $\eta = 0$ is

$$\int\limits_{0}^{1} \sqrt{1-h^2}^{n-3} dh = \frac{\lambda_{n-1}\{\omega_n\}}{2 \cdot \lambda_{n-2}\{\omega_{n-1}\}} = \frac{1}{(n-1)\cdot\mu_n}.$$

In order to confirm this behaviour also for $\eta > 0$, we perform additional transformations, which are feasible only here (for $\eta > 0$).

$$
(198) \quad = \quad \frac{-\int_{\eta}^{1} \sqrt{1-h^2}^{n-3} h\, dh \cdot (n-1) + \int_{\eta}^{1} \sqrt{1-h^2}^{n-3}\, dh \cdot \eta(n-1)}{(1-\eta^2)\int_{\eta}^{1} \sqrt{1-h^2}^{n-3} h\, dh \cdot (n-1)} =
$$

$$
= \quad -\frac{1}{1+\eta} \cdot \frac{\int_{\eta}^{1} \sqrt{1-h^2}^{n-3} \cdot \frac{h-\eta}{1-\eta}\, dh}{\int_{\eta}^{1} \sqrt{1-h^2}^{n-3} \cdot h\, dh} =
$$

$$
= \quad -\frac{1}{1+\eta} \cdot \left[\frac{\int_{\eta}^{1} \sqrt{1-h^2}^{n-3} \cdot \frac{h-\eta}{1-\eta}\, dh}{\int_{\eta}^{1} \sqrt{1-h^2}^{n-3}\, dh}\right] \cdot \left[\frac{\int_{\eta}^{1} \sqrt{1-h^2}^{n-3} \cdot h\, dh}{\int_{\eta}^{1} \sqrt{1-h^2}^{n-3}\, dh}\right]^{-1}.
$$

This product is negative, so (197) is decreasing. It remains to show convexity. The first quotient decreases while η grows. The rest gives a relation between two expectation values. First observe the quotient in the last brackets. For growing η we observe a transformation of weights in favour of higher values of h. Hence this expectation value grows. For the figure in the first expectation value (first brackets) $\frac{h-\eta}{1-\eta}$ the effect is just the opposite. The larger η becomes, the steeper is the relative descent of the density-function, while the range of the objective is always $[0,1]$. So, by increasing η, the large values of the objective variable get less and less weight. Therefore, this expectation value decreases. So we have

- a negative sign

- a positive, decreasing first factor

- a positive, decreasing second factor

- a positive, decreasing third factor (the inverse of an increasing one).

Hence (197) is convex, because its derivative (198) is a negative, but increasing product. $\square$

Lemma 6

$$
\eta + (1-\eta)\frac{2}{(n+1)} \leq \frac{\int_{\eta}^{1} \sqrt{1-h^2}^{n-3} \cdot h\, dh}{\int_{\eta}^{1} \sqrt{1-h^2}^{n-3}\, dh} \leq \eta + (1-\eta)\frac{2 \cdot \lambda_{n-2}\{\omega_{n-1}\}}{(n-1)\lambda_{n-1}\{\omega_n\}} \quad \forall\, \eta > 0. \quad (199)
$$

And for the correction factor Φ it is known that for $\eta > 0$

$$
\Phi(\eta) \geq \frac{1}{\eta + (1-\eta)\mu_n} = \frac{1}{\eta + (1-\eta)\frac{2 \cdot \lambda_{n-2}\{\omega_{n-1}\}}{(n-1)\lambda_{n-1}\{\omega_n\}}}. \qquad \square \qquad (200)
$$

<u>Proof</u>

The second line is just a reformulation of the right inequality in the first line.
Let us first prove the left inequality in the first line. We use that

$$\frac{\int\limits_{\eta}^{1}\sqrt{1-h^2}^{\,n-3}\cdot h\,dh}{\int\limits_{\eta}^{1}\sqrt{1-h^2}^{\,n-3}dh} = \frac{\int\limits_{\eta}^{1}\sqrt{1-h}^{\,n-3}\sqrt{1+h}^{\,n-3}\cdot h\,dh}{\int\limits_{\eta}^{1}\sqrt{1-h}^{\,n-3}\sqrt{1+h}^{\,n-3}dh}.$$

This is an expectation value of h. Hence the factor $\sqrt{1+h}^{\,n-3}$ will support higher values of h. Dropping or reducing this factor will provide a lower expectation value for h.

$$\frac{\int\limits_{\eta}^{1}\sqrt{1-h}^{\,n-3}h\,dh}{\int\limits_{\eta}^{1}\sqrt{1-h}^{\,n-3}dh} = \eta + \frac{\int\limits_{\eta}^{1}\sqrt{1-h}^{\,n-3}(h-\eta)dh}{\int\limits_{\eta}^{1}\sqrt{1-h}^{\,n-3}dh} = \eta + \frac{(1-\eta)^{\frac{n-3}{2}+2}\int\limits_{0}^{1}\sqrt{1-u}^{\,n-3}u\,du}{(1-\eta)^{\frac{n-3}{2}+1}\int\limits_{0}^{1}\sqrt{1-u}^{\,n-3}du} =$$

$$= \eta + (1-\eta)\cdot\frac{\frac{2}{n-1}-\frac{2}{n+1}}{\frac{2}{n-1}} = \eta + (1-\eta)\cdot\frac{2}{n+1}. \tag{201}$$

Now we proceed to the right inequality

$$\frac{\int\limits_{\eta}^{1}\sqrt{1-h^2}^{\,n-3}\cdot h\,dh}{\int\limits_{\eta}^{1}\sqrt{1-h^2}^{\,n-3}dh} = \eta + \frac{\int\limits_{0}^{1}\sqrt{1-[\eta+(1-\eta)x]^2}^{\,n-3}\cdot x\,dx}{\int\limits_{0}^{1}\sqrt{1-[\eta+(1-\eta)x]^2}^{\,n-3}dx}\cdot(1-\eta) \leq^{*}$$

$$\leq \eta + \frac{\int\limits_{0}^{1}\sqrt{1-[x]^2}^{\,n-3}\cdot x\,dx}{\int\limits_{0}^{1}\sqrt{1-[x]^2}^{\,n-3}dx}\cdot(1-\eta) = \eta + (1-\eta)\cdot\frac{2\cdot\lambda_{n-2}\{\omega_{n-1}\}}{(n-1)\lambda_{n-1}\{\omega_{n}\}}. \tag{202}$$

* holds, because we know that for $0 \leq x_k \leq x_g < 1$:

$$\frac{1-x_k^2}{1-x_g^2} = \frac{(1-x_k)(1+x_k)}{(1-x_g)(1+x_g)} \leq \frac{1-x_k}{1-x_g}$$

$$\Longrightarrow \frac{1-x_k^2}{1-x_g^2} \leq \frac{(1-\eta)^2(1-x_k^2)+2\eta(1-\eta)(1-x_k)}{(1-\eta)^2(1-x_g^2)+2\eta(1-\eta)(1-x_g)} = \frac{1-[\eta+(1-\eta)x_k]^2}{1-[\eta+(1-\eta)x_g]^2}.$$

This means that * results from a transformation of weights in favour of higher x-values. The expectation value of x is the largest when $\eta = 0$. $\qquad\square$

Lemma 10

There is a radial distribution function $\hat{\mathcal{F}}^{(t)}$ with $\hat{\mathcal{F}}^{(t)}(r) = 0 \;\forall\; r < t$ and $\hat{\mathcal{F}}^{(t)}(1) = 1$, such that the quotient corresponding to (142) when based on the distribution function $\hat{\mathcal{F}}^{(t)}$ becomes worse (larger) than (142), which is based on $\mathcal{F}$. One instance of such a distribution function is of the form

$$\hat{\mathcal{F}}^{(t)}(r) = \begin{cases} 0 & \text{for } r < t \\ \gamma_t \cdot \mathcal{F}(t) & \text{for } r = t \\ \gamma_t \cdot \mathcal{F}(t) + [1 - \gamma_t \cdot \mathcal{F}(t)] \cdot \frac{\mathcal{F}(r) - \mathcal{F}(t)}{1 - \mathcal{F}(t)} & \text{for } r > t, \end{cases} \tag{203}$$

with an appropriate (t-dependent) value of γ_t. $\qquad\square$

<u>Proof</u>
Consider the following integral-quotient from (142)

$$\frac{\int\limits_{\mu_n t}^{t} \int\limits_{\mu_n t}^{r} G(h)^{m-n} \frac{R^{n-3}}{r^{n-2}} \, dh \, d\mathcal{F}(r) + \int\limits_{t_+}^{1} \int\limits_{\mu_n t}^{t} G(h)^{m-n} [1 + \mu_{n-1} \frac{R}{T}] \frac{R^{n-3}}{r^{n-2}} \, dh \, d\mathcal{F}(r)}{\int\limits_{\mu_n t}^{t} \int\limits_{\mu_n t}^{r} G(h)^{m-n} \frac{T}{t} [1 + \frac{1}{n-1} \frac{R^2}{T^2}] \frac{R^{n-3}}{r^{n-2}} dh d\mathcal{F}(r) + \int\limits_{t_+}^{1} \int\limits_{\mu_n t}^{t} G(h)^{m-n} \frac{T}{r} \frac{R^{n-3}}{r^{n-2}} [1 + \frac{1}{n-1} \frac{R^2}{T^2}] dh d\mathcal{F}(r)} \tag{204}$$

Without loss of generality, we may assume that $\mathcal{F}(r) = 0 \;\forall\; r < \mu_n t$. Else, we could easily normalize, because $[0, \mu_n t)$ is irrelevant for the evaluation of (204).
First look at the quotient of the two left summands only. For every single radius quotient with $r \leq t$ we have

$$\frac{\int\limits_{\mu_n t}^{r} G(h)^{m-n} \frac{R^{n-3}}{r^{n-2}} \, dh}{\int\limits_{\mu_n t}^{r} G(h)^{m-n} \frac{T}{t} [1 + \frac{1}{n-1} \frac{R^2}{T^2}] \frac{R^{n-3}}{r^{n-2}} \, dh} \leq \frac{\int\limits_{\mu_n t}^{r} G(h)^{m-n} \frac{R^{n-3}}{r^{n-2}} [1] \, dh}{\int\limits_{\mu_n t}^{r} G(h)^{m-n} \frac{T}{t} \frac{R^{n-3}}{r^{n-2}} [1] \, dh} \leq$$

$$\leq \frac{\int\limits_{\mu_n t}^{r} G(h)^{m-n} \frac{R^{n-3}}{r^{n-2}} [1 + \mu_{n-1}] \, dh}{\int\limits_{\mu_n t}^{r} G(h)^{m-n} \frac{T}{t} \frac{R^{n-3}}{r^{n-2}} [1 + \frac{1}{n-1}] \, dh} \leq \frac{\int\limits_{\mu_n t}^{t} G(h)^{m-n} [1 + \mu_{n-1}] \frac{T^{n-3}}{t^{n-2}} \, dh}{\int\limits_{\mu_n t}^{t} G(h)^{m-n} \frac{T}{t} [1 + \frac{1}{n-1}] \frac{T^{n-3}}{t^{n-2}} \, dh}. \tag{205}$$

The second inequality exploits $1 + \frac{1}{n-1} < 1 + \mu_{n-1}$ from (92) and the third (205) is justified by comprehending that a transformation of weights has taken place to get from the new form to the old form according to

$$\begin{cases} \frac{R^{n-3}}{T^{n-3}} & \text{for } h \leq r \\ 0 & \text{for } r \leq h \leq t \end{cases} \text{ (monotonously decreasing with } h\text{)}.$$

Since $r \leq t$ and $R \leq T$, this transformation supports the lower values of h and the larger values of $\frac{T}{t}$. So the quotient is smaller in the old form.

This observation motivates a possible change of the radial distribution. Consider the radial distribution function

$$\tilde{\mathcal{F}}^{(t)}(r) := \begin{cases} 0 & r < t \\ \mathcal{F}(r) & r \geq t \end{cases} . \tag{206}$$

Let us use this function in (204) and let us do the replacement suggested in (205). This leads to a new quotient of integrals.

$$\frac{\int_{\mu_n t}^{t} G(h)^{m-n} \frac{T^{n-3}}{t^{n-2}}[1 + \mu_{n-1}]\, dh \cdot \tilde{\mathcal{F}}^{(t)}(t) + \int_{t_+}^{1} \int_{\mu_n t}^{t} G(h)^{m-n}[1 + \mu_{n-1}\frac{R}{T}]\frac{R^{n-3}}{r^{n-2}}\, dh\, d\tilde{\mathcal{F}}^{(t)}(t)}{\int_{\mu_n t}^{t} G(h)^{m-n}\frac{T}{t}[1 + \frac{1}{n-1}]\frac{T^{n-3}}{t^{n-2}}dh \cdot \tilde{\mathcal{F}}^{(t)}(t) + \int_{t_+}^{1} \int_{\mu_n t}^{t} G(h)^{m-n}\frac{T}{r}\frac{R^{n-3}}{r^{n-2}}[1 + \frac{1}{n-1}\frac{R^2}{T^2}]dhd\tilde{\mathcal{F}}^{(t)}(r)}. \tag{207}$$

The right summands in numerator and in denominator will not be involved by the modification, because we let them start at t_+. But the left summands will be changed. Our change from (204) to (207) has enlarged the left summand in the numerator of (204), i. e. the integral mass for [0,t]. The reasons are the factor $1 + \mu_{n-1}$ and the insight that the figure

$$\int_{\mu_n t}^{r} G(h)^{m-n} R^{n-3} r^{-n+2}\, dh = \int_{\frac{\mu_n t}{r}}^{1} G(qr)^{m-n}\sqrt{1-q^2}^{\,n-3} dq \tag{208}$$

increases monotonously with r. As a result of the monotony of G, we observe an increment of these terms for growing r until t. We conclude

$$\int_{\mu_n t}^{t} \int_{\mu_n t}^{r} G(h)^{m-n} R^{n-3} r^{-n+2}\, dh\, d\mathcal{F}(r) \leq \int_{\mu_n t}^{t} G(h)^{m-n}T^{n-3}t^{-n+2}\, dh \cdot [\mathcal{F}(t) - \mathcal{F}(0)]. \tag{209}$$

Consequently, the use of $\tilde{\mathcal{F}}^{(t)}$ in (207) instead of $\mathcal{F}$ as in (204) delivers a higher quotient of the left summands because of (205). And it provides a higher integral mass for the region $[0, t]$, as (209) shows.
Hence, this new quotient will not be smaller than the quotient in (204), if $\tilde{\mathcal{F}}^{(t)}(t) = \mathcal{F}(t) = 1$.

Let us now study the case $\tilde{\mathcal{F}}^{(t)}(t) = \mathcal{F}(t) < 1$. Up to now it is impossible to decide, whether the total quotient has changed and in which direction. Therefore, we simplify the notation for the moment and look at the structure of the modifications. Let us use $N_{\leq t}, N_{>t}, D_{\leq t}, D_{>t}$ for the two numerator- and the two denominator-terms in (204), which had been based on $\mathcal{F}$. Then the structure of (204) can be described as

$$\frac{N_{\leq t} + N_{>t}}{D_{\leq t} + D_{>t}} . \tag{210}$$

Accordingly, we use $\tilde{N}_{\leq t}, \tilde{N}_{>t}, \tilde{D}_{\leq t}, \tilde{D}_{>t}$ for the fragments of (207). Our change leads to an increment of the quotient of the left summands, i.e. there is a real number $\alpha \leq 1$ such that

$$\frac{N_{\leq t}}{D_{\leq t}} \leq \frac{\tilde{N}_{\leq t}}{\tilde{D}_{\leq t}} = \frac{N_{\leq t}}{\alpha \cdot D_{\leq t}} .$$

Besides that it leads to an increment of the left numerator-summand, i.e. there is a real number $\beta \geq 1$ such that

$$N_{\leq t} \leq \tilde{N}_{\leq t} = \beta \cdot N_{\leq t}.$$

Now we have

$$\frac{\tilde{N}_{\leq t} + \tilde{N}_{>t}}{\tilde{D}_{\leq t} + \tilde{D}_{>t}} = \frac{\beta \cdot N_{\leq t} + N_{>t}}{\alpha \cdot \beta \cdot D_{\leq t} + D_{>t}}. \tag{211}$$

Although the quotient in the area of "$\leq t$" is larger now, we cannot finally compare (211) with (210). To clarify the effect of the modification, we could introduce another real $\gamma \leq 1$, which restores the original balance of the two numerator-summands, with the property

$$\frac{\gamma \cdot \beta \cdot N_{\leq t}}{\frac{1 - \gamma \cdot \mathcal{F}(t)}{1 - \mathcal{F}(t)} \cdot N_{>t}} = \frac{N_{\leq t}}{N_{>t}}, \text{ respectively } \gamma \cdot \beta = \frac{1 - \gamma \cdot \mathcal{F}(t)}{1 - \mathcal{F}(t)}.$$

The existence of such a value is clear. Choose

$$\gamma = \frac{1}{1 + (1 - \mathcal{F}(t))(\beta - 1)} \leq 1.$$

Switching to another distribution function, namely

$$\hat{\mathcal{F}}^{(t)}(r) = \begin{cases} 0 & \text{for } r < t \\ \gamma \cdot \mathcal{F}(t) & \text{for } r = t \\ \gamma \cdot \mathcal{F}(t) + [1 - \gamma \cdot \mathcal{F}(t)] \cdot \frac{\mathcal{F}(r) - \mathcal{F}(t)}{1 - \mathcal{F}(t)} & \text{for } r > t, \end{cases} \tag{212}$$

with γ defined as above, would lead to the relation

$$\frac{N_{\leq t} + N_{>t}}{D_{\leq t} + D_{>t}} \leq \frac{\gamma \cdot \beta \cdot N_{\leq t} + \frac{1 - \gamma \cdot \mathcal{F}(t)}{1 - \mathcal{F}(t)} N_{>t}}{\gamma \cdot \alpha \cdot \beta \cdot D_{\leq t} + \frac{1 - \gamma \cdot \mathcal{F}(t)}{1 - \mathcal{F}(t)} D_{>t}} = \frac{N_{\leq t} + N_{>t}}{\alpha \cdot D_{\leq t} + D_{>t}} = \frac{\hat{N}_{\leq t} + \hat{N}_{>t}}{\hat{D}_{\leq t} + \hat{D}_{>t}} \text{ since } \alpha \leq 1.$$

Note that $\hat{N}_{\leq t}, \hat{N}_{>t}, \hat{D}_{\leq t}, \hat{D}_{>t}$ denote the fragments under use of the new distribution function $\hat{\mathcal{F}}^{(t)}(r)$ corresponding to (207), where $\tilde{\mathcal{F}}$ had been used.

So we arrive at the desired simplification $Q_3 \leq \hat{Q}_3$, as defined in (144) in the main text. $\square$

References

Adler, I., R. M. Karp, R. Shamir (1987). A Simplex Variant Solving an $m \times d$ Linear Program in $O(min(m^2, d^2))$ Expected Number of Steps. *Journal of Complexity* **3** 372–387.

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

Borgwardt, K.H.(1977). *Untersuchungen zur Asymptotik der mittleren Schrittzahl von Simplexverfahren in der linearen Optimierung*, Dissertation Universität Kaiserslautern.

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

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

Borgwardt, K.H. (1987). *The Simplex Method, A Probabilistic Analysis*, Springer Verlag, Heidelberg.

Borgwardt, K.H., R. Damm, R. Donig, G. Joas (1993). Empirical Studies on the Average Efficiency of Simplex Variants under Rotation Symmetry, *ORSA Journal on Computing* **5** 249–260.

Borgwardt, K. H.(1994a). Verschärfung des Polynomialitätsbeweises für die erwartete Anzahl von Schattenecken im Rotationssymmetrie-Modell. In *Beiträge zur Angewandten Analysis und Informatik, Festschrift zu Ehren von Helmut Brakhage*, E. Schock, ed., Shaker Verlag, Aachen.

Borgwardt, K.H.(1994b). *Improving the theoretical upper bound for the expected number of shadowvertices in the Rotation-Symmetry-Model*, Schwerpunktreport der Deutschen Forschungsgemeinschaft Nr. 537, Universität Augsburg.

Haimovich, M.(1983). *The Simplex Algorithm is Very Good! - On the Expected Number of Pivot Steps and Related Properties of Random Linear Programs*. Report Columbia University, New York.

Höfner, G. (1995). *Lineare Optimierung mit dem Schatteneckenalgorithmus –Untersuchungen zum mittleren Rechenaufwand und Entartungsverhalten*, Dissertation Universität Augsburg.

Huhn, P. (1997). *Schranken für die durchschnittliche Laufzeit von Simplexverfahren und Innere-Punkt-Verfahren*, Dissertation Universität Augsburg.

Küfer, K.(1995). On the variance of the number of pivot steps required by the simplex algorithm. *Mathematical Methods of Operations Research* **42** 1–24.

Küfer, K.(1996). An improved asymptotic analysis of the number of pivot steps required by the simplex algorithm. *Mathematical Methods of Operations Research* **44** 147–170.

Shamir, R.(1987). The Efficiency of the Simplex Method. *Management Science* **33** 301–334.

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

Todd, M.J. (1986). Polynomial Expected Behavior of a Pivoting Algorithm for Linear Complimentary and Linear Programming Problems. *Mathematical Programming* **35** 173–192.