Threshold phenomena for random cones

Daniel Hug and Rolf Schneider

Abstract

We consider an even probability distribution on the d-dimensional Euclidean space with the property that it assigns measure zero to any hyperplane through the origin. Given N independent random vectors with this distribution, under the condition that they do not positively span the whole space, the positive hull of these vectors is a random polyhedral cone (and its intersection with the unit sphere is a random spherical polytope). It was first studied by Cover and Efron. We consider the expected face numbers of these random cones and describe a threshold phenomenon when the dimension d and the number N of random vectors tend to infinity. In a similar way, we treat the solid angle, and more generally the Grassmann angles. We further consider the expected numbers of k-faces and of Grassmann angles of index $d-k$ when also k tends to infinity.

Keywords: Cover–Efron cone, face numbers, solid angle, Grassmann angle, high dimensions, threshold phenomenon

2020 Mathematics Subject Classification: Primary 60D05

1 Introduction

The following is a literal quotation from [5]: "Recent work has exposed a phenomenon of abrupt *phase transitions* in high-dimensional geometry. The phase transitions amount to a rapid shift in the likelihood of a property's occurrence when a dimension parameter crosses a critical level (a *threshold*)." Two early observations of this phenomenon were published in 1992. Dyer, Füredi and McDiarmid [8] considered the convex hull of $N = N(d)$ vertices chosen independently at random (with equal chances) from the vertices of the unit cube in $\mathbb{R}^d$. Let $V_{d,N}$ denote the volume of this random polytope. Then, for every $\varepsilon > 0$,

$$\lim_{d\to\infty} \mathbb{E}\, V_{d,N} = \begin{cases} 1, & \text{if } N \geq (2e^{-1/2}+\varepsilon)^d, \\ 0, & \text{if } N \leq (2e^{-1/2}-\varepsilon)^d. \end{cases}$$

Here $\mathbb{E}$ denotes mathematical expectation. The paper [8] has a similar result for the convex hull of i.i.d. uniform random points from the unit cube. We have quoted this example as an illustration of what we have in mind: for instance, a d-dimensional random polytope with its number N of vertices depending on d, where a small change of this dependence causes an abrupt change of some property as $d \to \infty$. In the work of Vershik and Sporyshev [15], a d-dimensional random polytope is obtained as a uniform random orthogonal projection of a fixed regular simplex with N vertices in a higher-dimensional space, and threshold phenomena are exhibited for the expected numbers of k-faces, under the assumption of a linearly coordinated growth of the parameters d, N, k. Similar models, also with the regular simplex replaced by the regular cross-polytope, and random projections extended to more general random linear mappings, have found important applications in the work of Donoho and collaborators. We refer to the paper of Donoho and Tanner [6], where also earlier work of these authors is cited and explained. The paper [7] of the same authors treats random projections of the cube and the positive orthant in a similar way. Generally in stochastic geometry, threshold phenomena have been investigated for face numbers, neighborliness properties, volumes, intrinsic volumes, more general measures, and for several different models of random polytopes. Different phase transitions were exhibited. We mention that [9] has extended the model of [8] by introducing more general distributions for the random points. The paper [13] considers convex hulls of i.i.d. random points with either Gaussian distribution or uniform distribution on the unit sphere. In [1], [2], the points have a beta or beta-prime distribution. The paper [3] studies facet numbers of convex hulls of random points on the unit sphere in different regimes. The papers [13] and [1] deal also with polytopes generated by intersections of random closed halfspaces.

In this note, we consider a model of random polyhedral convex cones (or, equivalently, of random spherical polytopes) that was introduced by Cover and Efron [4] (and more closely investigated in [11]). Let ϕ be a probability measure on the Euclidean space $\mathbb{R}^d$ which is even (invariant under reflection in the origin o) and assigns measure zero to each hyperplane through the origin. For $n \in \mathbb{N}$, the (ϕ, n)-*Cover–Efron cone* C_n is defined as the positive hull of n independent random vectors $X_1, \dots, X_n$ with distribution ϕ, under the condition that this positive hull is different from $\mathbb{R}^d$. The intersection $C_n \cap \mathbb{S}^{d-1}$ with the unit sphere $\mathbb{S}^{d-1}$ is a spherical random polytope, contained in some halfsphere. In the following it will be convenient to work with polyhedral cones instead of spherical polytopes.

For $k \in \{1, \dots, d-1\}$, let $f_k(C_n)$ denote the number of k-dimensional faces of the cone C_n (equivalently, the number of $(k-1)$-dimensional faces of the spherical polytope $C_n \cap \mathbb{S}^{d-1}$). We are interested in the asymptotic behavior of the expectation $\mathbb{E}\, f_k(C_n)$, as d tends to infinity and n grows suitably with d.

Convention. In the following, C_N is a Cover–Efron cone in $\mathbb{R}^d$, and $N = N(d) \geq d$ is an integer depending on the dimension d, but we will omit the dimension d in the notation. Further (until stated otherwise), k is a fixed positive integer, and we consider only dimensions $d > k$.

Any k-face of the cone C_N is a.s. the positive hull of k vectors from $X_1, \dots, X_N$, and there are $\binom{N}{k}$ possible choices. Therefore, we consider the quotient $\mathbb{E}\, f_k(C_N)/\binom{N}{k}$. For this, we can state the following phase transition.

Theorem 1. *Suppose that*

$$\frac{d}{N} \to \delta \quad \textit{as } d \to \infty,$$

with a number $\delta \in [0,1]$. Then the Cover–Efron cone C_N satisfies

$$\lim_{d\to\infty} \frac{\mathbb{E}\, f_k(C_N)}{\binom{N}{k}} = \begin{cases} 1 & \textit{if } 1/2 < \delta \leq 1, \\ (2\delta)^k & \textit{if } 0 \leq \delta < 1/2. \end{cases}$$

We do not know what happens if $\delta = 1/2$. However, if N is close to $2d$, or even equal to it, then more precise asymptotic statements are possible

Theorem 2. *If $N - 2d$ is bounded from above, then*

$$\lim_{d\to\infty} \frac{\mathbb{E}\, f_k(C_N)}{\binom{N}{k}} = 1.$$

If $N = 2d$, then

$$\lim_{d\to\infty} \sqrt{d}\left(1 - \frac{\mathbb{E}\, f_k(C_N)}{\binom{N}{k}}\right) = \frac{k}{\sqrt{\pi}}.$$

Of course, the first part of Theorem 2 implies the first part of Theorem 1.

As another functional of a closed convex cone $C \subset \mathbb{R}^d$, we consider the solid angle $v_d(C)$. This is the normalized spherical Lebesgue measure of $C \cap \mathbb{S}^{d-1}$. (We avoid the notation $V_d(C)$ used in [11], since V_d is often used for the volume of a convex body. The reader is warned that what we denote here by $v_d(C)$ was denoted by $v_{d-1}(C \cap \mathbb{S}^{d-1})$ in [14, Sect. 6.5].)

More generally, we consider the Grassmann angles. For a closed convex cone $C \subset \mathbb{R}^d$ which is not a subspace, the jth *Grassmann angle* of C, for $j \in \{1, \dots, d-1\}$, is defined by

$$U_j(C) := \frac{1}{2}\mathbb{P}(C \cap \mathcal{L} \neq \{o\}),$$

where $\mathcal{L}$ is a random $(d-j)$-dimensional subspace with distribution ν_{d-j}. The latter is the unique Haar probability measure on $G(d, d-j)$, the Grassmannian of $(d-j)$-dimensional linear subspaces of $\mathbb{R}^d$. Thus,

$$U_j(C) = \frac{1}{2}\int_{G(d,d-j)} \mathbb{1}\{C \cap L \neq \{o\}\}\, \nu_{d-j}(\mathrm{d}L).$$

Grassmann angles were introduced by Grünbaum [10], in a slightly different, though equivalent way. Grünbaum's Grassmann angles are given by $\gamma^{m,d} = 1 - 2U_{d-m}$. We note that $v_d = U_{d-1}$, and that $U_j(C) \leq 1/2$, with equality if C is a halfspace.

Theorem 3. *Suppose that*

$$\frac{d}{N} \to \delta \quad \text{as } d \to \infty,$$

with a number $\delta \in [0, 1]$. Then the Cover–Efron cone C_N satisfies

$$\lim_{d\to\infty} \mathbb{E}\, 2U_{d-k}(C_N) = \begin{cases} 0 & \text{if } 1/2 < \delta \leq 1, \\ 1 - \left(\frac{\delta}{1-\delta}\right)^k & \text{if } 0 \leq \delta < 1/2. \end{cases}$$

Again, the first part of this theorem has a stronger version, given by the first part of the following theorem.

Theorem 4. *If $N - 2d$ is bounded from above, then*

$$\lim_{d\to\infty} \mathbb{E}\, U_{d-k}(C_N) = 0.$$

If $N = 2d$, then

$$\lim_{d\to\infty} \sqrt{d} \cdot \mathbb{E}\, U_{d-k}(C_N) = \frac{k}{\sqrt{\pi}}.$$

Clearly, in Theorem 1 (and similarly in Theorem 3), the change when passing a threshold is not so abrupt as in the other mentioned examples: below the threshold $\delta = 1/2$, the limit in question increases (decreases) with the parameter to an extremal value, above the threshold, it remains constant. The situation changes if also the number k, the dimension of the considered faces, increases with the dimension; then a more abrupt phase transition is observed. Under a linearly coordinated growth, for k-faces we find the same threshold as established by Donoho and Tanner [7] in their investigation of random linear images of orthants. This comes unexpected, since the random cones considered in [7] and here have different distributions.

We replace the convention made earlier by the following one.

Convention. In the following theorems and in Sections 6 and 7, $N = N(d) \geq d$ and $k = k(d) < d$ are integers depending on the dimension d, but we will omit the argument d in the notation.

As in [7], we define

$$\rho_W(\delta) := \max\{0, 2 - \delta^{-1}\} \quad \text{for } 0 < \delta < 1$$

(the index W stands for 'weak' threshold).

Theorem 5. *Let $0 < \delta, \rho < 1$ be given. Let $k < d < N$ be integers such that*

$$\frac{d}{N} \to \delta, \qquad \frac{k}{d} \to \rho \qquad \text{as } d \to \infty.$$

Then

$$\lim_{d\to\infty} \frac{\mathbb{E}\, f_k(C_N)}{\binom{N}{k}} = \begin{cases} 1 & \text{if } \rho < \rho_W(\delta), \\ 0 & \text{if } \rho > \rho_W(\delta). \end{cases}$$

We note that the first assumption of this theorem, $\rho < \rho_W(\delta)$, implies that for large d we have $N + k < 2d$.

Adapting an argument of Donoho and Tanner [7] to the present situation, we can also replace the convergence of an expectation in the first part of Theorem 5 by the convergence of a probability, at the cost of a smaller threshold.

Theorem 6. *Let $0 < \delta, \rho < 1$ be given, where $\delta > 1/2$. Let $k < d < N$ be integers such that*

$$\frac{d}{N} \to \delta, \qquad \frac{k}{d} \to \rho \qquad \text{as } d \to \infty.$$

There exists a positive number $\rho_S(\delta)$ such that

$$\lim_{d\to\infty} \mathbb{P}\left(f_k(C_N) = \binom{N}{k}\right) = 1 \quad \text{if } \rho < \rho_S(\delta).$$

There is also a counterpart to Theorem 5 for Grassmann angles.

Theorem 7. *Let $0 < \delta, \rho < 1$ be given. Let $k < d < N$ be integers such that*

$$\frac{d}{N} \to \delta, \qquad \frac{k}{d} \to \rho \qquad \text{as } d \to \infty.$$

Then

$$\lim_{d\to\infty} \mathbb{E}\, 2U_{d-k}(C_N) = \begin{cases} 0 & \text{if } \rho < \frac{1}{2}\rho_W(\delta), \\ 1 & \text{if } \rho > \frac{1}{2}\rho_W(\delta). \end{cases}$$

After some preliminaries in the next section, we collect a number of auxiliary results about sums of binomial coefficients in Section 3. Then we prove the first two theorems in Section 4, Theorems 3 and 4 in Section 5, Theorems 5 and 6 in Section 6, and Theorem 7 in Section 7.

2 Preliminaries

First we recall two classical facts. For $n \in \mathbb{N}$, let $H_1, \dots, H_n \in G(d, d-1)$. Assume that these hyperplanes are in general position, that is, the intersection of any $m \leq d$ of them is of dimension $d - m$. Then the number of d-dimensional cones in the tessellation of $\mathbb{R}^d$ induced by these hyperplanes is given by

$$C(n, d) := 2 \sum_{i=0}^{d-1} \binom{n-1}{i}.$$

From this result of Steiner (in dimension three) and Schläfli, Wendel has deduced the following. If $X_1, \dots, X_n$ are i.i.d. random vectors in $\mathbb{R}^d$ with distribution ϕ (enjoying the properties mentioned above), then

$$P_{d,n} := \mathbb{P}\left(\mathrm{pos}\{X_1, \dots, X_n\} \neq \mathbb{R}^d\right) = \frac{C(n,d)}{2^n},$$

where $\mathbb{P}$ stands for probability and pos denotes the positive hull. For references and proofs, we refer to [14, Sect. 8.2.1]. Now we can write down the distribution of the Cover–Efron cone C_n, namely

$$\mathbb{P}\left(C_n \in B\right) = \frac{1}{P_{d,n}} \int_{(\mathbb{S}^{d-1})^n} \mathbb{1}\{\mathrm{pos}\{x_1, \dots, x_n\} \in B \setminus \{\mathbb{R}^d\}\} \, \phi^n(\mathrm{d}(x_1, \dots, x_n))$$

for $B \in \mathcal{B}(\mathcal{C}^d)$, where $\mathcal{C}^d$ denotes the space of closed convex cones in $\mathbb{R}^d$ (with the topology of closed convergence) and $\mathcal{B}(\mathcal{C}^d)$ is its Borel σ-algebra.

There is an equivalent representation of C_n. For this, we denote by ϕ^* the image measure of ϕ under the mapping $x \mapsto x^\perp$ from $\mathbb{R}^d \setminus \{o\}$ to $G(d, d-1)$. Let $\mathcal{H}_1, \dots, \mathcal{H}_n$ be i.i.d. random hyperplanes with distribution ϕ^*. They are almost surely in general position. The (ϕ^*, n)-*Schläfli cone* S_n is obtained by picking at random (with equal chances) one of the d-dimensional cones from the tessellation induced by $\mathcal{H}_1, \dots, \mathcal{H}_n$. Its distribution is given by

$$\mathbb{P}\left(S_n \in B\right) = \int_{G(d,d-1)^n} \frac{1}{C(n,d)} \sum_{C \in \mathcal{F}_d(H_1, \dots, H_n)} \mathbb{1}\{C \in B\} \, \phi^{*n}(\mathrm{d}(H_1, \dots, H_n))$$

for $B \in \mathcal{B}(\mathcal{C}^d)$, where $\mathcal{F}_d(H_1, \dots, H_n)$ is the set of d-cones in the tessellation induced by $H_1, \dots, H_n$. We have (see [11, Thm. 3.1])

$$C_n = S_n^\circ \quad \text{in distribution},$$

where S_n° denotes the polar cone of S_n.

For the expectations appearing in our theorems, explicit representations are available. The proofs of Theorems 1, 2 and 5 are based on the formula

$$\frac{\mathbb{E}\, f_k(C_N)}{\binom{N}{k}} = 2^k \frac{C(N-k, d-k)}{C(N,d)} = \frac{P_{d-k,N-k}}{P_{d,N}} \tag{1}$$

for $k \in \{0, \dots, d-1\}$ (see [4, Formula (3.3)] or [11, Formula (27)].) For the proofs of Theorems 3, 4 and 7, we use the explicit formula

$$\mathbb{E}\, U_j(C_N) = \frac{C(N,d) - C(N,j)}{2C(N,d)} \tag{2}$$

for $j \in \{1, \dots, d-1\}$ (see [11, (29)]). It is sometimes useful to write this in the form

$$\mathbb{E}\, 2U_{d-k}(C_N) = \frac{1 + \binom{N-1}{d-1}^{-1} \sum_{i=d-k}^{d-2} \binom{N-1}{i}}{1 + \binom{N-1}{d-1}^{-1} \sum_{i=0}^{d-2} \binom{N-1}{i}} \tag{3}$$

for $k \in \{1, \dots, d-1\}$ (where an empty sum is zero, by definition).

3 Auxiliary results on binomial coefficients

From the last expressions it is obvious that some information on binomial coefficients is required. First we note Stirling's formula

$$n! = \sqrt{2\pi n}\, e^{-n} n^n e^{\theta/12n}, \quad 0 < \theta < 1. \tag{4}$$

It implies, in particular, that

$$\binom{2n}{n} \sim \frac{2^{2n}}{\sqrt{\pi n}} \quad \text{as } n \to \infty. \tag{5}$$

We need some information on the Wendel probabilities

$$P_{d,n} = \frac{1}{2^{n-1}} \sum_{i=0}^{d-1} \binom{n-1}{i}.$$

The following lemma considers the reciprocal of the quotient appearing in (1).

Lemma 1. *For $k \in \{1, \dots, d-1\}$,*

$$\frac{P_{d,N}}{P_{d-k,N-k}} = 1 + \frac{1}{2^{N-1} P_{d-k,N-k}} \sum_{j=1}^{k} \binom{k}{j} \sum_{m=0}^{j-1} \binom{N-k-1}{d-k+m}.$$

Proof. Writing $\binom{N-1}{i} = \binom{N-2}{i-1} + \binom{N-2}{i}$ for $i = 1, \dots, d-1$, we obtain

$$P_{d,N} = \frac{1}{2} P_{d-1,N-1} + \frac{1}{2} P_{d,N-1}.$$

This and induction can be used to prove that

$$P_{d,N} = \frac{1}{2^k} \sum_{j=0}^{k} \binom{k}{j} P_{d-k+j,N-k}.$$

For $j \in \{1, \dots, k\}$ we have

$$P_{d-k+j,N-k} = P_{d-k,N-k} + \frac{1}{2^{N-k-1}} \sum_{m=0}^{j-1} \binom{N-k-1}{d-k+m}.$$

This gives

$$\begin{aligned} \frac{P_{d,N}}{P_{d-k,N-k}} &= \frac{1}{2^k} \sum_{j=0}^{k} \binom{k}{j} \frac{P_{d-k+j,N-k}}{P_{d-k,N-k}} \\ &= \frac{1}{2^k} \left[\binom{k}{0} + \sum_{j=1}^{k} \binom{k}{j} \left(1 + \frac{1}{2^{N-k-1} P_{d-k,N-k}} \sum_{m=0}^{j-1} \binom{N-k-1}{d-k+m} \right) \right] \end{aligned}$$

and thus the assertion. □

Lemma 2.

$$P_{d-k,N-k} \geq \frac{1}{2} - \frac{1}{2^{N-k}} \sum_{r=0}^{N-2d+k-1} \binom{N-k-1}{d-k+r}.$$

If $N-2d+k-1 < 0$, the sum is zero, by convention.

Proof. Let M, p be integers. If $1 \leq p \leq \frac{M}{2}$, then

$$\begin{aligned} 2^M = \sum_{i=0}^{M} \binom{M}{i} &= \sum_{i=0}^{p-1} \binom{M}{i} + \binom{M}{p} + \cdots + \binom{M}{M-p} + \sum_{i=M-p+1}^{M} \binom{M}{i} \\ &= 2\sum_{i=0}^{p-1} \binom{M}{i} + \sum_{r=0}^{M-2p} \binom{M}{p+r}, \end{aligned}$$

thus

$$\sum_{i=0}^{p-1} \binom{M}{i} = 2^{M-1} - \frac{1}{2} \sum_{r=0}^{M-2p} \binom{M}{p+r}.$$

If $p > \frac{M}{2}$, we have

$$2^M = \sum_{i=0}^{M} \binom{M}{i} \leq 2 \sum_{i=0}^{p-1} \binom{M}{i}.$$

Hence, for arbitrary $p \geq 1$ we may write

$$\sum_{i=0}^{p-1} \binom{M}{i} \geq 2^{M-1} - \frac{1}{2} \sum_{r=0}^{M-2p} \binom{M}{p+r},$$

with the convention that the last sum is zero if $2p > M$. The choice $M = N-k-1$ and $p = d-k$ now gives the assertion. □

The following lemma gives upper and lower bounds for the expressions appearing in (3). For the proof of the upper bound, we adjust and slightly refine the argument for Proposition 1(c) in [12], in the current framework. The improved lower bound in (8) will be crucial in the following.

Lemma 3. *Let $n \in \mathbb{N}$, $m \in \mathbb{N}_0$ and $2m \leq n+1$.*

(a) *If $2m < n+1$, then*

$$\frac{1}{\binom{n}{m+1}} \sum_{j=0}^{m} \binom{n}{j} \leq \frac{m+1}{n-m} \cdot \frac{n-m+1}{n-2m+1} \left(1 - \left(\frac{m}{n-m+1} \right)^{m+1} \right). \tag{6}$$

If $2m = n+1$, then

$$\frac{1}{\binom{n}{m+1}} \sum_{j=0}^{m} \binom{n}{j} \leq \frac{(m+1)^2}{n-m}. \tag{7}$$

(b) *If $2 \le \ell \le m$, then*

$$\frac{m-\ell+1}{n-2m+2\ell-1}\left(1-\left(\frac{m-\ell+1}{n-m+\ell}\right)^{\ell+1}\right) \le \frac{1}{\binom{n}{m+1}}\sum_{j=0}^{m}\binom{n}{j}. \tag{8}$$

Moreover,

$$\frac{m+1}{n-m} \le \frac{m+1}{n-m}\cdot\frac{n+1}{n+1-m} \le \frac{1}{\binom{n}{m+1}}\sum_{j=0}^{m}\binom{n}{j}. \tag{9}$$

Proof. (a) The cases $n=1$, $m \in \{0,1\}$ and $n \ge 2$, $m=0$ are easy to check. Now let $n \ge 2$, $m \ge 1$, and hence also $m \le n-1$. If $j \in \{0,\dots,m-1\}$, then

$$\frac{\binom{n}{j}}{\binom{n}{m}} = \frac{m}{n-m+1}\cdots\frac{j+1}{n-j} \le \left(\frac{m}{n-m+1}\right)^{m-j},$$

since

$$\frac{j+1}{n-j} = -1+\frac{n+1}{n-j} \le -1+\frac{n+1}{n-(m-1)} = \frac{m}{n-m+1} \in (0,1].$$

Therefore, if $2m < n+1$, then $0 < q_0 := m/(n-m+1) < 1$ and

$$\frac{1}{\binom{n}{m+1}}\sum_{j=0}^{m}\binom{n}{j} \le \frac{\binom{n}{m}}{\binom{n}{m+1}}\sum_{j=0}^{m} q_0^{m-j} = \frac{m+1}{n-m}\cdot\frac{1-q_0^{m+1}}{1-q_0},$$

which implies (6). If $2m = n+1$, then $m/(n-m+1)=1$, and (7) follows similarly.

(b) Note that $m \le (n+1)/2 \le n$, and $m \le n-1$ if $n \ge 2$. Hence, if $2 \le \ell \le m$, then $m \le n-1$, $n-2m+2\ell-1 \ge 2$, $(n+1)/(m+1) > 1$ and $0 < q_1 := (m-\ell+1)/(n-m+\ell) < 1$. Then, for $j \in \{m-\ell,\dots,m\}$ we obtain

$$\begin{aligned}\frac{\binom{n}{j}}{\binom{n}{m+1}} &= \left(\frac{n+1}{m+1}-1\right)^{-1}\cdots\left(\frac{n+1}{j+1}-1\right)^{-1} \ge \left(\frac{n+1}{j+1}-1\right)^{-(m+1-j)}\\ &\ge \left(\frac{m-\ell+1}{n-m+\ell}\right)^{m+1-j} = q_1^{m+1-j}\end{aligned}$$

and hence

$$\frac{1}{\binom{n}{m+1}}\sum_{j=m-\ell}^{m}\binom{n}{j} \ge q_1\sum_{r=0}^{\ell} q_1^r = \frac{q_1}{1-q_1}\left(1-q_1^{\ell+1}\right),$$

which yields (8).

If $m=n$, then (9) holds trivially, since $\binom{n}{n+1}=0$. It also holds for $m=0$. In the remaining cases, we have

$$\frac{1}{\binom{n}{m+1}}\sum_{j=0}^{m}\binom{n}{j} \ge \frac{\binom{n}{m}+\binom{n}{m-1}}{\binom{n}{m+1}} = \frac{m+1}{n-m}\cdot\frac{n+1}{n+1-m}.$$

This completes the proof (b). $\square$

From (6) and (9) with $n = N-1$ and $m = d-2$, we deduce that

$$\frac{d-1}{N-d+1} \le \frac{1}{\binom{N-1}{d-1}} \sum_{j=0}^{d-2} \binom{N-1}{j} \le \frac{d-1}{N-d+1} \cdot \frac{N-d+2}{N-2d+4}, \tag{10}$$

if $N > 2d-4$.

Lemma 4. *If $d/N \to \delta$ with $0 \le \delta < 1/2$, then*

$$\lim_{d\to\infty} \frac{1}{\binom{N-1}{d-1}} \sum_{j=0}^{d-2} \binom{N-1}{j} = \frac{\delta}{1-2\delta}.$$

Proof. Assume that $d/N \to \delta$ as $d \to \infty$, with a number $0 \le \delta < 1/2$. We write $N = \alpha d$, where α depends on d and satisfies $\alpha \to \delta^{-1}$ as $d \to \infty$. If $\delta = 0$, this means that $\alpha \to \infty$. We assume that d is so large that $\alpha > 2$. From (10) we have

$$\frac{1}{\binom{N-1}{d-1}} \sum_{j=0}^{d-2} \binom{N-1}{j} \le \frac{d-1}{(\alpha-1)d+1} \cdot \frac{(\alpha-1)d+2}{(\alpha-2)d+4}.$$

We conclude that

$$\limsup_{d\to\infty} \frac{1}{\binom{N-1}{d-1}} \sum_{j=0}^{d-2} \binom{N-1}{j} \le \frac{\delta}{1-2\delta}. \tag{11}$$

Lemma 3(b) provides the lower bound

$$\frac{1}{\binom{N-1}{d-1}} \sum_{j=0}^{d-2} \binom{N-1}{j} \ge \frac{d-\ell_1-1}{N-2d+2\ell_1+2} \left(1 - \left(\frac{d-\ell_1-1}{N-d+\ell_1+1}\right)^{\ell_1+1}\right),$$

if $2 \le \ell_1 \le d-2$. From this, we obtain

$$\liminf_{d\to\infty} \frac{1}{\binom{N-1}{d-1}} \sum_{j=0}^{d-2} \binom{N-1}{j} \ge \frac{\delta}{1-2\delta} \left(1 - \left(\frac{\delta}{1-\delta}\right)^{\ell_1+1}\right)$$

for each fixed $\ell_1 \ge 2$. Letting $\ell_1 \to \infty$, we find that

$$\liminf_{d\to\infty} \frac{1}{\binom{N-1}{d-1}} \sum_{j=0}^{d-2} \binom{N-1}{j} \ge \frac{\delta}{1-2\delta}.$$

Together with (11) this completes the proof. □

We state another simple lemma.

Lemma 5. *Let $m \in \mathbb{N}$. Then*

$$\sum_{i=0}^{m} \binom{2m}{i} = 2^{2m-1} + \frac{1}{2}\binom{2m}{m}, \qquad \sum_{i=0}^{m-1} \binom{2m-1}{i} = 2^{2m-2}.$$

Proof. We use $\binom{n}{\ell} = \binom{n}{n-\ell}$. If $x := \sum_{i=0}^{m} \binom{2m}{m}$, then

$$2x = \sum_{i=0}^{2m} \binom{2m}{i} + \binom{2m}{m} = 2^{2m} + \binom{2m}{m},$$

which gives the first relation. If $y := \sum_{i=0}^{m-1} \binom{2m-1}{i}$, then

$$2y = \sum_{i=0}^{2m-1} \binom{2m-1}{i} = 2^{2m-1},$$

which gives the second relation. $\square$

4 Proofs of Theorems 1 and 2

Proof of Theorem 1. As already mentioned, the first part of Theorem 1 follows from the first part of Theorem 2, which will be proved below.

To prove the second part of Theorem 1, we assume that $d/N \to \delta$ as $d \to \infty$, where $0 \le \delta < 1/2$. We write (1) in the form

$$\frac{\mathbb{E}\, f_k(C_N)}{\binom{N}{k}} = 2^k \frac{\binom{N-k-1}{d-k-1}\left[1 + \binom{N-k-1}{d-k-1}^{-1} \sum_{i=0}^{d-k-2} \binom{N-k-1}{i}\right]}{\binom{N-1}{d-1}\left[1 + \binom{N-1}{d-1}^{-1} \sum_{i=0}^{d-2} \binom{N-1}{i}\right]}, \tag{12}$$

and here

$$\frac{\binom{N-k-1}{d-k-1}}{\binom{N-1}{d-1}} = \frac{d-1}{N-1} \cdots \frac{d-k}{N-k} \to \delta^k \quad \text{as } d \to \infty.$$

Since also $(d-k)/(N-k) \to \delta$, we deduce from Lemma 4 that the normalized sums in the numerator and denominator of (12) tend to the same finite limit. It follows that $\lim_{d\to\infty} \mathbb{E}\, f_k(C_N)/\binom{N}{k} = (2\delta)^k$. $\square$

Proof of Theorem 2. First we assume that $N - 2d$ is bounded from above. From (1) and Lemma 1 we have

$$\begin{aligned} \left(\frac{\mathbb{E}\, f_k(C_N)}{\binom{N}{k}}\right)^{-1} &= \frac{P_{d,N}}{P_{d-k,N-k}} \\ &= 1 + \frac{1}{2^{N-1} P_{d-k,N-k}} \sum_{j=1}^{k} \binom{k}{j} \sum_{m=0}^{j-1} \binom{N-k-1}{d-k+m} \\ &=: 1 + A, \end{aligned} \tag{13}$$

where A depends on d, N, k. Here

$$A^{-1} = \frac{2^{N-1} P_{d-k,N-k}}{\sum_{j=1}^{k} \binom{k}{j} \sum_{m=0}^{j-1} \binom{N-k-1}{d-k+m}} \ge \frac{2^{N-2} - 2^{k-1} \sum_{r=0}^{N-2d+k-1} \binom{N-k-1}{d-k+r}}{\sum_{j=1}^{k} \binom{k}{j} \sum_{m=0}^{j-1} \binom{N-k-1}{d-k+m}}, \tag{14}$$

by Lemma 2.

The argument is different according to whether $N-2d+k-1<0$ or not. Therefore, we consider separately the subsequence of all dimensions d for which $N-2d+k-1<0$ and the subsequence for which $N-2d+k-1\geq 0$. Without loss of generality, we will assume that one of the assumptions holds for the whole sequence.

Assume, first, that $N-2d+k-1<0$ for all d. We recall that in this case the sum in the last numerator of (14) is zero, hence

$$A \leq \sum_{j=1}^{k} \binom{k}{j} \sum_{m=0}^{j-1} 2^{2-N} \binom{N-k-1}{d-k+m}.$$

Let $m \in \{0,\dots,k-1\}$. If $N-k-1$ is even, we write $N-k-1=2p$. Note that $N-k-1 < 2d-2k \leq 2d-2k+2m$, hence $d-k+m \geq p$. Therefore, we get

$$\binom{N-k-1}{d-k+m} \leq \binom{2p}{p} \sim \frac{2^{N-k-1}}{\sqrt{\pi p}}$$

as $d\to\infty$ (which implies $p\to\infty$), by (5). If $N-k-1$ is odd, we write $N-k=2p$. Then $N-k<2d-2k+1$, hence $N-k\leq 2(d-k)\leq 2(d-k+m)$, so that again $d-k+m\geq p$. This yields

$$\binom{N-k-1}{d-k+m} = \frac{N-d-m}{N-k}\binom{N-k}{d-k+m} < \binom{2p}{p} \sim \frac{2^{N-k}}{\sqrt{\pi p}}$$

as $d\to\infty$. In each case we get

$$\lim_{d\to\infty} 2^{2-N}\binom{N-k-1}{d-k+m} = 0$$

for $m=0,\dots,k-1$ and hence $\lim_{d\to\infty} A = 0$.

Now we assume that $N-2d+k-1\geq 0$ for all d. We write

$$N = 2d-k+1+s,$$

where $s=s(d)$ may depend on d. Thus, $s\geq 0$, and by assumption, $s(d)$ is bounded by a constant.

We compare the different binomial coefficients appearing in (14) with a special one. For this, we set

$$\frac{\binom{N-k-1}{d-k+r}}{\binom{2(d-k)}{d-k}} = \frac{\binom{2(d-k)+s}{d-k+r}}{\binom{2(d-k)}{d-k}} =: f(d,k,r,s).$$

Thus,

$$f(d,k,r,s) = \frac{(2d-2k+s)\cdots(2d-2k+1)}{(d-k+r)\cdots(d-k+1)\cdot(d-k+s-r)\cdots(d-k+1)} \quad \text{if } s\geq r,$$

$$f(d,k,r,s) = \frac{(2d-2k+s)\cdots(2d-2k+1)\cdot(d-k+s-r+1)\cdots(d-k)}{(d-k+r)\cdots(d-k+1)} \quad \text{if } s<r.$$

If $s'\geq 0$ is a fixed integer, then $\lim_{d\to\infty} f(d,k,r,s')$ is positive and finite. Since the nonnegative function s is bounded, it attains only finitely many values. Therefore, there exist two positive constants c_1, c_2 with

$$c_1 \leq f(d,k,r,s) \leq c_2 \quad \text{for all } d$$

and also

$$c_1 \le f(d,k,m,s) \le c_2 \quad \text{for } m \le k \text{ and all } d.$$

We have

$$2^{k-1} \sum_{i=0}^{N-2d+k-1} \binom{N-k-1}{d-k+r} = P(d,k)\binom{2(d-k)}{d-k}$$

with

$$P(d,k) := 2^{k-1} \sum_{r=0}^{s(d)} f(d,k,r,s(d)).$$

Similarly,

$$\sum_{j=1}^{k} \binom{k}{j} \sum_{m=0}^{j-1} \binom{N-k-1}{d-k+m} = Q(d,k)\binom{2(d-k)}{d-k}$$

with

$$Q(d,k) := \sum_{j=1}^{k} \binom{k}{j} \sum_{m=0}^{j-1} f(d,k,m,s(d)).$$

This gives

$$A^{-1} \ge \frac{1}{Q(d,k)} \left[\frac{2^{N-2}}{\binom{2(d-k)}{d-k}} - P(d,k) \right].$$

As $d \to \infty$, the terms $P(d,k)$ and $Q(d,k)$ remain between two positive constants, and

$$\frac{2^{N-2}}{\binom{2(d-k)}{d-k}} \ge \frac{2^{2d-k-1}}{\binom{2(d-k)}{d-k}} \sim \frac{2^{2d-k-1}}{2^{2d-2k}/\sqrt{\pi(d-k)}} \to \infty$$

by (5). It follows that $A \to 0$ as $d \to \infty$. By (13), this implies that $\lim_{d\to\infty} \mathbb{E}\, f_k(C_N)/\binom{N}{k} = 1$.

To prove the second part, suppose that $N = 2d$. Then by (1) we have

$$\frac{\mathbb{E}\, f_k(C_N)}{\binom{N}{k}} = 2^k \cdot \frac{\sum_{i=0}^{d-k-1} \binom{2d-k-1}{i}}{\sum_{i=0}^{d-1} \binom{2d-1}{i}}.$$

We distinguish two cases. If k is odd, say $k = 2\ell - 1$ with $\ell \in \mathbb{N}$, then

$$\begin{aligned} \sum_{i=0}^{d-k-1} \binom{2d-k-1}{i} &= \sum_{i=0}^{d-\ell} \binom{2d-2\ell}{i} - \sum_{i=d-2\ell+1}^{d-\ell} \binom{2d-2\ell}{i} \\ &= 2^{2d-2\ell-1} + \frac{1}{2}\binom{2d-2\ell}{d-\ell} - \sum_{j=0}^{\ell-1} \binom{2d-2\ell}{d-1-j}, \end{aligned}$$

where Lemma 5 was used. Since Stirling's formula (4) yields

$$\binom{2d-2\ell}{d-1-j} \sim \frac{4^{d-\ell}}{\sqrt{\pi d}}, \quad j = 0, \dots \ell - 1,$$

it follows (again using Lemma 5) that

$$\begin{aligned} \frac{\mathbb{E}\, f_k(C_N)}{\binom{N}{k}} &\sim 2^{2\ell-1} \frac{1}{4^{d-1}} \left(\frac{1}{2} 4^{d-\ell} + \frac{1}{2} \frac{4^{d-\ell}}{\sqrt{\pi d}} - \ell \frac{4^{d-\ell}}{\sqrt{\pi d}} \right) \\ &= 1 - 2 \cdot 4^{\ell-d} \left(\ell - \frac{1}{2} \right) \frac{4^{d-\ell}}{\sqrt{\pi d}} = 1 - (2\ell - 1) \frac{1}{\sqrt{\pi d}}. \end{aligned}$$

If k is even, say $k = 2\ell$ with $\ell \in \mathbb{N}$, then

$$\begin{aligned}\sum_{i=0}^{d-k-1} \binom{2d-k-1}{i} &= \sum_{i=0}^{d-\ell-1} \binom{2(d-\ell)-1}{i} - \sum_{j=1}^{\ell} \binom{2(d-\ell)-1}{d-j} \\ &= 2^{2d-2\ell-2} - \sum_{j=1}^{\ell} \binom{2(d-\ell)-1}{d-j}\end{aligned}$$

by Lemma 5. By Stirling's approximation (4),

$$\binom{2(d-\ell)-1}{d-j} \sim \frac{1}{2}\frac{4^{d-\ell}}{\sqrt{\pi d}}, \quad j = 1, \dots, \ell,$$

and hence we get

$$\begin{aligned}\frac{\mathbb{E}\, f_k(C_N)}{\binom{N}{k}} &\sim 2^{2\ell}\frac{1}{4^{d-1}}\left(2^{2d-2\ell-2} - \ell\frac{1}{2}\frac{4^{d-\ell}}{\sqrt{\pi d}}\right) \\ &= 1 - (2\ell)\frac{1}{\sqrt{\pi d}}.\end{aligned}$$

Thus in both cases the asymptotic relation is proved. $\square$

5 Proofs of Theorems 3 and 4

Proof of Theorem 3. The first part of Theorem 3 follows from the first part of Theorem 4, which will be proved below.

For the second part of the proof, we assume that $d/N \to \delta$ as $d \to \infty$, with $0 \le \delta < 1/2$. We note that relation (2) shows that

$$\mathbb{E}\, 2U_{d-k}(C_N) = 1 - \frac{C(N, d-k)}{C(N,d)}.$$

Let

$$R(N-1, d-\ell) := \frac{1}{\binom{N-1}{d-\ell}} \sum_{i=0}^{d-\ell-1} \binom{N-1}{i}, \quad \ell \in \{0, \dots, d-1\}.$$

Then

$$\frac{C(N, d-k)}{C(N,d)} = \frac{\binom{N-1}{d-k}}{\binom{N-1}{d}} \frac{R(N-1, d-k)}{R(N-1,d)},$$

and here

$$\frac{\binom{N-1}{d-k}}{\binom{N-1}{d}} = \frac{d}{N-d} \cdots \frac{d-(k-1)}{N-d+k-1} \to \left(\frac{\delta}{1-\delta}\right)^k \quad \text{as } d \to \infty.$$

Together with $d/N \to \delta$ we have $(d-k+1)/N \to \delta$ and $(d+1)/N \to \delta$, hence Lemma 4 yields that

$$R(N-1, d-k) \to \frac{\delta}{1-2\delta} \quad \text{and} \quad R(N-1,d) \to \frac{\delta}{1-2\delta}.$$

The assertion follows. $\square$

Proof of Theorem 4. First we assume that $N \le 2d+b$ with a fixed number b. We want to make use of (3). Writing $n := \lfloor \frac{N}{2} \rfloor$, we have

$$\sum_{i=0}^{n} \binom{N-1}{i} \ge 2^{N-2}$$

and hence

$$\sum_{i=0}^{d-2} \binom{N-1}{i} \ge 2^{N-2} - \sum_{i=d-1}^{n} \binom{N-1}{i}, \tag{15}$$

where an empty sum is zero. For $d-1 < i \le n$ we have $(N-i)/i \le N/i \le N/d$ and hence

$$\frac{\binom{N-1}{i}}{\binom{N-1}{d-1}} = \frac{(N-i)\cdots(N-d)}{i\cdots d} \le \left(\frac{N}{d}\right)^{i-d+1} \le (2+b)^{1+b/2}.$$

The last sum in (15) has at most $n-d+2 \le b+2$ summands. Therefore, there is a constant c, independent of d, such that

$$\frac{1}{\binom{N-1}{d-1}} \sum_{i=0}^{d-2} \binom{N-1}{i} \ge \frac{2^{N-2}}{\binom{N-1}{d-1}} - c \ge \frac{2^{N-2}}{\binom{N-1}{n'}} - c \sim \frac{2^{N-2}}{2^{N-1}/\sqrt{\pi n'}} - c \to \infty$$

as $d \to \infty$, where $n' := \lfloor \frac{N-1}{2} \rfloor$ and where (5) was used.

Moreover, for $d-k \le i \le d-2$ we have $(N-1)/i \le N/(d-k) \le 2N/d$ for sufficiently large d. This shows that

$$\frac{1}{\binom{N-1}{d-1}} \sum_{i=d-k}^{d-2} \binom{N-1}{i}$$

remains bounded as $d \to \infty$. Now (3) shows that $\lim_{d\to\infty} \mathbb{E}\, U_{d-k}(C_N) = 0$.

To prove the second part, let $N = 2d$. Using Lemma 5, we get

$$\frac{C(2d, d-k)}{C(2d,d)} = \frac{\sum_{i=0}^{d-1}\binom{2d-1}{i} - \sum_{i=d-k}^{d-1}\binom{2d-1}{i}}{\sum_{i=0}^{d-1}\binom{2d-1}{i}} = 1 - 4^{1-d} \sum_{j=0}^{k-1} \binom{2d-1}{d-1-j}.$$

By Stirling's approximation (4), we obtain for $j \in \{0, \dots, k-1\}$ that

$$\binom{2d-1}{d-1-j} \sim \frac{4^d}{2} \frac{1}{\sqrt{\pi d}}.$$

Hence we get

$$\frac{C(2d,d-k)}{C(2d,d)} \sim 1 - 4^{1-d} \cdot k \cdot \frac{1}{2} 4^d \frac{1}{\sqrt{\pi d}} = 1 - 2k \frac{1}{\sqrt{\pi d}}.$$

Thus we arrive at

$$\mathbb{E}\, U_{d-k}(C_{2d}) \sim \frac{1}{2} - \frac{1}{2}\left(1 - 2k\frac{1}{\sqrt{\pi d}}\right) = \frac{k}{\sqrt{\pi d}},$$

which completes the proof of the second part. $\square$

6 Faces of increasing dimension

In this section and the next, we allow also the number k to depend on the dimension d. In the present section, we are interested in a phase transition for the expectation $\mathbb{E}\, f_k(C_N)/\binom{N}{k}$. It turns out that it appears at the same threshold as it was observed earlier by Donoho and Tanner [7] for a different, but closely related class of random polyhedral cones. These authors considered a real random $d \times N$ matrix A of rank d, where $d < N$, the nonnegative orthant

$$\mathbb{R}^N_+ := \{x \in \mathbb{R}^N : x_i \geq 0 \text{ for } i = 1, \dots, N\}$$

of $\mathbb{R}^N$, and its image $\mathsf{A}\mathbb{R}^N_+$ in $\mathbb{R}^d$. Considering the column vectors of A as random vectors in $\mathbb{R}^d$, the image $\mathsf{A}\mathbb{R}^N_+$ is just the positive hull of these vectors. For a suitable distribution, the random cone $\mathsf{A}\mathbb{R}^N_+$ is obtained in a similar way as the Cover–Efron cone, just by omitting the condition that the cone is different from $\mathbb{R}^d$. Imposing this condition leads, of course, to different distributions of the random cones. Comparing formula (13) of [7] with our formula (1), where the right-hand side can be written as

$$\frac{1 - P_{N-d,N-k}}{P_{d,N}},$$

we see that it results in an additional denominator in the expression for the expected number of k-faces, thus increasing this expectation. Therefore our estimates, though leading to the same threshold, require more effort.

Proof of Theorem 5. We define

$$\frac{d}{N} =: \delta_d, \qquad \frac{k}{d} =: \rho_d,$$

then $\delta_d \to \delta$ and $\rho_d \to \rho$ as $d \to \infty$.

First we assume that $0 < \rho < \rho_W(\delta)$. For sufficiently large d, we then have $0 < \rho_d < \rho_W(\delta_d)$, which implies that $\delta_d > 1/2$ and $N - 2d + k < 0$. We assume in the following that d is so large that this holds.

In Section 4 we have proved that

$$\left(\frac{\mathbb{E}\, f_k(C_N)}{\binom{N}{k}} \right)^{-1} = 1 + A,$$

where (under the assumption that $N - 2d + k - 1 < 0$, which is now satisfied)

$$A \leq 2^{2-N} \sum_{j=1}^{k} \binom{k}{j} \sum_{m=0}^{j-1} \binom{N-k-1}{d-k+m}. \tag{16}$$

For $m \in \{0, \dots, k-1\}$ we have $\binom{N-k-1}{d-k+m} \leq \binom{N-k-1}{d-k}$ (since $d - k \geq (N-k-1)/2$). Therefore, the identities

$$\binom{N-k-1}{d-k} = \frac{N-d}{N-k} \cdot \frac{(N-k)!}{(N-d)!(d-k)!}, \qquad \sum_{j=1}^{k} \binom{k}{j} j = 2^{k-1} \cdot k, \tag{17}$$

and Stirling's formula (4) yield

$$
\begin{aligned}
A &\le 2\frac{\rho_d\delta_d N}{2^{(1-\rho_d\delta_d)N}}\frac{1-\delta_d}{1-\rho_d\delta_d}\frac{\sqrt{2\pi(N-k)}(N-k)^{N-k}e^{\frac{\theta}{12(N-k)}}}{\sqrt{2\pi(N-d)}(N-d)^{N-d}\sqrt{2\pi(d-k)}(d-k)^{d-k}}\\
&\le 2\frac{\rho_d\delta_d\sqrt{1-\delta_d}}{\sqrt{1-\rho_d\delta_d}}N\,F(\delta_d,\rho_d)^N,
\end{aligned}
$$

where we also used that $1 \le e^{\frac{\theta}{12}} \le 2$ for $\theta \in (0,1)$, $d-k \ge 1$ and $2/\sqrt{2\pi} < 1$. Here the function F is defined by

$$
F(a,b) := \left(\frac{1-ab}{2(1-a)}\right)^{1-ab}\left(\frac{1-a}{a(1-b)}\right)^{a(1-b)}, \quad a \in (0,1),\ b \in [0,1).
$$

We need only consider the case $a > 1/2$. For fixed $a > 1/2$, we consider $F_a(b) := F(a,b)$ for $0 < b \le 2-a^{-1}$. Note that $2-a^{-1} \in (0,1)$ for $a > 1/2$. Differentiation yields

$$
F_a'(b) = a\log\left(\frac{2a(1-b)}{1-ab}\right)F(a,b),
$$

and hence $F_a'(b) > 0$ whenever

$$
\frac{2a(1-b)}{1-ab} > 1, \quad \text{equivalently} \quad b < 2-a^{-1}.
$$

This implies that $F(a,b) = F_a(b) < F_a(2-a^{-1}) = 1$ for $b < 2-a^{-1}$.

Since $F(\delta_d,\rho_d) \to F(\delta,\rho)$ as $d \to \infty$, we conclude that $F(\delta_d,\rho_d) \le c < 1$ if d is large enough, with a constant c independent of d. It follows that $\lim_{d\to\infty} A = 0$ and hence that

$$
\lim_{d\to\infty}\frac{\mathbb{E}\,f_k(C_N)}{\binom{N}{k}} = 1.
$$

Now we assume that $\rho > \rho_W(\delta)$. Then $\rho > 2-\delta^{-1}$, irrespective of whether $\delta \ge 1/2$ or not. For sufficiently large d (which we assume in the following), we then have $\rho_d > 2-\delta_d^{-1}$, which implies $N-2d+k > 0$. Thus, we can apply (10) to the normalized sum in the numerator of (12). This yields

$$
\frac{d-k-1}{N-d+1} \le \binom{N-k-1}{d-k-1}^{-1}\sum_{i=0}^{d-k-2}\binom{N-k-1}{i} \le \frac{d-k-1}{N-d+1}\cdot\frac{N-d+2}{N-2d+k+4}.
$$

Here,

$$
\lim_{d\to\infty}\frac{d-k-1}{N-d+1} = \frac{\delta(1-\rho)}{1-\delta}, \qquad \lim_{d\to\infty}\frac{N-d+2}{N-2d+k+4} = \frac{1-\delta}{1-2\delta+\rho\delta},
$$

where the last denominator is positive. It follows that

$$
c_1 \le \binom{N-k-1}{d-k-1}^{-1}\sum_{i=0}^{d-k-2}\binom{N-k-1}{i} \le c_2 \tag{18}
$$

for all sufficiently large d. Here and below we denote by c_i a positive constant that is independent of d.

In view of (12), we now determine the asymptotic behavior of

$$2^k \frac{\binom{N-k-1}{d-k-1}}{\binom{N-1}{d-1}} = 2^k \frac{N}{d} \frac{d-k}{N-k} \frac{\binom{d}{k}}{\binom{N}{k}} \quad \text{as } d \to \infty.$$

Here,

$$\lim_{d\to\infty} \frac{N}{d} \frac{d-k}{N-k} = \frac{1-\rho}{1-\rho\delta}.$$

To treat the remaining terms, we use the Stirling formula (4). This gives

$$\begin{aligned} 2^k \frac{\binom{d}{k}}{\binom{N}{k}} &= 2^{\rho_d\delta_d N} \sqrt{\frac{1-\rho_d\delta_d}{1-\rho_d}} \cdot \frac{(\delta_d N)^{\delta_d N}(N-\rho_d\delta_d N)^{N-\rho_d\delta_d N}}{(\delta_d N - \rho_d\delta_d N)^{\delta_d N - \rho_d\delta_d N} N^N} \cdot e^{\varphi/12N} \\ &= e^{\varphi/12N} \sqrt{\frac{1-\rho_d\delta_d}{1-\rho_d}} H(\delta_d, \rho_d)^N, \end{aligned} \tag{19}$$

where φ is contained in a fixed interval independent of d, and

$$H(a,b) := \frac{(2a)^{ab}(1-ab)^{1-ab}}{(1-b)^{a(1-b)}}, \quad a \in (0,1),\, b \in [0,1).$$

We define

$$g(a) := H\left(a, 2-a^{-1}\right) = 2a^a(1-a)^{1-a} \qquad \text{for } a \in (0,1) \tag{20}$$

and remark that $H(a,b) = g(a)F(a,b)$. Note that for $a, b \in (0,1)$ we have $b > 2 - a^{-1}$ if and only if $a < 1/(2-b)$. Let $H_a(b) := H(a,b)$. Differentiation yields

$$H_a'(b) = a \log\left(\frac{2a(1-b)}{1-ab}\right) H(a,b).$$

Hence $H_a'(b) < 0$ for $b > 2 - a^{-1}$, since

$$\frac{2a(1-b)}{1-ab} < 1 \Leftrightarrow b > 2 - a^{-1}.$$

If $a \leq 1/2$, then

$$H(a,b) < H(a,0) = 1$$

for $b \in (0,1)$. On the other hand, if $a > 1/2$ and $b > 2 - a^{-1}$, we have

$$H(a,b) < H\left(a, 2-a^{-1}\right) = 2a^a(1-a)^{1-a}. \tag{21}$$

Since the function g defined by (20) satisfies $g(1/2) = 1$ and $g'(a) > 0$ for $1/2 < a < 1$, we also have

$$H(a, 2-a^{-1}) > 1 \quad \text{if } a > 1/2. \tag{22}$$

Now we distinguish two cases.

(1) Let $\delta \leq 1/2$. Then (12), (18) and (19) yield

$$\begin{aligned} \limsup_{d\to\infty} \frac{\mathbb{E}\, f_k(C_N)}{\binom{N}{k}} &\leq \frac{1-\rho}{1-\rho\delta} \frac{\sqrt{1-\rho\delta}}{\sqrt{1-\rho}} \limsup_{d\to\infty} H(\delta_d, \rho_d)^N \frac{1+c_2}{1} \\ &\leq (1+c_2) \limsup_{d\to\infty} c_3^N = 0, \end{aligned}$$

where $H(\delta_d, \rho_d) \leq c_3 < 1$, since $H(\delta_d, \rho_d) \to H(\delta, \rho) < H(\delta, 0) = 1$.

(2) Let $\delta > 1/2$. Then we can assume that $N/d < c_4 < 2$. We have

$$\sum_{i=0}^{d-2} \binom{N-1}{i} = 2^{N-1} - \sum_{j=0}^{N-d} \binom{N-1}{j}.$$

Since $2(N-d) < N$, Lemma 3 yields

$$\binom{N-1}{N-d+1}^{-1} \sum_{j=0}^{N-d} \binom{N-1}{j} \leq \frac{\frac{N}{d-1} - 1}{2 - \frac{N}{d}},$$

and hence

$$\binom{N-1}{d-1}^{-1} \sum_{i=0}^{d-2} \binom{N-1}{i} \geq \binom{N-1}{d-1}^{-1} 2^{N-1} - \frac{1}{2 - \frac{N}{d}}.$$

To estimate the last binomial coefficient, we use Stirling's approximation (4) together with (21). Thus, we get for large d the lower bound

$$\binom{N-1}{d-1}^{-1} \sum_{i=0}^{d-2} \binom{N-1}{i} \geq c_5 H\left(\delta_d, 2 - \delta_d^{-1}\right)^N - c_6.$$

Combining these estimates and starting again from (12), we finally obtain

$$\begin{aligned}
\limsup_{d\to\infty} \frac{\mathbb{E}\, f_k(C_N)}{\binom{N}{k}} &\leq c_7 \limsup_{d\to\infty} H(\delta_d, \rho_d)^N \frac{1 + c_2}{1 + c_5 H(\delta_d, 2 - \delta_d^{-1})^N - c_6} \\
&= c_8 \limsup_{d\to\infty} \left(\frac{H(\delta_d, \rho_d)}{H(\delta_d, 2 - \delta_d^{-1})} \right)^N \frac{1}{(1 - c_6) H(\delta_d, 2 - \delta_d^{-1})^{-N} + c_5} \\
&= 0.
\end{aligned}$$

Here we have used that

$$\frac{H(\delta_d, \rho_d)}{H(\delta_d, 2 - \delta_d^{-1})} \to \frac{H(\delta, \rho)}{H(\delta, 2 - \delta^{-1})} < 1 \quad \text{for } \rho > 2 - \delta^{-1}$$

by (21) and that

$$H(\delta_d, 2 - \delta_d^{-1}) \to H(\delta, 2 - \delta^{-1}) > 1 \quad \text{for } \delta > \frac{1}{2}$$

by (22). This completes the proof also in the case $\rho > \rho_W(\delta)$. $\square$

We prepare the proof of Theorem 6 by a lemma, which serves to establish the threshold ρ_S and to provide an upper estimate for it. By H we denote the binary entropy function with base e, that is

$$\mathsf{H}(x) := -x \log x - (1 - x) \log(1 - x) \quad \text{for } 0 \leq x \leq 1$$

(with $0 \log 0 := 0$). We note that $\mathsf{H}(0) = \mathsf{H}(1) = 0$ and that H attains its unique maximum, $\log 2$, at the point $1/2$. As in [7], we consider the function defined by

$$G(\delta, \rho) := \mathsf{H}(\delta) + \delta \mathsf{H}(\rho) - (1 - \rho\delta) \log 2, \quad \rho, \delta \in [0, 1].$$

For a later application, we remark that

$$e^{-G(\delta,\rho)} = (1 - \delta)^{1-\delta} \delta^\delta (1 - \rho)^{\delta(1-\rho)} \rho^{\delta\rho} 2^{1-\delta\rho}. \tag{23}$$

Lemma 6. *For $\delta \in (1/2, 1)$, the function G_δ defined by $G_\delta(x) := G(\delta, x)$ has a unique zero $x_0 \in [0, 1]$. Moreover, $x_0 \in (0, \min\{\frac{2}{3}, 2 - \delta^{-1}\})$.*

Proof. Clearly, $G_\delta(0) = \mathsf{H}(\delta) - \log 2 < 0$ since $\delta \neq 1/2$. We have

$$G'_\delta(x) = \delta \log\left(\frac{2(1-x)}{x}\right).$$

Hence $x_0 = 2/3$ is the unique zero of G'_δ in $(0, 1)$, and $G'_\delta > 0$ in $(0, 2/3)$ and $G'_\delta < 0$ in $(2/3, 1)$. We will show that

(a) $G_\delta(2/3) > 0$,

(b) $G_\delta(1) > 0$,

(c) $G_\delta(2 - \delta^{-1}) > 0$,

which then implies that G_δ has a unique zero x_0 in $[0, 1]$ and $x_0 < 2/3$ as well as $x_0 < 2 - \delta^{-1}$.

For (a) we define $\kappa_1(\delta) := G_\delta(2/3)$ for $\delta \in [1/2, 1]$. Then a simple calculation shows that $\kappa_1(1/2) = 1/2 \log 3 > 0$, $\kappa_1(1) = \log 3 - \log 2 > 0$ and

$$\kappa'_1(\delta) = \log\left(\frac{3(1-\delta)}{\delta}\right), \quad \delta \in (1/2, 1).$$

Hence $\kappa'_1(3/4) = 0$, $\kappa'_1 > 0$ on $(1/2, 3/4)$ and $\kappa'_1 < 0$ on $(3/4, 1)$. This shows that $\kappa_1(\delta) = G_\delta(2/3) > 0$ for $\delta \in (1/2, 1)$ (with maximal value $\kappa_1(3/4) = \log 2$).

For (b) we consider $\kappa_2(\delta) := G_\delta(1)$ for $\delta \in [1/2, 1]$. We have $\kappa_2(1/2) = 1/2 \log 2 > 0$, $\kappa_2(1) = 0$ and

$$\kappa'_2(\delta) = \log\left(\frac{2(1-\delta)}{\delta}\right), \quad \delta \in (1/2, 1).$$

Hence, $\kappa'_2(2/3) = 0$, $\kappa'_2 > 0$ on $(1/2, 2/3)$ and $\kappa'_2 < 0$ on $(2/3, 1)$. In particular, this yields $G_\delta(1) = \kappa_2(\delta) > 0$ for $\delta \in (1/2, 1)$.

For (c) we consider $\kappa_3(\delta) = G_\delta(2 - \delta^{-1})$ for $\delta \in [1/2, 1]$. Then

$$\kappa_3(\delta) = -2(1-\delta)\log(1-\delta) - (2\delta - 1)\log(2\delta - 1) - 2(1-\delta)\log 2$$

and

$$\kappa'_3(\delta) = 2\log\left(\frac{2(1-\delta)}{2\delta - 1}\right), \quad \delta \in (1/2, 1).$$

We have $\kappa_3(1/2) = \kappa_3(1) = 0$, $\kappa'_3(3/4) = 0$, $\kappa'_3 > 0$ on $(1/2, 3/4)$ and $\kappa'_3 < 0$ on $(3/4, 1)$. But then $\kappa_3 > 0$ on $(1/2, 1)$ and therefore $G_\delta(2 - \delta^{-1}) = \kappa_3(\delta) > 0$ for $\delta \in (1/2, 1)$. $\square$

Now we denote the zero x_0 of the function G_δ provided by Lemma 6 by $\rho_S(\delta)$.

Proof of Theorem 6. Again we define

$$\frac{d}{N} =: \delta_d, \qquad \frac{k}{d} =: \rho_d.$$

Let $\delta > 1/2$ and $0 < \rho < \rho_S(\delta)$. Then $G(\delta, \rho) < 0$. For sufficiently large d, we have $G(\delta_d, \rho_d) < 0$ as well as $\delta_d > 1/2$ and $\rho_d < \rho_S(\delta_d)$. We assume that d is large enough in this sense. Since $\rho_s(\delta_d) < \rho_W(\delta_d)$, we have $N - 2d + k < 0$. We can, therefore, apply the estimates from the first part of the proof of Theorem 5.

Let $X_1, \dots, X_N$ be i.i.d. unit vectors with distribution ϕ. By definition, C_N is the positive hull of $X_1, \dots, X_N$ under the condition that this positive hull is different from $\mathbb{R}^d$. For $k \in \{1, \dots, d-1\}$, choose $1 \le i_1 < \dots < i_k \le N$ and let $M = \{X_{i_1}, \dots, X_{i_k}\}$. Then the distribution of $\operatorname{pos} M$, under the condition that $\operatorname{pos}\{X_1, \dots, X_N\} \neq \mathbb{R}^d$, is independent of the choice of $i_1, \dots, i_k$, hence

$$\binom{N}{k} \mathbb{P}\left(\operatorname{pos} M \in \mathcal{F}_k(C_N)\right) = \mathbb{E}\, f_k(C_N).$$

Therefore,

$$p := \mathbb{P}\left(\operatorname{pos} M \notin \mathcal{F}_k(C_N)\right) = 1 - \frac{\mathbb{E}\, f_k(C_N)}{\binom{N}{k}} = \frac{A}{1+A}$$

by (13) (and with the notation used there). By Boole's inequality,

$$\mathbb{P}\left(\operatorname{pos} M \notin \mathcal{F}_k(C_N) \text{ for some } k\text{-element subset } M\right) \le \binom{N}{k} p$$

and thus

$$\mathbb{P}\left(f_k(C_N) = \binom{N}{k}\right) \ge 1 - \binom{N}{k} p.$$

Here,

$$\binom{N}{k} p \le \binom{N}{k} A.$$

Now we use (16), (17) and (4) to get

$$\begin{aligned}\binom{N}{k} A &\le 2^{2-N} \frac{N-d}{N-k} k 2^{k-1} \frac{N!}{(N-d)!(d-k)!k!} \\ &= 2 \frac{(1-\delta_d)\delta_d \rho_d}{1-\delta_d \rho_d} N \frac{\sqrt{N}}{2\pi\sqrt{N-d}\sqrt{d-k}\sqrt{k}} e^{G(\delta_d, \rho_d)N} e^{\frac{\varphi}{12N}},\end{aligned}$$

where (23) was used and where $\varphi \in (0,1)$. Since $G(\delta_d, \rho_d) \to G(\delta, \rho) < 0$ as $d \to \infty$, it follows that

$$\lim_{d \to \infty} \mathbb{P}\left(f_k(C_N) = \binom{N}{k}\right) = 1,$$

as stated. □

7 Proof of Theorem 7

We use the representation

$$\mathbb{E}\, 2U_{d-k}(C_N) = 1 - \frac{C(N, d-k)}{C(N, d)}$$

and show the convergence of the quotient, under different assumptions.

First we deal with the second part of the proof and point out that our argument requires to distinguish whether $\rho > \rho_W(\delta)$ or not. We begin with the case $\rho > \rho_W(\delta)$; then $\rho \ge 2 - \delta^{-1}$. Clearly,

$$\frac{C(N, d-k)}{C(N, d)} = \frac{\binom{N-1}{d-k-1}}{\binom{N-1}{d-1}} \cdot \frac{1 + \binom{N-1}{d-k-1}^{-1} \sum_{i=0}^{d-k-2} \binom{N-1}{i}}{1 + \binom{N-1}{d-1}^{-1} \sum_{i=0}^{d-2} \binom{N-1}{i}}.$$

Since (a fortiori) $\rho > 1-(2\delta)^{-1}$, we have $N-2d+2k>0$ for sufficiently large d, hence Lemma 3 yields

$$\frac{1}{\binom{N-1}{d-k-1}}\sum_{i=0}^{d-k-2}\binom{N-1}{i} \le \frac{d-k-1}{N-d+k+1}\cdot\frac{N-d+k+2}{N-2d+2k+4} \to \frac{\delta(1-\rho)}{1-2\delta(1-\rho)} \tag{24}$$

as $d\to\infty$, and the last denominator is positive. Hence, if d is large enough (which is always assumed in the following), there are constants C_1, C_2, independent of d, such that

$$\begin{aligned}0 \le \frac{C(N,d-k)}{C(N,d)} &\le C_1\frac{d-k}{d}\frac{d!(N-d)!}{(d-k)!(N-d+k)!}\\ &\le C_2\frac{d^d(N-d)^{N-d}}{(d-k)^{d-k}(N-d+k)^{N-d+k}} \le C_2\cdot K(\delta_d,\rho_d)^N,\end{aligned}$$

where

$$K(a,b) := \frac{g(a)}{g(a(1-b))} = \frac{a^{ab}(1-a)^{1-a}}{(1-b)^{a(1-b)}(1-a(1-b))^{1-a(1-b)}}, \quad a\in(0,1),\, b\in[0,1).$$

We have $K(a,0)=1$, and also $K(a,2-a^{-1})=1$ if $a\ge 1/2$. We write $K_a(b) := K(a,b)$. Then

$$K_a'(b) = a\log\left(\frac{a(1-b)}{1-a(1-b)}\right)K(a,b) < 0 \quad \text{for} \quad b > 1-(2a)^{-1} \text{ (where } b\ge 0).$$

Thus, for any $a\in(0,1)$, we have $K(a,b)<1$ for all $b>2-a^{-1}$ in $(0,1)$.

Since $\rho > 2-\delta^{-1}$, we deduce that $K(\delta,\rho)<1$ and hence that $K(\delta_d,\rho_d)\le c<1$ for all sufficiently large d, with c independent of d. It follows that

$$\lim_{d\to\infty}\frac{C(N,d-k)}{C(N,d)} = 0. \tag{25}$$

Now we suppose that $\frac{1}{2}\rho_W(\delta) < \rho \le \rho_W(\delta)$, then $1-(2\delta)^{-1} < \rho \le 2-\delta^{-1}$. Since $\rho>0$, we have $\delta>1/2$, further $\delta(1-\rho)<1/2$.

We use repeatedly that

$$2^{N-1} = \sum_{i=0}^{d-1}\binom{N-1}{i} + \sum_{i=0}^{N-d-1}\binom{N-1}{i}.$$

We note that still $N-2d+2k>0$ for sufficiently large d, so that (24) can be applied. It yields

$$\begin{aligned}\frac{C(N,d-k)}{C(N,d)} &= \frac{\binom{N-1}{d-k-1}\left[1+\binom{N-1}{d-k-1}^{-1}\sum_{i=0}^{d-k-2}\binom{N-1}{i}\right]}{2^{N-1}-\sum_{i=0}^{N-d-1}\binom{N-1}{i}}\\ &\le \frac{\binom{N-1}{d-k-1}}{2^{N-1}}\cdot\frac{1+C_3}{1-\frac{1}{2^{N-1}}\sum_{i=0}^{N-d-1}\binom{N-1}{i}}.\end{aligned}$$

Here and below, C_m denotes a positive constant independent of d.

To estimate the last denominator, we can again use Lemma 3, since $\delta > 1/2$ and hence $2(N-d-1) < N$, if d is large enough, to get

$$\begin{aligned}
\frac{1}{2^{N-1}} \sum_{i=0}^{N-d-1} \binom{N-1}{i} &= \frac{\binom{N-1}{N-d}}{2^{N-1}} \frac{1}{\binom{N-1}{N-d}} \sum_{i=0}^{N-d-1} \binom{N-1}{i} \\
&\leq \frac{\binom{N-1}{N-d}}{2^{N-1}} \frac{N-d}{d} \cdot \frac{d+1}{-N+2d+2} \\
&\leq C_4 \frac{\binom{N-1}{d-1}}{2^{N-1}} \leq C_5 g(\delta_d)^{-N}
\end{aligned}$$

with g defined by (20). As already observed, $g(1/2) = 1$ and $g(a) > 1$ for $a \in [0,1] \setminus \{1/2\}$. Since $\delta > 1/2$, we have $g(\delta) > 1$ and hence $g(\delta_d) \geq c > 1$ for sufficiently large d, with c independent of d. It follows that $g(\delta_d)^{-N} \to 0$ as $d \to \infty$.

Finally, we observe that

$$\frac{\binom{N-1}{d-k-1}}{2^{N-1}} \leq C_6 \frac{N^N}{2^N (d-k)^{d-k} (N-d+k)^{N-d+k}} = C_7 \cdot g(\delta_d(1-\rho_d))^{-N} \to 0,$$

since $\delta_d(1-\rho_d) \to \delta(1-\rho) < 1/2$ and $g(\delta(1-\rho)) > 1$. Thus, (25) is obtained again.

To prove the first part of the theorem, we assume that $0 < \rho < \frac{1}{2}\rho_W(\delta)$; then $\rho < 1-(2\delta)^{-1}$ and $\delta(1-\rho) > 1/2$ and $\delta > 1/2$. We can write

$$\frac{C(N,d-k)}{C(N,d)} = \frac{1 - \frac{1}{2^{N-1}} \sum_{i=0}^{N-d+k-1} \binom{N-1}{i}}{1 - \frac{1}{2^{N-1}} \sum_{i=0}^{N-d-1} \binom{N-1}{i}}. \tag{26}$$

Note that by the assumed range of ρ, δ it follows that $2k + N < 2d$ and hence $2(N-d+k-1) < N$. Therefore, Lemma 3 can be applied, and we obtain

$$0 \leq \frac{1}{\binom{N-1}{N-d+k}} \sum_{i=0}^{N-d+k-1} \binom{N-1}{i} \leq \frac{N-d+k}{d-k} \cdot \frac{d-k+1}{-N+2d-2k+2} \to \frac{1-\delta(1-\rho)}{-1+2\delta(1-\rho)}.$$

Thus we find that

$$\frac{1}{2^{N-1}} \sum_{i=0}^{N-d+k-1} \binom{N-1}{i} \leq C_8 \frac{\binom{N-1}{N-d+k}}{2^{N-1}} \leq C_9 g(\delta_d(1-\rho_d))^{-N} \to 0,$$

since $\delta_d(1-\rho_d) \to \delta(1-\rho) > 1/2$ and $g(\delta(1-\rho)) > 1$. A fortiori, this implies that also

$$\frac{1}{2^{N-1}} \sum_{i=0}^{N-d-1} \binom{N-1}{i} \to 0.$$

We conclude that the quotient in (26) tends to one as $d \to \infty$, and hence $\mathbb{E}\, U_{d-k}(C_N) \to 0$. $\square$

References

[1] Bonnet, G., Chasapis, G., Grote, J., Temesvari, D., Turchi, N., Threshold phenomena for high-dimensional random polytopes. *Commun. Contemp. Math.* **21** (2019), no. 5, 1850038, 30 pp.

[2] Bonnet, G., Kabluchko, Z., Turchi, N., Phase transition for the volume of high-dimensional random polytopes. arXiv:1911.12696v1

[3] Bonnet, G., O'Reilly, E., Facets of spherical random polytopes. arXiv:1908.04033 (2019).

[4] Cover, T.M., Efron, B., Geometrical probability and random points on a hypersphere. *Ann. Math. Stat.* **38** (1967), 213–220.

[5] Donoho, D., Tanner, J., Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing. *Phil. Trans. R. Soc. A* **367** (2009), 4273–4293.

[6] Donoho, D., Tanner, J., Counting faces of randomly projected polytopes when the projection radically lowers dimension. *J. Amer. Math. Soc.* **22** (2009), 1–53.

[7] Donoho, D., Tanner, J., Counting the faces of randomly-projected hypercubes and orthants, with applications. *Discrete Comput. Geom.* **43** (2010), 522–541.

[8] Dyer, M.E., Füredi, Z., McDiarmid, C., Volumes spanned by random points in the hypercube. *Random Structures Algorithms* **3** (1992), 91–106.

[9] Gatzouras, D., Giannopoulos, A., Threshold for the volume spanned by random points with independent coordinates. *Israel J. Math.* **169** (2009), 125–153.

[10] Grünbaum, B., Grassmann angles of convex polytopes. *Acta Math.* **121** (1968), 293–302.

[11] Hug, D., Schneider, R., Random conical tessellations. *Discrete Comput. Geom.* **56** (2016), 395–462.

[12] Klar, B., Bounds on tail probabilities of discrete distributions. *Probab. Engrg. Inform. Sci.* **14** (2000), 161–171.

[13] Pivovarov, P., Volume thresholds for Gaussian and spherical random polytopes and their duals. *Studia Math.* **183** (2007), 15–34.

[14] Schneider, R., Weil, W., *Stochastic and Integral Geometry.* Springer, Berlin, 2008.

[15] Vershik, A.M., Sporyshev, P.V., Asymptotic behavior of the number of faces of random polyhedra and the neighborliness problem. *Selecta Math. Soviet.* **11**, vol. 2 (1992), 181–201.

Authors' addresses:

Daniel Hug
Karlsruhe Institute of Technology, Department of Mathematics
D-76128 Karlsruhe, Germany
E-mail: daniel.hug@kit.edu

Rolf Schneider

Mathematisches Institut, Albert-Ludwigs-Universität
D-79104 Freiburg i. Br., Germany
E-mail: rolf.schneider@math.uni-freiburg.de