

High-Dimensional Centrally-Symmetric Polytopes With Neighborliness Proportional to Dimension

David L. Donoho

January 2005

Abstract

Let A be a d by n matrix, $d < n$. Let C be the regular cross polytope (octahedron) in \mathbf{R}^n . It has recently been shown that properties of the centrosymmetric polytope $P = AC$ are of interest for finding sparse solutions to the underdetermined system of equations $y = Ax$; [9]. In particular, it is valuable to know that P is centrally k -neighborly.

We study the face numbers of randomly-projected cross-polytopes in the proportional-dimensional case where $d \sim \delta n$, where the projector A is chosen uniformly at random from the Grassmann manifold of d -dimensional orthoprojectors of \mathbf{R}^n . We derive $\rho_N(\delta) > 0$ with the property that, for any $\rho < \rho_N(\delta)$, with overwhelming probability for large d , the number of k -dimensional faces of $P = AC$ is the same as for C , for $0 \leq k \leq \rho d$. This implies that P is centrally $\lfloor \rho d \rfloor$ -neighborly, and its skeleton $Skel_{\lfloor \rho d \rfloor}(P)$ is combinatorially equivalent to $Skel_{\lfloor \rho d \rfloor}(C)$. We display graphs of ρ_N .

Two weaker notions of neighborliness are also important for understanding sparse solutions of linear equations: facial neighborliness and sectional neighborliness [9]; we study both. The weakest, (k, ϵ) -facial neighborliness, asks if the k -faces are all simplicial and if the numbers of k -dimensional faces $f_k(P) \geq f_k(C)(1 - \epsilon)$. We characterize and compute the critical proportion $\rho_F(\delta) > 0$ at which phase transition occurs in k/d . The other, (k, ϵ) -sectional neighborliness, asks whether all, except for a small fraction ϵ , of the k -dimensional intrinsic sections of P are k -dimensional cross-polytopes. (Intrinsic sections intersect P with k -dimensional subspaces spanned by vertices of P .) We characterize and compute a proportion $\rho_S(\delta) > 0$ guaranteeing this property for $k/d \sim \rho < \rho_S(\delta)$. We display graphs of ρ_S and ρ_F .

Key Words and Phrases: Centrosymmetric Polytopes. Neighborly Polytopes. Cross-Polytope. Randomly-Projected Polytope. Grassmann Angle. Internal Angles of Polytopes. External Angles of Polytopes. Large Deviations of Half-Normal random variables. Octahedron. Quotient spaces of ℓ^1 .

Acknowledgements. Partial support from NSF DMS 00-77261, and 01-40698 (FRG), ONR-MURI, and NIH. Thanks to the Institute for Pure and Applied Mathematics (IPAM) and its ‘neighborly’ hospitality during the program ‘Analysis and Geometry in High Dimensions’ in the Fall of 2004, while this was prepared. Thanks to G.R. Burton and R. Schneider for help with references and Amir Dembo for discussions about large deviations.

1 Introduction

1.1 Neighborliness and Central-Neighborliness

In the classical theory of convex polytopes, the notion of neighborliness offers a beautiful glimpse of the surprises of high dimensions. Neighborliness asks if every $k+1$ vertices of a polytope span a k -face. In low dimensions this is difficult for beginners to arrange – outside the trivial case of the simplex – because it seems that some candidate edges easily get ‘swallowed up’ crossing ‘inside’ the polytope. It can be surprising to students that in higher dimensions $d > 3$ this can be managed easily, by simply taking $n > d$ points $x_i = M(t_i)$ along the moment curve $M(t) = (1, t, t^2, \dots, t^{d-1})$ [10, 12]. The convex hull of these points is a polytope with n vertices which are $\lfloor d/2 \rfloor$ neighborly, for each $n > d$; and this is the maximum possible value. See eg. [12, Chapter 7] for more.

For centrosymmetric polytopes, a modified notion of neighborliness is needed; one asks if every $k+1$ vertices *not including an antipodal pair* span a k -face. The slight modification detracts a bit from the beauty of the notion; and perhaps also from the interest in studying it. There is no known general construction of centrally k -neighborly for large n and d , and the achievable upper bound is smaller: $k \leq \lfloor (d+1)/3 \rfloor$, according to McMullen and Shephard [14]. For n not much larger than d , Schneider [17] showed the existence of centrally-symmetric polytopes which are k -centrally-neighborly for $k \approx .2309d$; however Schneider’s polytopes have only $2d(1 + o(1))$ vertices. Burton [4] showed that for fixed d and large enough n , even 2-central-neighborliness is impossible. Not much else seems to have been published.

1.2 Central Neighborliness and Optimization

In a companion paper [9], the author shows that central-neighborliness of centrally-symmetric polytopes is important for understanding solvability of certain combinatorial optimization problems by convex relaxation. Specifically, suppose A is a d -by- n matrix with $d < n$ and we are interested in finding the solution to the underdetermined system $y = Ax$ having *fewest nonzeros*. Although this problem is NP-hard in general, the sparsest solution can be often found by solving the *convex* optimization problem $\min \|x\|_1$ subject to $y = Ax$. The conditions on A and y guaranteeing success are: first, that a solution with at most k nonzeros exist; and secondly, that the convex polytope $P = AC$ be centrally k -neighborly. Here C denotes the cross-polytope (ℓ^1 ball) in \mathbf{R}^n .

The relation to optimization brings new significance into the study of neighborliness in the centrosymmetric case. As [9] shows, we can interpret recent results in the study of sparse solutions by ℓ^1 optimization as constructions of centrosymmetric polytopes which are neighborly for reasonably large k . For example, a result of the author [8] relying on Banach space geometry techniques implies that for large d and n , d proportional to n , if we randomly take points x_1, \dots, x_n from the uniform distribution on the unit sphere in \mathbf{R}^d , then the centrosymmetric polytope generated by taking the convex hull of these points and their antipodes is overwhelmingly likely to be k -neighborly, for $k \sim \rho d$. Here ρ is a positive constant depending on n/d ; until now little was known about the possible values for ρ .

Clearly, we would like to know more about the possible/prevalent ranges of neighborliness.

1.3 Analysis in High Dimensions

In this paper we adopt the high-dimensional viewpoint, and construct polytopes by projecting from n dimensions down to d dimensions, n large, d proportional to n . The resulting families

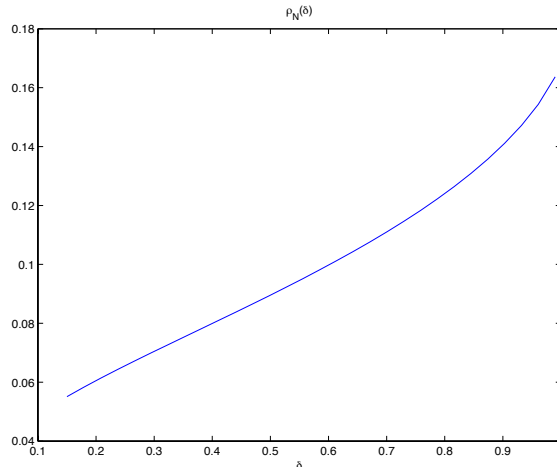


Figure 1.1: The lower bound $\rho_N(\delta)$ on the neighborliness threshold, computed by methods of this paper. Matlab software available from the author.

of high-dimensional centrosymmetric polytopes are *proportionally-neighborly*, in the sense that for some $\rho > 0$ and large d , they are typically ρd -centrally neighborly. Our approach gives quantitative information about the size of ρ achievable. We present numerical evidence that $k \geq .089d$ when $n = 2d$ and n is large.

Our analysis considers the ensemble of polytopes $P = AC$ where A is a random projection from \mathbf{R}^d to \mathbf{R}^n and $C = C^n$ is the standard crosspolytope. We study a function $\rho_N : (0, 1] \mapsto [0, 1]$, depicted in Figure 1.1 and defined in detail in later sections. ρ_N provides a lower bound on the proportional central-neighborliness of the random polytope P .

Corollary 1.1 *Let A be a uniform random projection from \mathbf{R}^n to \mathbf{R}^d with $d = \lfloor \delta n \rfloor$. Fix $\epsilon > 0$. With overwhelming probability for large n , $P = AC$ is centrally k -neighborly with $k/d \geq \rho_N(\delta) - \epsilon$.*

1.4 Face Numbers

In fact, this article does not much discuss neighborliness *per se*. Instead, we consider the properties of face numbers of the projected cross-polytope, getting the following result:

Theorem 1 *Let $\rho < \rho_N(\delta)$ and let $A = A_{d,n}$ be a uniformly-distributed random projection from \mathbf{R}^n to \mathbf{R}^d , with $d \geq \delta n$. Then*

$$\text{Prob}\{f_\ell(AC) = f_\ell(C), \quad \ell = 0, \dots, \lfloor \rho d \rfloor\} \rightarrow 1, \quad \text{as } n \rightarrow \infty. \quad (1.1)$$

Central k -neighborliness follows from this equality of face numbers; see Section 2 below.

Our proof of Theorem 1 starts from work of Böröczky and Henk [2], who considered face numbers of the randomly projected cross-polytope with d fixed and $n \rightarrow \infty$. We modify the analysis, letting d and n both go to infinity in a proportional way. The approach of [2] depends on the framework for computing Grassmann angles of a polytope due to Affentranger and Schneider [1] and Vershik and Sporyshev [19]. This uses exact analytical work in integral geometry of convex sets by McMullen [13] (nonlinear sum/angle relations), Grünbaum [11] (Grassmann Angles), and Harold Ruben [15] (volumes of spherical simplices).

Our approach is to develop formulas for the internal and external angles of cross-polytope faces in the n -proportional-to- d setting, obtaining inequalities of a substantially different form than in the d -fixed setting. We use these inequalities to characterize and compute $\rho_N(\delta)$.

The study of face numbers in the proportional-dimensional case, where $d \sim \delta n$, was pioneered by Vershik and Sporyshev [19] in the ‘projection of simplex’ case $P = AT^n$, with T^n the regular simplex in \mathbf{R}^n . Most importantly, Vershik and Sporyshev [19] developed, in addition to the proportional-to-dimension viewpoint, several analytical tools relevant to the proportional-dimensional case, for studying internal and external angles of simplices; these are also used here.

1.5 Weaker Notions of Neighborliness

Vershik and Sporyshev [19] were interested in the question of whether, for k in a fixed proportion to n , the face numbers $f_k(AT^n) = f_k(T^n)(1 + o(1))$ or not. The answer obeyed a threshold phenomenon for k in the vicinity of $\rho_T d$, for some implicitly characterized $\rho_T = \rho_T(d/n, k/n)$.

For comparison to Theorem 1, note that the question of *approximate* equality of face numbers $f_k(AT^n) = f_k(T^n)(1 + o(1))$ is weaker than the *exact* equality studied here in Theorem 1; it changes at a different threshold in k/d . The comparable question in our setting is approximate equality of face numbers $f_k(AC^n) = f_k(C^n)(1 + o(1))$. Figure 1.2 displays thresholds computed based on the following result.

Theorem 2 *There is a function $\rho_F(\delta)$, characterised below, with the following property. Let $d = d(n) \sim \delta n$ and let $A = A_{d,n}$ be a uniform random projection from \mathbf{R}^n to \mathbf{R}^d . Then for a sequence $k = k(n)$ with $k/d \sim \rho$, $\rho < \rho_F(\delta)$, we have*

$$f_k(AC) = f_k(C)(1 + o_P(1)). \tag{1.2}$$

This result is sharp in the sense that for sequences with $k/d \sim \rho > \rho_F$, we do not have the approximate equality (1.2); but we do not prove this here. Thus, we distinguish between ρ_F which is really a threshold and ρ_N which is a lower bound on a threshold.

As explained in [9], (1.2) can itself be justified as a weak kind of neighborliness – *facial* neighborliness – in which the overwhelming majority of (rather than all) k -tuples span $(k - 1)$ -faces. This notion of neighborliness is easier to satisfy than orthodox central neighborliness and so $\rho_F > \rho_N$. [9] also defines a notion of *sectional* neighborliness, intermediate between facial and central neighborliness. In this notion, we take any k vertices not including an antipodal pair and section P by the linear subspace spanned by those vertices. If the overwhelming majority of such sections are k -dimensional cross-polytopes, we say that P is typically *sectionally k -neighborly*. In Figure 1.2 we also display a bound on the sectional neighborliness of quotient polytopes, based on the following result.

Theorem 3 *There is a function $\rho_S(\delta)$, characterised below, with the following property. Let $\rho < \rho_S(\delta)$ and let A be a uniform random projection from \mathbf{R}^n to \mathbf{R}^d , with $d \geq \delta n$. Then for $k \sim \rho d$, we have with overwhelming probability for large d that $P = AC$ is typically *sectionally k -neighborly*.*

All three theorems are proved in more or less the same way; we spend the bulk of this article on the proof Theorem 1 and in a final section indicate the changes needed to prove Theorems 2-3.

Figure 1.2 depicts substantial numerical differences in the critical proportion ρ_F and the lower bounds ρ_N and ρ_S . The most striking differences between ρ_F and the other two proportions

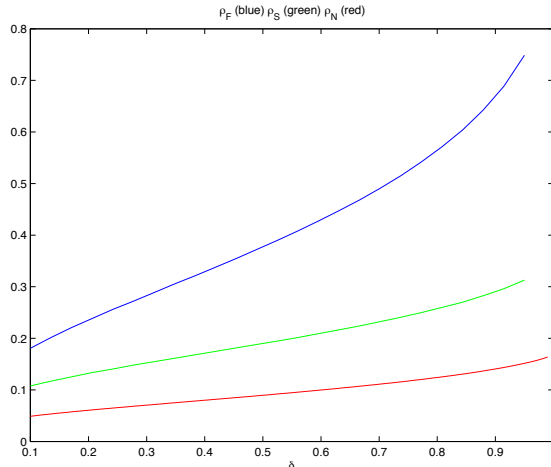


Figure 1.2: The threshold $\rho_F(\delta)$ for approximate equality of ρd -dimensional face numbers of C and AC (blue), and the lower bound $\rho_S(\delta)$ for sectional neighborliness (green). Plot of ρ_N overlaid in red for comparison

are that ρ_F crosses the line $\rho = 1/2$ near $\delta = .701$ and increases to 1 as $\delta \rightarrow 1$. The Appendix proves the following.

Theorem 4

$$\lim_{\delta \rightarrow 1} \rho_F(\delta) = 1. \tag{1.3}$$

For some $\delta_0 \in (0, 1)$,

$$\rho_F(\delta) > 1/2, \quad \delta_0 < \delta < 1. \tag{1.4}$$

For comparison, one can compute that

$$.168 \approx \lim_{\delta \rightarrow 1} \rho_N(\delta), \tag{1.5}$$

and

$$.352 \approx \lim_{\delta \rightarrow 1} \rho_S(\delta). \tag{1.6}$$

Such features can be important from the applications viewpoint, where they can be interpreted as saying that average case behavior is far more favorable than worst-case behavior. See the discussion in [9].

2 Neighborliness and Face Numbers

We first justify our claim that face numbers of the quotient polytope alone are enough to determine neighborliness.

We also fix notation concerning convex polytopes; see [12] for more details. In discussing the (closed, convex) polytope P we commonly refer to its vertices $v \in \text{vert}(P)$ and k -dimensional faces $F \in \mathcal{F}_k(P)$. $v \in P$ will be called a vertex of P if there is a linear functional λ_v separating v from $P \setminus \{v\}$, i.e. a value c so that $\lambda_v(v) = c$ and $\lambda_v(x) < c$ for $x \in P, x \neq v$. We write conv for the convex hull operation; thus $P = \text{conv}(\text{vert}(P))$. Vertices are just 0-dimensional faces, and

a k -dimensional face is a set F for which there exists a separating linear functional λ_F , so that $\lambda_F(x) = c$, $x \in F$ and $\lambda_F(x) < c$, $x \notin F$. Faces are convex polytopes, each one representable as the convex hull of a subset $\text{vert}(F) \subset \text{vert}(P)$; thus if F is a face, $F = \text{conv}(\text{vert}(F))$. A k -dimensional face will be called a k -simplex if it has $k + 1$ vertices.

Lemma 2.1 *Let A be an arbitrary linear transformation. Let $P = AC$ have the same face numbers as C , up to dimension $k - 1$:*

$$f_\ell(P) = f_\ell(C), \ell = 0, 1, \dots, k - 1.$$

Then

- All the ℓ -faces of P are ℓ -simplices, for $\ell = 0, \dots, k - 1$.
- P is centrally k -neighborly.

Proof. We first note the very elementary:

$$\text{vert}(P) \subset A \text{vert}(C).$$

Indeed, every element of C is a convex combination of its vertices. Every element of P is the image under A of such a convex combination and hence is a convex combination of the signed columns of A . Hence the vertices of P are among the signed columns of A , and

$$f_0(P) \leq f_0(C). \tag{2.7}$$

Since $f_0(P) = f_0(C)$, we conclude

$$\text{vert}(P) = A \text{vert}(C).$$

Thus the vertices of P are made of n antipodal pairs. No antipodal pair can be an edge of P if $n > 1$, because the origin 0 serves as the common midpoint of all line segments connecting antipodes. Avoiding such pairs forces $f_1(P) \leq 4\binom{n}{2}$. Now $f_1(C) = 4\binom{n}{2}$. Hence, the hypothesis $f_1(P) = f_1(C)$ implies that $\mathcal{F}_1(P)$ contains every possible edge formed from the vertex set which does not connect antipodal vertices. But this means P is centrally-2-neighborly.

Consider now a 2-face $F \in \mathcal{F}_2(P)$. We will show that it is simplicial. We have $\text{vert}F \subset \text{vert}P$. Also, such a 2-face of F cannot contain an antipodal pair of vertices from P . Hence, every pair of vertices of F , being a non-antipodal pair of vertices in P , generates an edge in $\mathcal{F}_1(P)$, and hence in $\mathcal{F}_1(F)$. It follows that every 2-face F is 2-neighborly, in the stronger sense of neighborliness appropriate to asymmetric sets – i.e. without any proviso about avoiding antipodes (because there are no antipodal pairs in F to avoid!). We now invoke Theorem 4, Chapter 7 of Grünbaum [12]; *a d -neighborly d -polytope is a d -simplex*. Hence all 2-faces are 2-simplices.

Now if all 2-faces are 2-simplices, and no such face can contain an antipodal pair, there are at most $8\binom{n}{3}$ such faces. But $f_2(C) = 8\binom{n}{3}$. Hence $f_2(P) = f_2(C)$ implies that all allowable combinations of 3 vertices generate faces. So P is centrally-3-neighborly.

We continue in this way to higher dimensional faces. Each ℓ -face F contains no antipodal pairs; by previous steps, all subsets of ℓ vertices span faces of P , and therefore of F , and so F is ℓ -neighborly, and therefore an ℓ -simplex. The hypothesis $f_\ell(P) = f_\ell(C)$ implies that all allowable combinations of $\ell + 1$ vertices not containing an antipodal pair generate faces of P , and so P is $\ell + 1$ -centrally-neighborly.

We continue through stage $k - 1$, and the lemma is proved. □

3 Random Projections of Cross-Polytopes

We now outline the proof of Theorem 1. Key lemmas and inequalities will be justified in later sections.

3.1 Angle Sums

As remarked in the introduction, our proof proceeds by refining a line of research in convex integral geometry. Affentranger and Schneider [1] (see also Vershik and Sporyshev [19] and Böröczky and Henk [2]) studied the properties of random projections $R = AQ$ where Q is an n -polytope and R is its d -dimensional orthogonal projection. [1] derived the formula

$$Ef_k(R) = f_k(Q) - 2 \sum_{s \geq 0} \sum_{F \in \mathcal{F}_k(Q)} \sum_{G \in \mathcal{F}_{d+1+2s}(Q)} \beta(F, G) \gamma(G, Q);$$

where E denotes the expectation over realizations of the random orthogonal projection, and the sum is over pairs (F, G) where F is a face of G . In this display, $\beta(F, G)$ is the internal angle at face F of G and $\gamma(G, Q)$ is the external angle of Q at face G ; for definitions of these terms see eg. Grünbaum, Chapter 14.

The slogan underlying the formula is that each face $F \in \mathcal{F}_k(Q)$ will either ‘survive’ under projection, so that AF is a k -face of R , or it will get ‘swallowed up’ inside R . The expected number of faces in R is thus the number of faces in Q minus the expected number faces ‘swallowed up’ in projection. The chance of a particular face’s getting ‘swallowed up’, is exactly the chance that the subspace spanned by columns of A^t in \mathbf{R}^n intersects trivially with the cone of separating linear functionals associated to face $F \in Q$. The chance that a uniform random subspace hits a cone is precisely the so-called Grassmann angle as defined by Grünbaum [11]. Hence the expected number of faces $f_k(R)$ involves a sum of Grassman angles, one for each k -face F of Q , evaluating the probability that AF is a k -face of R . McMullen [13] developed nonlinear angle-sum relations which are used to decompose these Grassmann Angles into the above sums involving internal and external angles.

Specializing to the case where $Q = C$, the n -dimensional Cross-Polytope, we write

$$Ef_k(P) = f_k(C) - \Delta(k, d, n) \tag{3.8}$$

with

$$\Delta(k, d, n) = 2 \sum_{s \geq 0} \sum_{F \in \mathcal{F}_k(C)} \sum_{G \in \mathcal{F}_{d+1+2s}(C)} \beta(F, G) \gamma(G, C). \tag{3.9}$$

3.2 Exact Equality from Expectation

Because of (2.7) we view (3.8) as showing that on average $f_k(P)$ is about the same as $f_k(C)$, except for a nonnegative ‘discrepancy’ Δ . We will show that under the stated conditions on k, d , and n , for some $\epsilon > 0$

$$\Delta(k, d, n) \leq n \exp(-n\epsilon). \tag{3.10}$$

Now as $f_k(P) \leq f_k(C)$,

$$Prob\{f_k(P) \neq f_k(C)\} \leq E(f_k(C) - f_k(P)) = \Delta(k, d, n).$$

Hence (3.10) implies that with overwhelming probability we get equality of $f_k(P)$ with $f_k(C)$, as claimed in the theorem. To extend this into the needed simultaneous result - that $f_\ell(P) = f_\ell(C)$,

$\ell = 0, \dots, k-1$ – one defines events $E_k = \{f_k(P) \neq f_k(C)\}$ and notes that by Boole's inequality

$$\text{Prob}(\cup_0^{k-1} E_\ell) \leq \sum_0^{k-1} \text{Prob}(E_k) \leq \sum_{\ell=0}^{k-1} \Delta(\ell, d, n).$$

The exponential decay of $\Delta(k, d, n)$ will guarantee that the sum converges to 0 whenever the $k-1$ -th term does. Hence by establishing (3.10) we get

$$\text{Prob}\{f_\ell(P) = f_\ell(C), \quad \ell = 0, \dots, k-1\} \rightarrow 1$$

as is to be proved.

To establish (3.10), we rewrite (3.9) as

$$\Delta(k, d, n) = \sum_{s \geq 0} D_s$$

where, for $\ell = d+1+2s$, $s = 0, 1, 2, \dots$

$$D_s = 2 \cdot \sum_{F \in \mathcal{F}_k(C)} \sum_{G \in \mathcal{F}_{d+1+2s}(C)} \beta(F, G) \gamma(G, C).$$

We will show that, for $\rho < \rho_N$ (still to be defined) and for sufficiently small $\epsilon > 0$, then for $n > n_0(\epsilon; \rho, \delta)$

$$n^{-1} \log(D_s) \leq -\epsilon.$$

This implies (3.10) and hence our main result follows.

3.3 Decay and Growth Exponents

Böröczky and Henk [2] studied exactly the setting $P = AC$ with C the cross-polytope - though for a different range of k, d, n (they considered k, d fixed and $n \rightarrow \infty$), and also used a different formula for $Ef_k(P)$, so they did not directly study the term $\Delta(k, d, n)$. They did, however, make the following useful observations.

- There are $2^{k+1} \binom{n}{k+1}$ k -faces of C .
- For $\ell > k$, there are $2^{\ell-k} \binom{n-k-1}{\ell-k}$ ℓ -faces of C containing a given k -face of C .
- The faces of C are all simplices, and the internal angle $\beta(F, G) = \beta(T^k, T^\ell)$, where T^d denotes the standard d -simplex.
- The external angle $\gamma(G^\ell, C^n)$ is the same for all ℓ -faces of C ; it has a closed form integral expression very similar to $\gamma(T^\ell, T^n)$.

Thus we can write

$$\begin{aligned} D_s &= 2 \cdot 2^\ell \cdot \binom{n}{k+1} \binom{n-k-1}{\ell-k} \beta(T^k, T^\ell) \gamma(F^\ell, C) \\ &= C_s \beta(T^k, T^\ell) \gamma(F^\ell, C), \end{aligned}$$

say, with C_s the combinatorial prefactor.

We plan now to estimate $n^{-1} \log(D_s)$, decomposing it into a sum of terms involving logarithms of the combinatorial prefactor, the internal angle and the external angle. Define the Shannon entropy:

$$H(p) = p \log(1/p) + (1-p) \log(1/(1-p));$$

noting that here the logarithm base is e , rather than the customary base 2. As did Vershik and Sporyshev [19], we also remark that

$$n^{-1} \log \binom{n}{\lfloor pn \rfloor} \rightarrow H(p), \quad p \in [0, 1], \quad n \rightarrow \infty \quad (3.11)$$

so this provides a convenient summary for combinatorial terms. Defining $\nu = \ell/n \geq \delta$, we have

$$n^{-1} \log(C_s) = \nu \log_e(2) + H(\rho\delta) + H\left(\frac{\nu - \rho\delta}{1 - \rho\delta}\right)(1 - \rho\delta) + R_1 \quad (3.12)$$

with remainder $R_1 = R_1(s, k, d, n)$. Define then the *growth* exponent

$$\Psi_{com}(\nu; \rho, \delta) \equiv \nu \log_e(2) + H(\rho\delta) + H\left(\frac{\nu - \rho\delta}{1 - \rho\delta}\right)(1 - \rho\delta),$$

describing the exponential growth of the combinatorial factors. It is banal to apply (3.11) and see that the remainder R_1 in (3.12) is $o(1)$ uniformly in the range $k - \ell > (\delta - \rho)n$, $n > n_0$.

Section 4.1 below defines a so-called *decay* exponent $\Psi_{ext}(\nu)$. Section 5 shows that $\gamma(F^\ell, C^n)$ decays exponentially at least at the rate $\Psi_{ext}(\nu)$; for each $\epsilon > 0$,

$$n^{-1} \log(\gamma(F^\ell, C)) \leq -\Psi_{ext}(\nu) + \epsilon,$$

uniformly in $\ell \geq \delta n$, $n \geq n_0(\delta, \epsilon)$. The graph of Ψ_{ext} is depicted in Figure 4.1.

Similarly, Section 4.2 below defines a decay exponent $\Psi_{int}(\nu; \rho\delta)$. Section 6 below shows that the internal angle $\beta(T^k, T^\ell)$ indeed decays with this exponent; along sequences $k \sim \rho\delta n$, $\ell \sim \nu n$,

$$n^{-1} \log(\beta(T^k, T^\ell)) = -\Psi_{int}(\nu; \rho\delta) + R_2,$$

where the remainder $R_2 \leq o(1)$ uniformly in $k - \ell \geq (\delta - \rho)n$.

Hence for any fixed choice of ρ, δ , for $\epsilon > 0$, and for $n \geq n_0(\rho, \delta, \epsilon)$ we have the inequality

$$n^{-1} \log(D_s) \leq \Psi_{com}(\nu; \rho, \delta) - \Psi_{int}(\nu; \rho\delta) - \Psi_{ext}(\nu) + 3\epsilon, \quad (3.13)$$

valid uniformly in s .

3.4 Defining ρ_N

Define now the *net exponent* $\Psi_{net}(\nu; \rho, \delta) = \Psi_{com}(\nu; \rho, \delta) - \Psi_{int}(\nu; \rho\delta) - \Psi_{ext}(\nu)$. We can define at last the mysterious ρ_N as the threshold where the net exponent changes sign. We will see that the components of Ψ_{net} are all continuous over sets $\{\rho \in [\rho_0, 1], \delta \in [\delta_0, 1], \nu \in [\delta, 1]\}$, and so Ψ_{net} has the same continuity properties.

Definition 1 *Let $\delta \in (0, 1]$. The critical proportion $\rho_N(\delta)$ is the supremum of $\rho \in [0, 1]$ obeying*

$$\Psi_{net}(\nu; \rho, \delta) < 0, \quad \nu \in [\delta, 1].$$

Continuity of Ψ_{net} shows that if $\rho < \rho_N$ then, for some $\epsilon > 0$,

$$\Psi_{net}(\nu; \rho, \delta) < -4\epsilon, \quad \nu \in [\delta, 1].$$

Combine this with (3.13). Then for all $s = 0, 2, \dots, (n-d)/2$ and all $n > n_0(\delta, \rho, \epsilon)$

$$n^{-1} \log(D_s) \leq -\epsilon.$$

This implies (3.10) and our main result follows.

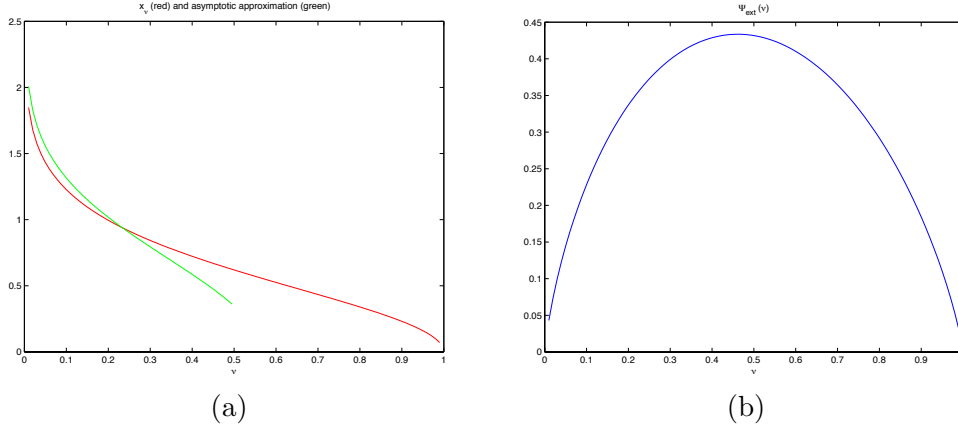


Figure 4.1: Panel (a): The minimizer x_ν of ψ_ν , as a function of ν (red) and the asymptotic approximation $\sqrt{\log(\frac{1}{\sqrt{\pi\nu}})}$ (green); Panel (b): The exponent Ψ_{ext} , a function of ν .

4 Properties of Exponents

We now define the exponents Ψ_{int} and Ψ_{ext} and discuss properties of ρ_N .

4.1 Exponent for External Angle

Let G denote the cumulative distribution function of a half-normal $HN(0, 1/2)$ random variable, i.e. a random variable $X = |Z|$ where $Z \sim N(0, 1/2)$, and $G(x) = \text{Prob}\{X \leq x\}$. It has density $g(x) = 2/\sqrt{\pi} \exp(-x^2)$. Writing this out,

$$G(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2} dy; \quad (4.1)$$

so G is just the classical error function *erf*. For $\nu \in (0, 1]$, define x_ν as the solution of

$$\frac{2xG(x)}{g(x)} = \frac{1-\nu}{\nu}. \quad (4.2)$$

Since xG is a smooth strictly increasing function ~ 0 as $x \rightarrow 0$ and $\sim x$ as $x \rightarrow \infty$, and $g(x)$ is strictly decreasing, the function $2xG(x)/g(x)$ is one-one on the positive axis, and x_ν is well-defined, and a smooth, decreasing function of ν . See Figure 4.1 for a depiction.

This has limiting behavior $x_\nu \rightarrow 0$ as $\nu \rightarrow 1$ and $x_\nu \sim \sqrt{\log((1-\nu)/(2\nu))}$ as $\nu \rightarrow 0$. Define now

$$\Psi_{ext}(\nu) = -(1-\nu) \log(G(x_\nu)) + \nu x_\nu^2.$$

This is depicted in Figure 4.1.

This function is smooth on the interior of $(0, 1)$ and concave, with endpoints $\Psi_{ext}(1) = 0$, $\Psi_{ext}(0) = 0$. A useful fine point is the asymptotic

$$\Psi_{ext}(\nu) \sim \nu \log\left(\frac{1}{\nu}\right) - \frac{1}{2}\nu \log\left(\log\left(\frac{1}{\nu}\right)\right) + O(\nu), \quad \nu \rightarrow 0. \quad (4.3)$$

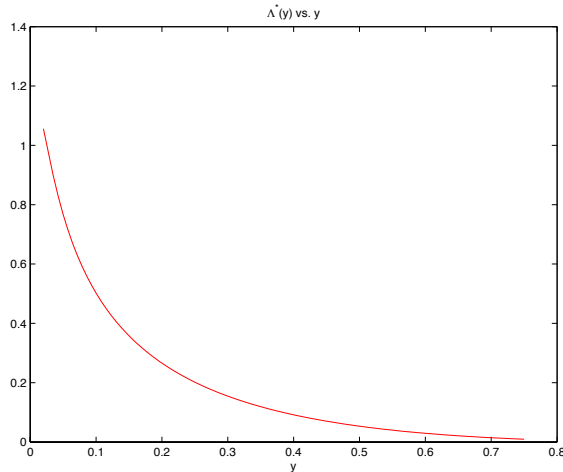


Figure 4.2: $\Lambda^*(y)$, rate function for Half-normal distribution; only the ‘left-half’ $0 < y < \mu$ is depicted. The function diverges at 0.

4.2 Exponent for Internal Angle

Let Y be a standard half-normal random variable $HN(0, 1)$; this has cumulant generating function $\Lambda(s) = \log(E \exp(sY))$. Very convenient for us is the exact formula

$$\Lambda(s) = s^2/2 + \log(2\Phi(s)),$$

where Φ is the usual cumulative distribution function of a standard Normal $N(0, 1)$. The cumulant generating function Λ has a rate function (Fenchel-Legendre dual [6])

$$\Lambda^*(y) = \max_s sy - \Lambda(s).$$

This is smooth and convex on $(0, \infty)$, strictly positive except at $\mu = EY = \sqrt{2/\pi}$. More details are provided in Section 6. See Figure 4.2.

For $\gamma \in (0, 1)$ let

$$\xi_\gamma(y) = \frac{1-\gamma}{\gamma} y^2/2 + \Lambda^*(y).$$

The function $\xi_\gamma(y)$ is strictly convex and positive on $(0, \infty)$ and has a minimum at a unique y_γ in the interval $(0, \sqrt{2/\pi})$. We define, for $\gamma = \frac{\rho\delta}{\nu} \leq \rho$,

$$\Psi_{int}(\nu; \rho\delta) = \xi_\gamma(y_\gamma)(\nu - \rho\delta) + \log(2)(\nu - \rho\delta).$$

This is depicted in Figure 4.3. For fixed ρ, δ , Ψ_{int} is continuous in $\nu \geq \delta$. Most importantly, in Section 6.4 below we get the asymptotic formula

$$\xi_\gamma(y_\gamma) \sim \frac{1}{2} \cdot \log\left(\frac{1-\gamma}{\gamma}\right), \quad \gamma \rightarrow 0. \quad (4.4)$$

Since $\gamma = \rho\delta/\nu \leq \rho$, (4.4) implies that for given $\eta > 0$ and small ρ ,

$$\Psi_{int}(\nu; \rho\delta) \geq \left(\frac{1}{2} \cdot \log\left(\frac{1-\rho}{\rho}\right)(1-\eta) + \log(2) \right) (\nu - \rho\delta), \quad \nu \in [\delta, 1]. \quad (4.5)$$

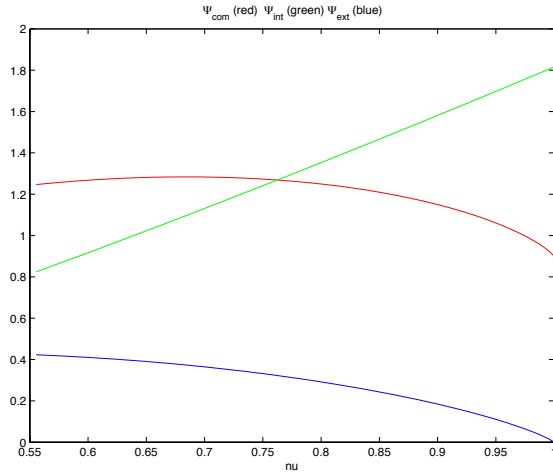


Figure 4.3: The exponents $\Psi_{com}(\nu; \rho, \delta)$ (red) and $\Psi_{int}(\nu; \rho\delta)$ (green), for $\rho = .095$, $\delta = .5555$. For comparison, Ψ_{ext} is displayed in blue.

4.3 Combining the Exponents

We now consider the combined behavior of Ψ_{com} , Ψ_{int} and Ψ_{ext} . We think of these as functions of ν with ρ, δ as parameters. The combinatorial exponent Ψ_{com} is the sum of a linear function in ν , and a scaled, shifted version of the Shannon entropy, which is a symmetric, roughly parabolic shaped function. This is the exponent of a growing function which must be outweighed by the sum $\Psi_{ext} + \Psi_{int}$. It is depicted in Figure 4.3.

Figure 4.4 shows both Ψ_{com} and $\Psi_{ext} + \Psi_{int}$ with $\delta = .5555$ and $\rho = .095$. The desired condition $\Psi_{net} < 0$ is the same as $\Psi_{com} < \Psi_{ext} + \Psi_{int}$, and this is distinctly obeyed except near $\nu = \delta$, where the two curves are close. We have $\rho_N(\delta) \approx .095$.

4.4 Properties of ρ_N

The asymptotic relations (4.5) and (4.3) allow us to see two key facts about ρ_N , both proved in the Appendix. Firstly, the concept is nontrivial:

Lemma 4.1

$$\rho_N(\delta) > 0, \quad \delta \in (0, 1). \quad (4.6)$$

This result was to be expected. Exploiting [9] and [8, 5] it could have previously been inferred that, for some $\rho = \rho(\delta) > 0$ such random polytopes are, with overwhelming probability, ρd -neighborly. Effectively, (4.6) shows that the techniques of this paper are at least as strong as those of [8, 5].

Secondly, one can show that, although $\rho_N(\delta) \rightarrow 0$ as $\delta \rightarrow 0$, it goes to zero slowly. We prove the following in the appendix.

Lemma 4.2 For $\eta > 0$,

$$\rho_N(\delta) \geq \log(1/\delta)^{-(1+\eta)}, \quad \delta \rightarrow 0.$$

Again this result could have been anticipated. Exploiting the connection with studies of ℓ^1 optimization [9], we could have inferred from [7, 5], that in the case $n \sim d^\gamma$, $\gamma > 1$, AC is centrally k -neighborly with $k \geq cd/\log(n)$.

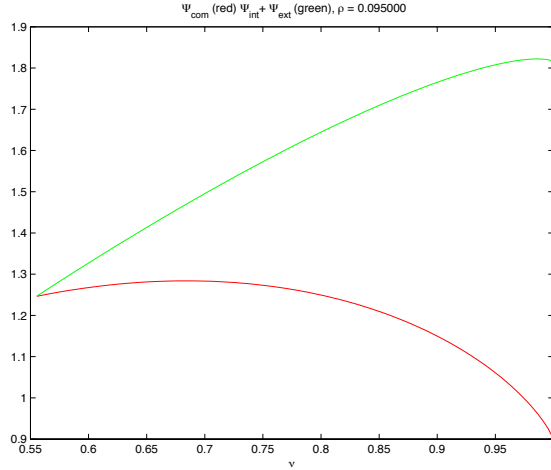


Figure 4.4: The exponents $\Psi_{com}(\nu; \rho, \delta)$ and $\Psi_{int}(\nu; \rho\delta) + \Psi_{ext}(\nu)$, for $\rho = .095$, $\delta = .5555$. The graph of Ψ_{com} (red) falls below that of $\Psi_{int} + \Psi_{ext}$ (green) and so $\Psi_{net} < 0$.

5 Bounds on the External Angle

We now justify the use of Ψ_{ext} .

Lemma 5.1 *Fix $\delta, \epsilon > 0$.*

$$n^{-1} \log(\gamma(F^\ell, C^n)) \leq -\Psi_{ext}(\ell/n) + \epsilon, \quad (5.1)$$

uniformly in $\ell \geq \delta n$, $n \geq n_0(\delta, \epsilon)$.

We start from an exact identity. Böröczky and Henk [2], building on work of Vershik and Sporyshev [18] and ultimately of H. Ruben [15], give the integral formula

$$\gamma(F^\ell, C) = \sqrt{\frac{\ell+1}{\pi}} \int_0^\infty e^{-(\ell+1)x^2} \left(\frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2} dy \right)^{n-\ell-1} dx.$$

We recognize the term in braces as the error function G from (4.1). Set $\nu_{\ell,n} = (\ell+1)/n$. The integral formula can be rewritten as

$$\sqrt{\frac{n\nu_{\ell,n}}{\pi}} \int_0^\infty \exp\{-n\nu_{\ell,n}x^2 + n(1-\nu_{\ell,n}) \log G(x)\} dx. \quad (5.2)$$

The appearance of n in the exponent suggests to use Laplace's method; we define, for ν fixed,

$$f_{\nu,n}(y) = \exp\{-n\psi_\nu(y)\} \cdot \sqrt{\frac{n\nu}{\pi}}$$

with

$$\psi_\nu(y) \equiv \nu y^2 - (1-\nu) \log G(y).$$

We note that ψ_ν is smooth and convex and (in the appendix) develop expressions for its second and third derivatives. Applying Laplace's method to ψ_ν in the usual way, but taking care about regularity conditions and remainders, gives a result with the uniformity in ν , which is crucial for us.

Lemma 5.2 For $\nu \in (0, 1)$ let x_ν denote the minimizer of ψ_ν . Then

$$\int_0^\infty f_{\nu,n}(x) dx \leq \exp(-n\psi_\nu(x_\nu))(1 + R_n(\nu))$$

where, for $\delta, \eta > 0$,

$$\sup_{\nu \in [\delta, 1-\eta]} R_n(\nu) = o(1) \text{ as } n \rightarrow \infty.$$

Of course the minimizer x_ν mentioned in this lemma is the same x_ν defined earlier in (4.2) in terms of the error function, and that the minimum value identified in this Lemma as driving the exponential rate is the same as our exponent Ψ_{ext} :

$$\Psi_{ext}(\nu) = \psi_\nu(x_\nu). \quad (5.3)$$

In fact Lemma 5.2 easily leads to Lemma 5.1. We first note that $\Psi_{ext}(\nu) \rightarrow 0$ as $\nu \rightarrow 1$. For given $\epsilon > 0$ in the statement of the Lemma, there is (a largest) $\nu_\epsilon < 1$ with $\Psi_{ext}(\nu_\epsilon) = \epsilon$. Note that $\gamma(F^\ell, C) \leq 1$, so that for $\ell > \nu_\epsilon n$,

$$n^{-1} \log(\gamma) \leq 0 < -\Psi_{ext}(\nu) + \epsilon,$$

for $n \geq 1$. Consider now $\ell \in [\delta n, \nu_\epsilon n)$. Taking into account (5.2), we now have

$$\gamma(F^\ell, C) = \int_0^\infty f_{\nu_\ell, n}(x) dx.$$

Applying the uniformity in ν given in Lemma 5.2, we conclude

$$n^{-1} \log(\gamma(F^\ell, C)) = \psi_{\nu_\ell, n}(x_{\nu_\ell, n}) + o(1), \quad \ell \geq \delta n.$$

Then invoking the identity (5.3) and the uniform continuity of ψ_ν in x and of x_ν in $\nu \in [\delta, 1]$, we get

$$n^{-1} \log(\gamma(F^\ell, C)) \leq -\Psi_{ext}(\ell/n) + o(1).$$

Lemma 5.1 follows.

6 Bounds on the Internal Angle

In this section we justify

Lemma 6.1 For $\epsilon > 0$, and $n > n_0(\epsilon, \delta, \rho)$

$$n^{-1} \log(\beta(T^k, T^\ell)) \leq -\Psi_{int}(\ell/n; k/n) + \epsilon,$$

uniformly in $\ell \geq \delta n$, $k \geq \rho n$, $(\ell - k) \geq (\delta - \rho)n$.

6.1 Background

By definition, the internal angle $\beta(F, G)$ is the fraction of the vector space $\text{span}(G - x^F)$ taken up by the positive cone $\text{pos}(G - x^F)$ where x^F is the barycenter of face F . Let now V_{m-1} denote $m - 1$ -dimensional surface measure on the sphere S^{m-1} , while $\Sigma_{m-1}(\alpha)$ denotes the regular spherical simplex (a generalization of triangle) with m vertices on the sphere and sides of angle α . Böröczky and Henk [2] give the formula

$$\beta(T^k, T^\ell) = \frac{V_{\ell-k}(\Sigma_{\ell-k}(\frac{1}{k+2}))}{V_{\ell-k}(S^{\ell-k})}; \quad (6.1)$$

see also Affentranger and Schneider [1] and Vershik and Sporyshev [19].

Defining

$$B(\alpha, m) = \frac{V_{m-1}(\Sigma_{m-1}(\alpha))}{V_{m-1}(S^{m-1})}$$

it is thus of interest to evaluate/bound $B(\frac{1}{k+2}, \ell - k + 1)$. By [2],

$$B(\alpha, m) = \theta^{(m-1)/2} \sqrt{(m-1)\alpha + 1} \pi^{-m/2} \alpha^{-1/2} J(m, \theta) \quad (6.2)$$

where $\theta \equiv (1 - \alpha)/\alpha$ and

$$J(m, \theta) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \left(\int_0^{\infty} e^{-\theta v^2 + 2iv\lambda} dv \right)^m e^{-\lambda^2} d\lambda. \quad (6.3)$$

Vershik and Sporyshev [19] also considered $\beta(T^k, T^\ell)$ in the proportional-dimensional setting $k \sim \rho n$ and $\ell \sim \rho n$. They analyzed a seemingly different integral expression based on contour integration. They sketched out an approach to the asymptotics of that integral by the saddlepoint method. We pursue here a probabilistic approach, and later correlate our findings to theirs.

6.2 Probabilistic Analysis

Seemingly taking a different tack from previous authors, we recognize in the expression for $J(m, \theta)$ a convolution of $m + 1$ probability densities being expressed in the Fourier domain. This leads to the following probabilistic interpretation, proved in the Appendix.

Lemma 6.2 *Put $\theta = \frac{1-\alpha}{\alpha}$. Let T be a random variable with the $N(0, 1/2)$ distribution, and let W_m be a sum of *i.i.d.* half-normals $U_i \sim HN(0, \frac{1}{2\theta})$. Let T and W_m be stochastically independent, and let g_{T+W_m} denote the probability density function of the random variable $T + W_m$. Then*

$$B(\alpha, m) = \sqrt{\frac{\alpha(m-1)+1}{1-\alpha}} \cdot 2^{1-m} \cdot \pi^{1/2} \cdot g_{T+W_m}(0). \quad (6.4)$$

Using the probabilistic interpretation, and applying large deviations techniques, we obtain effective bounds on $g_{T+W_m}(0)$. By the convolution formula, symmetry of the standard normal, support properties of the half-normal, and integration by parts we have

$$\begin{aligned} g_{T+W_m}(0) &= \int_{-\infty}^{\infty} g_T(0-v)g_{W_m}(v)dv = \int_0^{\infty} g_T(v)g_{W_m}(v)dv \\ &= \int_0^{\infty} (-g'_T(v))G_{W_m}(v)dv, \end{aligned}$$

where G_{W_m} denotes cumulative distribution function of the random variable W_m . Since $g'_T(t) = (-2t)g_T(t)$,

$$g_{T+W_m}(0) = 2 \int_0^\infty v g_T(v) G_{W_m}(v) dv.$$

Let $\mu_m = EW_m$ denote the mean of W_m . The part of the integral occurring above μ_m obeys

$$2 \int_{\mu_m}^\infty v g_T(v) G_{W_m}(v) dv \leq 2 \int_{\mu_m}^\infty v g_T(v) dv = g_T(\mu_m) \leq \exp(-\mu_m^2) \cdot 2/\sqrt{\pi}. \quad (6.5)$$

Since $\mu_m = m \cdot \sqrt{1/\pi\theta}$, it will turn out that this part of the integral is typically negligible. The part of the integral below the mean can be controlled by the basic inequalities of large deviations theory [6]; for $v \leq \mu_m \equiv E(W_m)$,

$$G_{W_m}(v) \leq \exp(-\Lambda_{W_m}^*(v)). \quad (6.6)$$

Here $\Lambda_{W_m}^*$ is the rate function of the random variable W_m (i.e. the Fenchel-Legendre dual of the cumulant generating function Λ_{W_m}).

Now as $W_m = U_1 + \dots + U_m$ with the U_i iid, then $\Lambda_{W_m}^*(v) = m\Lambda_{U_1}^*(v/m)$; also $\Lambda_{aU_1}^*(av) = \Lambda_{U_1}^*(v)$. Combining these, we have that if Y denotes a standard half-normal random variable $Y \sim HN(0, 1)$, and if Λ^* denotes the rate function associated to Y , then from $U_1 = Y/\sqrt{2\theta}$ we get

$$g_{T+W_m}(0) \leq \frac{2}{\sqrt{\pi}} \cdot \int_0^{\mu_m} v \exp\left(-v^2 - m\Lambda^*\left(\frac{\sqrt{2\theta}}{m}v\right)\right) dv + \exp(-\mu_m^2) \cdot \frac{2}{\sqrt{\pi}} = I_m + II_m, \text{ say}$$

We have already argued that II_m is negligible. We now change variables $y = \frac{\sqrt{2\theta}}{m}v$, getting

$$I_m \leq 2/\sqrt{\pi} \cdot m^2/2\theta \cdot \int_0^{\sqrt{2/\pi}} y \exp\left(-m\left(\frac{m}{2\theta}\right)y^2 - m\Lambda^*(y)\right) dy. \quad (6.7)$$

6.3 Laplace's Method

The m in the exponent of (6.7) was defined by $m = \ell - k + 1$ which we think of as $(\nu - \rho\delta)n$. Thus m is growing with n , suggesting again a recourse to Laplace's method. We recognize that essentially the function ξ_γ appears in the exponent, with $\gamma = \frac{\theta}{m+\theta-1}$. Recalling that $\theta = (1 - \alpha)/\alpha = k + 1$, and $m = \ell - k + 1$ this then has the form $\gamma_{k,\ell} = \frac{k+1}{\ell+2}$, - for $k \sim \rho n$ and $\ell \sim \nu n$ this is essentially the constant $\gamma = \rho\delta/\nu$. So define the integral

$$f_{\gamma,m} = \int_0^\infty y \exp(-m\xi_\gamma(y)) dy.$$

Again applying Laplace with careful attention to uniformity gives

Lemma 6.3 *For $\gamma \in (0, 1)$ let $y_\gamma \in (0, 1)$ denote the minimizer of ξ_γ . Then*

$$\int_0^\infty f_{\gamma,m}(x) dx \leq \exp(-m\xi_\gamma(y_\gamma)) \cdot R_m(\gamma)$$

where, for $\eta > 0$,

$$m^{-1} \sup_{\gamma \in [\eta, 1]} \log(R_m(\gamma)) = o(1) \text{ as } m \rightarrow \infty.$$

The proof is very similar to that of Lemma 5.2 and we omit it.

We conclude that

$$g_{T+W_m}(0) \leq \exp(-m\xi_\gamma(y_\gamma))R_m(\gamma),$$

where $n^{-1} \log(R_m(\gamma_{k,\ell})) = o(1)$. Applying 6.2 we get

$$n^{-1} \log(\beta(T^k, T^\ell)) \leq -(\xi_\gamma(y_\gamma) + \log(2))(\nu - \rho\delta) + o(1),$$

where the $o(1)$ is uniform over a range of k and ℓ . Arguing much as in the section on Internal Angles, we get Lemma 6.1.

6.4 Properties of ξ_γ

We now briefly review the properties of ξ_γ relevant to computing Ψ_{int} .

By standard convex duality, associated to the cumulant generating function $\Lambda(s)$ and its dual Λ^* , we have the duality relations

$$y = \Lambda'(s), \quad s = (\Lambda^*)'(y), \quad (6.8)$$

defining a one-one relationship $s = s(y)$ and $y = y(s)$ between $s < 0$ and $0 < y < \sqrt{2/\pi}$.

In particular, from $\Lambda(s) = s^2/2 + \zeta_0(s)$, $\zeta_0(s) = \log(2\Phi(s))$,

$$y(s) = s + \zeta_1(s), \quad \zeta_1 = \frac{d}{ds}\zeta_0 = \frac{\phi}{\Phi}, \quad (6.9)$$

where again ϕ and Φ are the normal density and cumulative. It is well known (search under ‘‘Mills’ Ratio’’) that there is a function $M(s)$ defined on $s < 0$ with $0 < M(s) < 1$ and $M(s) \rightarrow 1$ as $s \rightarrow -\infty$ so that

$$\Phi(s) = M(s) \cdot (\phi(s)/|s|), \quad s < 0.$$

There are convenient rational approximations to $M(s)$ in Bryc’s article [3]; we found these essential for numerical work. It follows that $\zeta_1(s) = |s|/M(s)$. Using (6.9) we conclude

$$y(s) = s \cdot (1 - 1/M(s)), \quad s < 0. \quad (6.10)$$

We can also obtain a useful identity for ζ_0 :

$$\zeta_0(s) = \log(2\Phi(s)) = \log(2 \cdot M(s) \cdot (\phi(s)/|s|)) = \log(M(s)) + \log(2/\pi)/2 - s^2/2 - \log(|s|).$$

Substituting in $\Lambda^*(y) = ys - \Lambda(s) = ys - s^2/2 - \zeta_0(s)$, $s = s(y)$, gives

$$\Lambda^*(y(s)) = s^2(1 - 1/M(s)) - \log(M(s)) - \log(2/\pi)/2 + \log(|s|). \quad (6.11)$$

Applying duality (6.8) to $\frac{d}{dy}\xi_\gamma(y) = 0$, we see that the minimizer y_γ of ξ_γ obeys

$$\frac{1 - \gamma}{\gamma} y_\gamma = -s_\gamma. \quad (6.12)$$

Inserting this in (6.10) gives the convenient characterization

$$M(s_\gamma) = 1 - \gamma.$$

Combining with (6.11) we obtain the exact formula

$$\Lambda^*(y_\gamma) = -y_\gamma^2 \frac{1 - \gamma}{\gamma} - \log(2/\pi)/2 + \log(y_\gamma/\gamma),$$

and the identity

$$\xi_\gamma(y_\gamma) = -\frac{1}{2}y_\gamma^2 \frac{1-\gamma}{\gamma} - \log(2/\pi)/2 + \log(y_\gamma/\gamma). \quad (6.13)$$

We now consider asymptotics as $\gamma \rightarrow 0$. We have

$$E \exp(sY) \sim \sqrt{\frac{2}{\pi}} \cdot \frac{-1}{s}, \quad s \rightarrow -\infty.$$

Hence $\Lambda(s) \sim -\log(|s|) + \log(2/\pi)/2$ as $s \rightarrow -\infty$. Strengthening this with remainders shows $\Lambda'(s) \sim -1/s$. Duality then implies the crude asymptotics

$$y(s) \sim -1/s, \quad s \rightarrow -\infty; \quad s(y) \sim -1/y, \quad y \rightarrow 0,$$

as well as

$$y_\gamma \sim \sqrt{\frac{\gamma}{1-\gamma}}, \quad s_\gamma \sim -\sqrt{\frac{1-\gamma}{\gamma}}, \quad \gamma \rightarrow 0. \quad (6.14)$$

From this the leading-order asymptotics of $\xi_\gamma(y_\gamma)$ given in (4.4) follow immediately. Compare also Lemma 8.3.

6.5 Reconciliation with Vershik and Sporyshev

As we have pointed out, Vershik and Sporyshev, in pioneering work [19], considered (in our notation) the internal angle $\beta(T^k, T^\ell)$ with $k \sim \rho n$, $\ell \sim \nu n$, $n \rightarrow \infty$. (To compare our papers, use the dictionary: $\nu \leftrightarrow \beta$, $\rho \leftrightarrow \epsilon$, $\delta \leftrightarrow \alpha$ and $s_\gamma \leftrightarrow t_\epsilon$, our paper \leftrightarrow [19].) They work with an expression apparently different from (6.3) defined by contour integration and sketch a saddlepoint procedure. Their results are not stated in terms of large-deviations properties of half-normal random variables. From the viewpoint of this paper, we understand their results as implicitly working with what we call the dual variables s_γ . It appears that their formula can be rewritten using our notation as

$$n^{-1} \log \beta(T^k, T^\ell) \leq \rho \delta s_\gamma^2 / 2 + \log \left(\frac{1-\gamma}{\sqrt{2\pi}|s_\gamma|} \right) (\nu - \rho \delta) + o(1).$$

Our results can be reexpressed, using (6.11) and (6.12), as follows:

$$\begin{aligned} n^{-1} \log \beta(T^k, T^\ell) + o(1) &\leq -\Psi_{int}(\ell/n, k/n) \\ &= -(\xi_\gamma(y_\gamma) + \log(2)) (\nu - \rho \delta) \\ &= -\left(-\frac{\gamma}{1-\gamma} s_\gamma^2 / 2 + \log\left(\sqrt{\frac{\pi}{2}} \cdot \frac{|s_\gamma|}{1-\gamma}\right) + \log(2) \right) (\nu - \rho \delta) \\ &= \frac{\gamma}{1-\gamma} \frac{s_\gamma^2}{2} (\nu - \rho \delta) - \log\left(\frac{\sqrt{2\pi}|s_\gamma|}{1-\gamma}\right) (\nu - \rho \delta) \\ &= \rho \delta s_\gamma^2 / 2 + \log \left(\frac{1-\gamma}{\sqrt{2\pi}|s_\gamma|} \right) (\nu - \rho \delta). \end{aligned} \quad (6.15)$$

Thus the two answers agree - as of course they would if both approaches were equally precise in determining exponential order. Our systematic approach, based on a probabilistic interpretation, provides a sound motivation for the correctness of this answer. It also provides detailed information useful in other ways, for example in proving Lemmas 4.1 and 4.2.

7 Theorems 2 and 3

Given our considerable efforts in proving Theorem 1, it may be a welcome relief to not that no serious work is required to get the other results announced in the introduction.

7.1 Proof of Theorem 2

Observe that $f_{k-1}(C) = 2^k \binom{n}{k}$; this combinatorial factor has exponential growth with n according to an exponent $\Psi_{face}(\rho\delta) \equiv \rho\delta \log(2) + H(\rho\delta)$; thus, if $k = k(n) \sim \rho\delta n$,

$$n^{-1} \log(f_{k-1}(C^n)) \rightarrow \Psi_{face}(\rho\delta), \quad n \rightarrow \infty.$$

We again define Ψ_{net} as in the proof of Theorem 1.

Definition 2 Let $\delta \in (0, 1]$. The critical proportion $\rho_F(\delta)$ is the supremum of $\rho \in [0, 1]$ obeying

$$\Psi_{net}(\nu; \rho, \delta) < \Psi_{face}(\rho\delta), \quad \nu \in [\delta, 1). \quad (7.1)$$

Recall Section 3's definition $\Delta(k, d, n) = f_{k-1}(C) - f_{k-1}(AC) \geq 0$. The proof of Theorem 2 is based on observing that (7.1) implies

$$\Delta(k, d, n) = o(f_{k-1}(C^n)). \quad (7.2)$$

We immediately get (1.2). Showing that (7.1) implies (7.2) requires no new ideas; one proceeds as in Section 3 almost line-by-line; we omit the exercise. \square

7.2 Proof of Theorem 3

We retrace slightly the discussion of sectional neighborliness in [9]. With probability one, the projected cross-polytope $P = AC$ has $2n$ vertices. Every subset K of k of these vertices which does not contain an antipodal pair defines a k -dimensional subspace V_K of \mathbf{R}^d . Every such V_K defines a so-called *intrinsic section* $P_K = V_K \cap P$. There are $\binom{n}{k}$ different intrinsic sections of P . Define the exponent $\Psi_{sect} = H(\rho\delta)$. Then for $k \sim \rho\delta n$, we have

$$n^{-1} \log(\#\{\text{intrinsic } k\text{-sections}\}) \rightarrow \Psi_{sect}(\rho\delta), \quad n \rightarrow \infty.$$

We again define Ψ_{net} as in the proof of Theorem 1.

Definition 3 Let $\delta \in (0, 1]$. $\rho_S(\delta)$ is the supremum of $\rho \in [0, 1]$ obeying

$$\Psi_{net}(\nu; \rho, \delta) < \Psi_{sect}(\rho\delta), \quad \nu \in [\delta, 1). \quad (7.3)$$

Recall again the definition $\Delta(k, d, n) = f_{k-1}(C) - f_{k-1}(AC) \geq 0$. The proof of Theorem 3 is based on the observation that (7.3) implies, for $k \sim \rho\delta n$,

$$\Delta(k, d, n) = o(\#\{\text{intrinsic } k\text{-sections}\}). \quad (7.4)$$

We have with probability one that the columns of A are in general position. As in [9, Section 6], it follows that all the k -faces of P are k -simplices, $k < d/2$. In order for a given section P_K not to be a k -dimensional cross polytope, it must have

$$f_{k-1}(P_K) < f_{k-1}(C^k),$$

where C^k denotes the k -dimensional cross-polytope. In words, to not have a cross-polytope, a section must 'lose at least one $(k-1)$ -face'. However, as there are $\binom{n}{k}$ intrinsic sections P_K , and the condition (7.4) permits us to 'lose' only $o(1)$ faces per section, we conclude that there are relatively few sections which can 'lose' any faces. Hence the overwhelming majority are intact cross-polytopes. \square

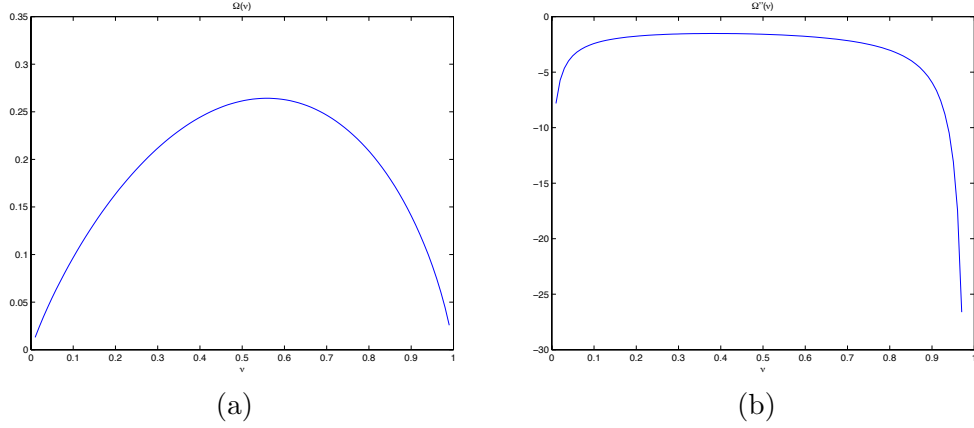


Figure 8.1: Panel (a): $\Omega(\nu) = H(\nu) - \Psi_{ext}(\nu)$ is concave; Panel (b): $\Omega'' < 0$.

8 Appendix

Proof of Lemma 4.1

Fix $\delta > 0$. Apply (4.5) with $\eta = 1/2$. For small enough ρ , the lower bound forces $\Psi_{int}(\nu; \rho\delta)$ to be quite big, uniformly in $\nu \in [\delta, 1]$ and $\delta < \delta_0$, simply by picking ρ small.

For $\delta < \delta_0$ small positive, the terms Ψ_{com} and Ψ_{ext} can be bounded independently of $\rho < \rho_0$ and $\nu \geq \delta$. Hence, for ρ small, the large size of Ψ_{int} can be big enough to force the net exponent Ψ_{net} negative, uniformly in ν . \square

Proof of Lemma 4.2

Proof. We will show that, with $\rho(\delta) = \log(1/\delta)^{-1-\eta}$ and $\delta_0 > 0$ to be chosen below.

$$\Psi_{net}(\nu; \rho(\delta), \delta) < -\delta, \quad \delta < \delta_0, \quad \nu \in [\delta, 1]. \quad (8.1)$$

Hence $\rho_N(\delta) \geq \rho(\delta)$ for $\delta < \delta_0$.

Define $\Omega(\nu) = H(\nu) - \Psi_{ext}(\nu)$. This is concave; see Figure 8.1.

Based on $H(\nu) \sim \nu \log(1/\nu)$ as $\nu \rightarrow 0$ and (4.3), we have the following asymptotic as $\nu \rightarrow 0$:

$$\Omega(\nu) \sim \frac{1}{2} \cdot \log(\log(1/\nu))\nu, \quad \nu \rightarrow 0. \quad (8.2)$$

Also define $K(\nu; \rho, \delta) = \Omega(\nu) - \xi_\gamma(y_\gamma)(\nu - \rho\delta)$. Our proof of (8.1) will be reduced to showing, via (8.2), that $K < 0$ for small ν .

Note the combinatorial identity

$$\binom{n}{k+1} \binom{n-k+1}{\ell-k} = \binom{n}{\ell} \binom{\ell}{k+1}.$$

This implies

$$H(\nu) + H(\rho\delta/\nu)\nu = H(\rho\delta) + H\left(\frac{\nu - \rho\delta}{1 - \rho\delta}\right)(1 - \rho\delta). \quad (8.3)$$

Use this to rewrite the net exponent as

$$\Psi_{net}(\nu; \rho, \delta) = K(\nu; \rho, \delta) + \rho\delta \log(2) + H(\rho\delta/\nu)\nu.$$

Lemma 8.1 gives for $\delta < \delta_1$ that $H(\rho\delta/\nu)\nu \leq H(\rho)\delta + 2\rho(\nu - \delta)$. Then

$$\Psi_{net} \leq K(\nu; \rho, \delta) + [\rho\delta \log(2) + H(\rho)\delta] + 2\rho(\nu - \delta), \quad (8.4)$$

say. Lemma 8.2 establishes concavity of K in ν . Suppose that for $\delta < \delta_1$, $K'(\delta) \leq 0$, then we would have

$$K(\nu; \rho, \delta) \leq K(\delta), \quad \nu \in [\delta, 1].$$

Additionally, suppose that for some $\epsilon_2 > 0$ $K'(\delta) \leq -\epsilon_2 < 0$, then we would have for any $\epsilon_1 > 0$

$$K(\nu; \rho, \delta) + \epsilon_1 + \epsilon_2(\nu - \delta) \leq K(\delta) + \epsilon_1, \quad \nu \geq \delta. \quad (8.5)$$

Below we will show that for $\delta < \delta_2$,

$$K'(\delta; \rho, \delta) \leq -\eta/4 \log(\log(1/\delta)) \quad (8.6)$$

$$K(\delta; \rho, \delta) \leq -\delta \cdot \eta/4 \cdot \log(\log(1/\delta)) \quad (8.7)$$

where $K'(\nu; \rho, \delta) \equiv \frac{\partial}{\partial \nu} K(\nu; \rho, \delta)$. It follows, recalling (8.4) and setting $\epsilon_1 = \rho\delta \log(2) + H(\rho)\delta$ and $\epsilon_2 = 2\rho$, that

$$K'(\delta; \rho, \delta) \leq -\epsilon_2, \quad \delta < \delta_3,$$

and so (8.5) gives, for $\delta < \delta_3$,

$$\Psi_{net}(\nu; \rho, \delta) \leq K(\nu; \rho, \delta) + \epsilon_1, \quad \nu \in [\delta, 1].$$

Now (8.7) shows that $K(\delta)/\delta \rightarrow -\infty$ as $\delta \rightarrow 0$ while evidently $\epsilon_1 = O(\delta)$, hence $K(\delta) + \epsilon_1 < -\delta$ for $\delta < \delta_4$. (8.1) follows with $\delta_0 = \min(\delta_1, \delta_2, \delta_3, \delta_4)$.

It remains to verify (8.6)-(8.7). Writing out

$$\Omega(\nu) = \frac{1}{2}\nu \log(\log(\frac{1}{\nu} + R(\nu))) + (1 - \nu) \log(1/(1 - \nu)) + \nu \log(\frac{1 - \nu}{\sqrt{\pi}}) - \log(2/\sqrt{\pi}),$$

where $R(\nu) = -\log(x_\nu G(x_\nu))$, we compute that

$$\Omega'(\nu) \sim \frac{1}{2} \log(\log(\frac{1}{\nu})), \quad \nu \rightarrow 0.$$

We also recall (8.2). Now by (6.13)-(6.14), we have that as $\delta \rightarrow 0$,

$$\xi_\rho(y_\rho) \sim \frac{1}{2} \log \frac{1}{\rho} \sim \frac{1}{2} \log(\log \frac{1}{\delta})(1 + \eta).$$

Hence for $\delta < \delta_{1,1}$,

$$\xi_\rho(y_\rho)(1 - \rho) \geq (1 + \eta/2)\Omega(\delta)/\delta, \quad (8.8)$$

$$\xi_\rho(y_\rho)(1 - \rho) \geq (1 + \eta/2)\Omega'(\delta). \quad (8.9)$$

This gives

$$K(\delta; \rho(\delta), \delta) < -\frac{\eta}{2} \log(\log(1/\delta))(1 + o(1));$$

and of course for $\delta < \delta_{1,2}$, the $(1 + o(1))$ factor on the right side exceeds $1/2$. Hence (8.7) holds for $\delta_1 = \min(\delta_{1,1}, \delta_{1,2})$. \square

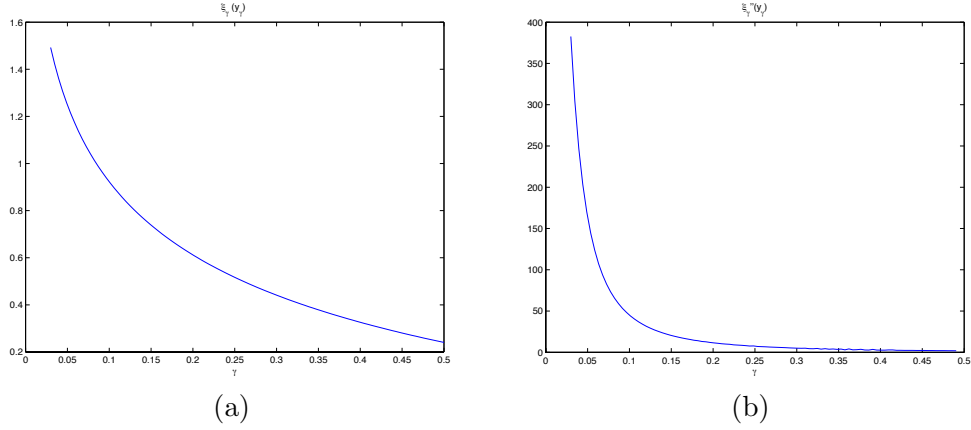


Figure 8.2: Panel (a): $\xi_\gamma(y_\gamma)$ is convex; Panel (b): $\frac{d^2}{d\gamma^2}\xi_\gamma(y_\gamma) > 0$.

Lemma 8.1 For $0 \leq \rho < 1/2$,

$$H(\rho\delta/\nu)\nu \leq H(\rho)\delta + 2\rho(\nu - \delta), \quad \nu \in [\delta, 1). \quad (8.10)$$

Proof. One can compute that, with $\gamma = \rho\delta/\nu$,

$$\frac{\partial}{\partial \nu}[H(\gamma)\nu] = -\log(1 - \gamma).$$

Now for $0 \leq \rho < 1/2$, $-\log(1 - \rho) \leq 2\rho$. Also $\gamma = \rho\delta/\nu \leq \rho$ for $\nu \geq \delta$. Hence

$$\frac{\partial}{\partial \nu}[H(\gamma)\nu] \leq 2\rho, \quad \nu \in [\delta, 1).$$

(8.10) follows. □

Lemma 8.2 For $\rho < \rho_0$ and $\delta < \delta_0$, $K(\nu; \rho, \delta)$ is concave as a function of ν .

Proof. Let

$$\Xi(\nu; \rho, \delta) \equiv \xi_\gamma(y_\gamma)(\nu - \rho\delta).$$

Then $K = \Omega - \Xi$. Since Ω is concave, it is sufficient to show that Ξ is convex. This involves properties of $\xi_\gamma(y_\gamma)$; the reader may be helped by Figure 8.2.

Using $\frac{\partial \gamma}{\partial \nu} = -\gamma/\nu$ we write

$$\begin{aligned} \frac{\partial}{\partial \nu}\Xi(\nu; \rho, \delta) &= \xi_\gamma(y_\gamma) + \frac{\partial}{\partial \gamma}\xi_\gamma(y_\gamma) \cdot \frac{\partial \gamma}{\partial \nu} \cdot (\nu - \rho\delta) \\ &= \xi_\gamma(y_\gamma) - \frac{\partial}{\partial \gamma}\xi_\gamma(y_\gamma) \cdot \gamma(1 - \gamma). \end{aligned}$$

Hence

$$\begin{aligned}
\frac{\partial^2}{\partial \nu^2} \Xi(\nu; \rho, \delta) &= \frac{\partial}{\partial \gamma} \xi_\gamma(y_\gamma) \frac{\partial \gamma}{\partial \nu} - \frac{\partial^2}{\partial \gamma^2} \xi_\gamma(y_\gamma) \cdot \frac{\partial \gamma}{\partial \nu} \cdot \gamma \cdot (1 - \gamma) \\
&\quad - \frac{\partial}{\partial \gamma} \xi_\gamma(y_\gamma) \cdot \frac{\partial \gamma}{\partial \nu} \cdot (1 - \gamma) - \frac{\partial}{\partial \gamma} \xi_\gamma(y_\gamma) \cdot \frac{\partial \gamma}{\partial \nu} \cdot (-\gamma) \\
&= -\frac{\partial^2}{\partial \gamma^2} \xi_\gamma(y_\gamma) \cdot \frac{\partial \gamma}{\partial \nu} \cdot \gamma \cdot (1 - \gamma) + 2 \frac{\partial}{\partial \gamma} \xi_\gamma(y_\gamma) \frac{\partial \gamma}{\partial \nu} \gamma \\
&= \frac{\partial^2}{\partial \gamma^2} \xi_\gamma(y_\gamma) \cdot \gamma^2 \cdot (1 - \gamma) / \nu - 2 \frac{\partial}{\partial \gamma} \xi_\gamma(y_\gamma) \gamma^2 / \nu.
\end{aligned}$$

We next show that the first term is positive for all small γ ; it will also emerge that it dominates the second, and the claimed convexity follows. The key to evaluating the required derivatives is to express $\xi_\gamma(y_\gamma)$ in terms of the dual variables s_γ , as in (6.15). We then have

$$\xi_\gamma(y_\gamma) = \log(|s_\gamma|) + R(\gamma), \quad (8.11)$$

where the remainder $R(\gamma)$ involves other terms, also explicitly expressed in terms of γ and s_γ :

$$R(\gamma) = \log(\sqrt{2\pi}/(1 - \gamma)) - s_\gamma^2/2 \cdot (\gamma/(1 - \gamma)).$$

The term $\log(|s_\gamma|)$ gives rise to large derivatives. Thus, using (8.11) and Lemma 8.3,

$$\frac{\partial}{\partial \gamma} \log(|s_\gamma|) = s_\gamma^{-1} \frac{d}{d\gamma} s_\gamma \sim (-\gamma^{1/2}) \left(\frac{1}{2} \gamma^{-3/2}\right) \sim -\gamma^{-1}/2, \quad \gamma \rightarrow 0, \quad (8.12)$$

while

$$\frac{\partial^2}{\partial \gamma^2} \log(|s_\gamma|) = \left(\frac{d}{d\gamma} s_\gamma\right)^2 / s_\gamma^2 + \left(\frac{d^2}{d\gamma^2} s_\gamma\right) / s_\gamma \sim \gamma^{-4}/4, \quad \gamma \rightarrow 0. \quad (8.13)$$

The remainder R turns out to have relatively negligible influence – principally because $s_\gamma^2 \gamma = O(1)$ does not have such large derivatives in γ ; indeed

$$\frac{\partial^2}{\partial \gamma^2} s_\gamma^2 \gamma = 2s_\gamma'' s_\gamma \gamma + 2(s_\gamma')^2 \gamma + 4s_\gamma' s_\gamma;$$

applying Lemma 8.3 below, this is $O(\gamma^{-2}) \ll O(\gamma^{-4})$. Hence we conclude that for some $\gamma_0 > 0$ sufficiently small, $\frac{\partial^2}{\partial \gamma^2} \xi_\gamma(y_\gamma)$ is positive uniformly in $\gamma \leq \gamma_0$, and, since $\gamma \leq \rho$, uniformly in $\rho \leq \rho_0 \equiv \gamma_0$. The dominance of this term flows from the different asymptotic orders of (8.12)-(8.13). \square

Lemma 8.3 *As $\gamma \rightarrow 0$,*

$$\begin{aligned}
s_\gamma &\sim -\gamma^{-1/2}, \\
\frac{d}{d\gamma} s_\gamma &\sim \frac{1}{2} \gamma^{-3/2}, \\
\frac{d^2}{d\gamma^2} s_\gamma &\sim -\frac{3}{4} \gamma^{-5/2}.
\end{aligned}$$

Proof. We use the classical asymptotic series

$$R_x \equiv \Phi(-x)/\phi(x) \sim \frac{1}{x} - \frac{1}{x^3} + \frac{1 \cdot 3}{x^5} - \frac{1 \cdot 3 \cdot 5}{x^7} + \dots, \quad x > 0;$$

Harold Ruben [16] attributes this to Laplace. Since $M(s) = |s|R_{|s|}$, we have

$$M(-x) \sim 1 - \frac{1}{x^2} + \frac{1 \cdot 3}{x^4} - \frac{1 \cdot 3 \cdot 5}{x^6} + \dots, \quad x > 0.$$

Because there is no $1/x$ term in this formula, the solution s_γ of

$$M(s_\gamma) = 1 - \gamma \tag{8.14}$$

obeys

$$s_\gamma^{-2} \sim \gamma, \quad \gamma \rightarrow 0,$$

i.e. $s_\gamma \sim -\gamma^{-1/2}$. To take derivatives of s_γ we use the inverse function theorem:

$$\frac{d}{d\gamma}s_\gamma = -1/M'(s_\gamma); \quad \frac{d^2}{d\gamma^2}s_\gamma = M''(s_\gamma)/(M'(s_\gamma))^3.$$

We can explicitly compute the derivatives of $N(x) = M(-x)$, getting

$$N' = N/x + x + xN; \quad N'' = 1 + (N - 1)(3 + x^2)$$

Using this and the asymptotic series for M given above, we obtain asymptotic expressions for $M'(s_\gamma)$ and $M''(s_\gamma)$, which give the desired result. \square

Proof of Lemma 5.1

The uniform Laplace method ultimately boils down to this statement

Lemma 8.4 *Let ψ be convex and C^2 on an interval I and suppose that it takes its minimum at an interior point $x_0 \in I$ where $\psi'' > 0$ and that in a vicinity $(x_0 - \epsilon, x_0 + \epsilon)$ of x_0 :*

$$|\psi''(x) - \psi''(x_0)| \leq C|\psi''(x_0)||x - x_0|. \tag{8.15}$$

Let $\bar{\psi}$ be the quadratic approximation $\psi(x_0) + \psi''(x_0)(x - x_0)^2/2$. Then

$$\int_I \exp(-n\psi(x))dx \leq \int_{-\infty}^{\infty} \exp(-n\bar{\psi}(x))dx \cdot (S_{1,n} + S_{2,n})$$

where

$$S_{1,n} = \exp(n\psi''(x_0)C\epsilon^2/6),$$

$$S_{2,n} = 2/\left(n\epsilon(|\psi''(0)|2\pi)^{1/2}(1 - C\epsilon^2/2)\right).$$

We defer the proof to the next subsection.

The constant C referred to in the lemma is effectively a (scaled) third derivative, since if ψ is C^3 , we may take in (8.15)

$$C = C(\epsilon) = \sup_{(x_0 - \epsilon, x_0 + \epsilon)} \psi^{(3)}(x)/\psi''(x).$$

The lemma has the following significance for Lemma 5.2. Pick $\epsilon_n = n^{-2/5}$. Pick $n_0 = n_0(\psi''(x_0), C)$ so that

$$\psi''(x_0)Cn^{-1/5}/6 < 1, \quad Cn^{-2/5}/3 < 1/2.$$

Then for $n > n_0$, $\exp(u) \leq 1 + (e-1)u$ gives

$$S_{1,n} \leq 1 + (e-1)(\psi''(x_0)C/6) \cdot n^{-1/5}$$

while $(1 - Cn^{-2/5}/3) > 1/2$ gives

$$S_{2,n} \leq 2/\psi''(0)^{1/2}n^{-3/5}.$$

Hence

$$S_{1,n} + S_{2,n} \leq 1 + o(1),$$

We conclude that

$$\int_I \exp(-n\psi(x))dx \leq \int_{-\infty}^{\infty} \exp(-n\bar{\psi}(x))dx \cdot (1 + o(1)) \quad (8.16)$$

Here the $o(1)$ is uniform over any collection of convex functions with a given $\psi''(x_0)$ and C . Now we consider the collection of convex functions ψ_ν defined in Section 5. Recall $\eta > 0$ and $\delta > 0$ in the statement of Lemma 5.2. Over the interval $\nu \in [\delta, 1-\eta]$, they obey uniform bounds on ψ''_ν and ψ'''_ν of the type needed for Lemma 8.4. We record this in the next lemma. Lemma 5.2 then follows from this and the uniformity in (8.16). \square

Lemma 8.5 *The function ψ_ν is C^∞ with second derivative at the minimum*

$$\psi''_\nu(x_\nu) = 2\nu(1 + x_\nu^2 \frac{4\nu}{1-\nu}) \quad (8.17)$$

and third derivative at the minimum

$$\psi'''_\nu(x_\nu) = (1-\nu)((2-4x_\nu^2)z + 6x_\nu z^2 + 2z^2), \quad (8.18)$$

where $z = z_\nu = 2\nu x_\nu / (1-\nu)$. We have

$$0 < 2\delta \leq \inf_{\nu \in (\delta, 1]} \psi''_\nu(x_\nu).$$

For $\epsilon < \min(\delta/2, \eta/2)$, the ratio

$$C(\epsilon; \delta, \eta) = \sup_{\nu \in (\delta, 1-\eta]} \sup_{|x-x_\nu| < \epsilon} \psi'''_\nu(x) / \psi''_\nu(x)$$

is finite.

Proof. We calculate that

$$\begin{aligned} \psi'(x) &= -(1-\nu)g/G + 2\nu x; & \psi''(x) &= -(1-\nu)(g'/G - g^2/G^2) + 2\nu; \\ \psi'''(x) &= -(1-\nu)(g''/G - 3g'g/G^2 - 2g^3/G^3). \end{aligned}$$

We also remark that $g' = (-2x)g$ and $g'' = (-2 + 4x^2)g$. At the point x_ν , we have

$$g(x_\nu)/G(x_\nu) = \frac{2\nu x_\nu}{1-\nu} = z_\nu, \text{ say.}$$

We combine these to get (8.17) and (8.18).

The formula for $\psi''_\nu(x_\nu)$ immediately gives $\psi''_\nu(x_\nu) \geq 2\nu$, so it is bounded away from zero on any interval $\nu \in [\delta, 1]$, $\delta > 0$.

Now as for $\psi^{(3)}$, we note that clearly z_ν and x_ν are continuous functions on $(0, 1)$. Note that as $\nu \rightarrow 1$,

$$z_\nu \sim \sqrt{\frac{2\nu}{1-\nu}},$$

which diverges, and as $\nu \rightarrow 0$,

$$x_\nu \sim \sqrt{\log((1-\nu)/2\nu)},$$

which also diverges. However, both are bounded on any interval $\nu \in [\delta, 1-\eta]$. As a polynomial in ν , x_ν and z_ν , $\psi''_\nu(x_\nu)$ is also bounded. Boundedness also holds, by inspection if we consider the behavior at x near to x_ν . \square

Proof of Lemma 8.4

We start with a preliminary lemma.

Lemma 8.6 *Let $\psi(x)$ be a C^2 function defined on $[-\epsilon, \epsilon]$ so that*

$$\psi'(0) = 0, \quad \psi''(0) > 0, \quad |\psi''(x) - \psi''(0)| \leq C \cdot \psi''(0)|x|.$$

Then

$$\|\psi - \psi''(0)x^2/2\|_{L^\infty(-\epsilon, \epsilon)} \leq C\psi''(0)\epsilon^3/6. \quad (8.19)$$

and

$$|\psi'(x)| \geq \psi''(0)|x|(1 - Cx^2/2), \quad x \in [-\epsilon, \epsilon]. \quad (8.20)$$

This involves standard estimates, and we omit the proof.

To apply this lemma, we split the integral

$$\int_I \exp(-n\psi(x))dx = \int_{-\epsilon}^\epsilon + \int_{[-\epsilon, \epsilon]^c} = I + II.$$

Near x_0 , we use (8.19), i.e. the fact that the quadratic approximation has ϵ^3 -order accuracy:

$$\begin{aligned} I &\leq \left(\int_{-\epsilon}^\epsilon \exp(-n\bar{\psi}(x))dx \right) \cdot \exp\{n\|\psi - \bar{\psi}\|_{L^\infty(-\epsilon, \epsilon)}\} \\ &\leq \left(\int_{-\infty}^\infty \exp(-n\bar{\psi}(x))dx \right) \cdot \exp\{n\psi''(0)C\epsilon^3/6\} \end{aligned}$$

Away from x_0 we use convexity. To simplify notation, suppose $x_0 = 0$. Convexity gives

$$\psi(x) \geq \psi(\epsilon) + \psi'(\epsilon)(x - \epsilon), \quad x > \epsilon.$$

Hence, using the formula $\int_0^\infty \exp(-n(a + bu))du = \exp(-na)/(nb)$,

$$\int_\epsilon^\infty \exp(-n\psi(x))dx \leq \exp(-n\psi(0))/(n\psi'(\epsilon)).$$

Now by (8.20), $\psi'(\epsilon) \geq \psi''(0)\epsilon(1 - C\epsilon^2/2)$. Also,

$$\exp(-n\psi(0)) = \int \exp(-n\bar{\psi}(x))dx \cdot \sqrt{\frac{\psi''(0)}{2\pi}}$$

so that

$$\int_{\epsilon}^{\infty} \exp(-n\psi(x))dx \leq \exp(-n\psi(0))/(n\epsilon|\psi''(0)|^{1/2}(2\pi)^{1/2}(1 - C\epsilon^2/2)).$$

The lemma follows. \square

Proof of Lemma 6.2

Define half-normal R.V.'s $U_i = |Z_i|$, $Z_i \sim_{iid} N(0, \frac{1}{2\theta})$, and the scaled normal $T \sim N(0, 1/2)$. Use the notation g_X for probability density of RV X and \hat{g}_X for characteristic function (Fourier transform) $\hat{g}_X(\lambda) = \int e^{i\lambda x} g_X(x) dx$. We remark that

$$g_{U_1}(u) = \begin{cases} e^{-\theta u^2} \sqrt{\frac{2\theta}{\pi}} & u \geq 0 \\ 0 & u < 0 \end{cases}$$

and

$$\hat{g}_{U_1}(\lambda) = 2\sqrt{\frac{\theta}{\pi}} \int_0^{\infty} e^{-\theta u^2 + iu\lambda} du.$$

Hence

$$J(m, \theta) = 2^{-m} \left(\sqrt{\frac{\pi}{\theta}}\right)^m \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \hat{g}_T(\lambda) (\hat{g}_{U_1}(2\lambda))^m d\lambda.$$

Now we define $W_m = 2 \sum U_i$ and note that the convolution theorem gives

$$\hat{g}_{W_m}(\lambda) = \hat{g}_{U_1}(2\lambda)^m.$$

Hence

$$J(m, \theta) = 2^{-m} (\pi/\theta)^{m/2} \pi^{-1/2} \int_{-\infty}^{\infty} \hat{g}_T(\lambda) \hat{g}_{W_m}(\lambda) d\lambda.$$

Again by the convolution theorem $\hat{g}_{T+W_m} = \hat{g}_T \hat{g}_{W_m}$, and from the Fourier inversion formula $g_X(0) = (2\pi)^{-1} \int \hat{g}_X(\lambda) d\lambda$ we recognize that

$$\int_{-\infty}^{\infty} \hat{g}_T(\lambda) \hat{g}_{W_m}(\lambda) d\lambda = 2\pi g_{T+W_m}(0),$$

giving

$$J(m, \theta) = 2^{-m} (\pi/\theta)^{-m/2} 2\sqrt{\pi} g_{T+W_m}(0). \quad \square$$

Proof of Theorem 4

Using (8.3) in the same way as in the proof of Lemma 4.2, we have

$$\Psi_{net} - \Psi_{face} = H\left(\frac{\nu - \rho\delta}{1 - \rho\delta}\right)(1 - \rho\delta) - \xi_{\gamma}(y_{\gamma}) \cdot (\nu - \rho\delta) - \psi_{\nu}(x_{\nu}).$$

We make three elementary observations:

$$0 = \lim_{\delta \rightarrow 1} \sup_{\nu \in [\delta, 1]} H\left(\frac{\nu - \rho\delta}{1 - \rho\delta}\right)(1 - \rho\delta) \quad (8.21)$$

$$0 = \lim_{\delta \rightarrow 1} \sup_{\nu \in [\delta, 1]} \psi_{\nu}(x_{\nu}) \quad (8.22)$$

$$\xi_\gamma(y_\gamma) \cdot (\nu - \rho\delta) \geq \xi_\rho(y_\rho) \cdot (1 - \rho)\delta, \quad \nu \geq \delta. \quad (8.23)$$

At the same time $\xi_\rho(y_\rho) > 0$ for $\rho < 1$; see Figure 8.2. Now fix $\rho < 1$ and $0 < \epsilon < \xi_\rho(y_\rho)(1 - \rho)/2$. Combining (8.21)-(8.23), we have for sufficiently large $\delta < 1$,

$$\Psi_{net}(\nu; \rho, \delta) - \Psi_{face}(\nu; \rho, \delta) < \epsilon - \xi_\rho(y_\rho)(1 - \rho)\delta < -\xi_\rho(y_\rho)(1 - \rho)/2 < 0, \quad \nu \in [\delta, 1) \quad (8.24)$$

We conclude from the definition of ρ_N that $\rho_N(\delta) > \rho$. The desired relation (1.3) follows.

It follows in particular from (8.24) that for some $\delta < 1$, there is $\rho > 1/2$ with

$$\Psi_{net}(\nu; \rho, \delta) - \Psi_{face}(\nu; \rho, \delta) < 0, \quad \nu \in [\delta, 1)$$

taking δ_0 to be the infimum of all such δ gives (1.4). □

References

- [1] Fernando Affentranger and Rolf Schneider. Random projections of regular simplices. *Discrete Comput. Geom.*, 7(3):219–226, 1992.
- [2] Károly Böröczky, Jr. and Martin Henk. Random projections of regular polytopes. *Arch. Math. (Basel)*, 73(6):465–473, 1999.
- [3] Włodzimierz Bryc. A uniform approximation to the right normal tail integral. *Appl. Math. Comput.*, 127(2-3):365–374, 2002.
- [4] Geoffrey R. Burton. The nonneighbourliness of centrally symmetric convex polytopes having many vertices. *J. Combin. Theory Ser. A*, 58(2):321–322, 1991.
- [5] Emmanuel J. Candès and Terence Tao. Near optimal signal recovery from random projections and universal encoding strategies. Technical report, Applied and Computational Mathematics, California Institute of Technology, 2004.
- [6] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 1998.
- [7] David L. Donoho. Compressed sensing. Technical report, Department of Statistics, Stanford University, 2004.
- [8] David L. Donoho. For most large systems of underdetermined equations, the minimum ℓ^1 -norm solution is the sparsest solution. Technical report, Department of Statistics, Stanford University, 2004.
- [9] David L. Donoho. Neighborly polytopes and sparse solutions of underdetermined linear equations. Technical report, Department of Statistics, Stanford University, 2004.
- [10] David Gale. Neighboring vertices on a convex polyhedron. In *Linear inequalities and related system*, Annals of Mathematics Studies, no. 38, pages 255–263. Princeton University Press, Princeton, N.J., 1956.
- [11] Branko Grünbaum. Grassmann angles of convex polytopes. *Acta Math.*, 121:293–302, 1968.

- [12] Branko Grünbaum. *Convex polytopes*, volume 221 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 2003. Prepared and with a preface by Volker Kaibel, Victor Klee and Günter M. Ziegler.
- [13] Peter McMullen. Non-linear angle-sum relations for polyhedral cones and polytopes. *Math. Proc. Cambridge Philos. Soc.*, 78(2):247–261, 1975.
- [14] Peter McMullen and Geoffrey C. Shephard. Diagrams for centrally symmetric polytopes. *Mathematika*, 15:123–138, 1968.
- [15] Harold Ruben. On the geometrical moments of skew-regular simplices in hyperspherical space, with some applications in geometry and mathematical statistics. *Acta Math.*, 103:1–23, 1960.
- [16] Harold Ruben. A convergent asymptotic expansion for Mill’s ratio and the normal probability integral in terms of rational functions. *Math. Ann.*, 151:355–364, 1963.
- [17] Rolf Schneider. Neighbourliness of centrally symmetric polytopes in high dimensions. *Mathematika*, 22(2):176–181, 1975.
- [18] A. M. Vershik and P. V. Sporyshev. An asymptotic estimate for the average number of steps in the parametric simplex method. *Zh. Vychisl. Mat. i Mat. Fiz.*, 26(6):813–826, 958, 1986.
- [19] A. M. Vershik and P. V. Sporyshev. Asymptotic behavior of the number of faces of random polyhedra and the neighborliness problem. *Selecta Math. Soviet.*, 11(2):181–201, 1992. Selected translations.