A moment ratio bound for polynomials and some extremal properties of Krawchouk polynomials and Hamming spheres

Naomi Kirshner and Alex Samorodnitsky

Abstract

Let $p \geq 2$. We improve the bound $\frac{\|f\|_p}{\|f\|_2} \leq (p-1)^{s/2}$ for a polynomial f of degree s on the boolean cube $\{0,1\}^n$, which comes from hypercontractivity, replacing the right hand side of this inequality by an explicit bivariate function of p and s, which is smaller than $(p-1)^{s/2}$ for any $p > 2$ and $s > 0$. We show the new bound to be tight, within a smaller order factor, for the *Krawchouk polynomial* of degree s.

This implies several nearly-extremal properties of Krawchouk polynomials and Hamming spheres (equivalently, Hamming balls). In particular, Krawchouk polynomials have (almost) the heaviest tails among all polynomials of the same degree and ℓ_2 norm[1]. The Hamming spheres have the following approximate edge-isoperimetric property: For all $1 \leq s \leq \frac{n}{2}$, and for all even distances $0 \leq i \leq \frac{2s(n-s)}{n}$, the Hamming sphere of radius s contains, up to a multiplicative factor of $O(i)$, as many pairs of points at distance i as possible, among sets of the same size[2]. This also implies that Hamming spheres are (almost) stablest with respect to noise among sets of the same size. In coding theory terms this means that a Hamming sphere (equivalently a Hamming ball) has the maximal probability of undetected error, among all binary codes of the same rate.

We also describe a family of hypercontractive inequalities for functions on $\{0,1\}^n$, which improve on the 'usual' "$q \to 2$" inequality by taking into account the concentration of a function (expressed as the ratio between its ℓ_r norms), and which are nearly tight for characteristic functions of Hamming spheres.

1 Introduction

We prove upper bounds on the moments of polynomials on the discrete cube $\{0,1\}^n$ endowed with uniform measure. Let H be the binary entropy function, and let $\psi(p,x)$ be a function on $[2,\infty) \times [0,1/2]$, defined by

$$\psi(p,x) \quad = \quad (p-1) + \log_2\left((1-\delta)^p + \delta^p\right) - \frac{p}{2}H(x) - px\log_2(1-2\delta),$$

where δ is determined by $x = \left(\frac{1}{2} - \delta\right) \cdot \frac{(1-\delta)^{p-1} - \delta^{p-1}}{(1-\delta)^p + \delta^p}$. (It will be shown that δ is well-defined.)

[1] This has to be interpreted with some care.

[2] There is a similar, but slightly weaker and somewhat more complicated claim for general distances.

Then, for $p \geq 2$, $0 \leq s \leq \frac{n}{2}$,[3] and for a homogeneous polynomial f of degree s on $\{0,1\}^n$, holds

$$\frac{\mathbb{E}\,|f|^p}{(\mathbb{E}\,f^2)^{\frac{p}{2}}} \;\leq\; 2^{\psi\left(p,\frac{s}{n}\right)\cdot n}. \tag{1}$$

We will show this to be an improvement over the well-known bound

$$\frac{\mathbb{E}\,|f|^p}{(\mathbb{E}\,f^2)^{\frac{p}{2}}} \;\leq\; (p-1)^{\frac{ps}{2}}, \tag{2}$$

which follows from the hypercontractive inequality (4) below (see e.g. [34]). Let $\psi_1(p,x) = \frac{p\log_2(p-1)}{2}\cdot x$, so that $\psi_1\left(p,\frac{s}{n}\right) = \frac{1}{n}\log_2\left((p-1)^{\frac{ps}{2}}\right)$. We will show that for any fixed $p > 2$ the functions ψ and ψ_1 and their first derivatives coincide at $x = 0$ and, moreover, that the function ψ is strongly concave in x. This will imply $\psi(p,x) < \psi_1(p,x)$ for any $p > 2$ and $x > 0$.

For a fixed $p > 2$ and for $s \ll n$, the bounds in (1) and in (2) are very close to each other, in accord with the fact ([26]) that if s is a slowly growing function of n, the RHS in (2) cannot be replaced by C^s with $C < (p-1)^{\frac{p}{2}}$. However, if we allow p to grow with n, the two bounds can be significantly different, even for small s. This will be important in estimates which take into account higher moments of polynomials, as is the cases we discuss below.

Let us also observe that both bounds hold in somewhat higher generality - for all polynomials of degree *at most* s on $\{0,1\}^n$ (see Corollary 1.4 below).

We proceed with an informal description of several applications of (1). The formal statements and a more extensive discussion of these results will be given below, in Section 1.2. First, it will be shown that (1) is "nearly tight" (in the sense that will be clarified below) if f is the *Krawchouk polynomial* K_s defined by

$$K_s(x) \;=\; \sum_{S\subseteq[n],|S|=s} (-1)^{\sum_{i\in S} x_i}, \quad \text{for} \quad x = (x_1, x_2, ...x_n) \in \{0,1\}^n.$$

Recalling that K_s is proportional to the Fourier transform of the characteristic function of the Hamming sphere of radius s around zero, this says, alternatively, that Fourier transforms of Hamming spheres are nearly extremal with respect to (1). This will be shown to imply that Krawchouk polynomials and Hamming spheres have certain nearly extremal properties, compared to other objects with similar characteristics. Specifically, we will show that, up to at most polynomial in n error, the following facts hold for functions on $\{0,1\}^n$:

- Krawchouk polynomials have (almost) the heaviest tails among all polynomials of the same degree and ℓ_2 norm. That is, for a polynomial f of degree s with $\|f\|_2 = \|K_s\|_2$, and for a threshold $T > 0$ holds

  $$\Pr\{|K_s| \geq T\} \;\gtrsim\; \Pr\{|f| \geq T'\},$$

  where T' is not much larger than T. For the exact formulation see Theorem 1.5.

[3] This is the interesting range of parameters in terms of s, since the spaces of homogeneous polynomials of degree s and $n-s$ on $\{0,1\}^n$ are isometric, for any ℓ_p norm, see Section 1.1.2.

- For any $0 \le s \le \frac{n}{2}$ and any even $0 \le i \le \frac{2s(n-s)}{n}$, the Hamming sphere of radius s around 0 contains (almost) the "maximal" number of pairs of points at distance i, among all sets of the same size. For a general distance i, the same holds for the union of two Hamming spheres of consecutive radii.

  For the exact formulation see Theorem 1.6.

- For any $p \ge 2$, characteristic functions of Hamming spheres are (almost) stablest with respect to noise among all functions with the same ℓ_1 and ℓ_p norms. That is, let $0 \le s \le \frac{n}{2}$, let f_s be the characteristic function of the Hamming sphere of radius s around 0, and let f be a function with $\|f\|_1 = \|f_s\|_1$ and $\|f\|_p = \|f_s\|_p$. Then, for the noise operator T_ϵ, $0 \le \epsilon \le \frac{1}{2}$, holds

  $$\langle f_s, T_\epsilon f_s \rangle \ \gtrsim \ \langle f, T_\epsilon f \rangle .$$

  For the exact formulation see Corollary 1.9 and the discussion after it.

- For any $p \ge 2$, characteristic functions of Hamming spheres have (almost) the largest spectral projections among all functions with the same ℓ_1 and ℓ_p norms. That is, in the notation of the previous item, for any $0 \le s \le \frac{n}{2}$ and 'many' $0 \le k \le \frac{n}{2}$ holds

  $$\|\Pi_k f_s\|_2 \ \gtrsim \ \|\Pi_k f\|_2.$$

  Here $\Pi_k f$ is the orthogonal projection on the span of Walsh-Fourier characters of weight k. For the exact formulation see Theorem 1.10.

Let us make several comments about these results.

– In all the statements above 'homogeneous polynomials of degree s' can be replaced with 'polynomials of degree s', and 'Hamming spheres of radius s' with 'Hamming balls of radius s' (we do not go into details for lack of space, but see Corollary 1.4.)

– It can be seen that the last three of the claims above are essentially equivalent to each other.

– The exact formulations of the claims above will be in terms of functional inequalities (for functions on $\{0,1\}^n$) involving certain explicit, but rather complicated, functions of two variables. These bivariate functions describe the relevant aspects of behavior of Hamming spheres or of Krawchouk polynomials. For instance, consider the function ψ defined above. As will be seen, $\psi\left(p, \frac{s}{n}\right)$ is the right constant in the exponent of the ratio between the p^{th} moment of the Krawchouk polynomial K_s and the $p/2$-power of its second moment. We point out that the appearance of these functions in the statements of the results indicates that Hamming spheres / Krawchouk polynomials are indeed (almost) extremal objects for these results.

– Continuing from the preceding comment, we observe that while these bivariate functions describe the correct exponential behavior of Hamming spheres or Krawchouk polynomials, they do introduce error, which is polynomial in the dimension n of the discrete cube. This is the cause of imprecision in all of the results above. Let us provide some details. Krawchouk polynomials on $\{0,1\}^n$ and n-dimensional Hamming spheres are discrete objects (if we view a polynomial as a vector of its coefficients), whose behavior is described by expressions involving binomial

coefficients. Hence it cannot be reduced to a simply exponential expression without incurring a certain loss. In our case this (lossy) reduction is achieved by replacing the binomial coefficient $\binom{b}{a}$ by a larger exponential expression $2^{H\left(\frac{a}{b}\right)\cdot b}$ (see (3) below). This is the main source of loss we incur. For an illustration see Example 1.1 below and observe that the gap between the upper bound and the lower bound given by a Hamming sphere is due solely to replacing two binomial coefficients by corresponding exponential expressions.

– Finally, we observe that a polynomial error will typically be much smaller than the main term in the estimates we discuss, since the approximation of $\binom{b}{a}$ by $2^{H\left(\frac{a}{b}\right)\cdot b}$ is usually a very good one. However, this fact has to be interpreted with some care, since the significance of an inaccuracy depends on the context. Consider the following two examples.

Example 1.1:

We will show in Theorem 1.6 that if A a subset of $\{0,1\}^n$ with $|A| \leq \binom{n}{s}$, and if $0 \leq i \leq \frac{2s(n-s)}{n}$, then the number of pairs of points at distance i in A is bounded from above by $|A| \cdot 2^{H\left(\frac{i}{2s}\right)\cdot s + H\left(\frac{i}{2(n-s)}\right)\cdot(n-s)}$. On the other hand, if A is a Hamming sphere of radius s, and if i is even, this number is $|A|\cdot\binom{s}{i/2}\binom{n-s}{i/2}$. This, by (3), is at least $\Omega\left(\frac{1}{i}\right)\cdot|A|\cdot 2^{H\left(\frac{i}{2s}\right)\cdot s + H\left(\frac{i}{2(n-s)}\right)\cdot(n-s)}$. So here the error is of order i, which is significant if we view this as an isoperimetric-type result, since in this context one is typically interested in almost tight results. (With that, to the best of our knowledge, the bounds we obtain are new. In particular, for $i = 2$ we seem to obtain some new estimates related to the Kleitman-West problem. See the discussion in Section 1.2.3.) ∎

Example 1.2:

We will show (as a corollary of Theorem 1.6) that if A is a binary code of length n used over a binary symmetric channel, then the undetected error probability of A is at most $O\left(n^2\right)$ times that of the union of two Hamming spheres of adjacent radii, whose size is roughly that of A. So here the error is of order n^2. However, in this type of coding estimates sub-exponential errors are ignored. Hence this result implies that unions of Hamming spheres (one can also take a Hamming sphere or a Hamming ball of an appropriate size) have, asymptotically, the largest undetected error probability over the binary symmetric channel. (See (8) and the discussion preceding it, and also Section 3.2.) ∎

– Finally, let us draw attention to the special case of the third of the claims above (it is also closely related to the second example above) in which f is a characteristic function of a set. The claim then is that characteristic functions of Hamming spheres (or Hamming balls) are almost stablest with respect to noise among all sets of the same cardinality. To say this differently, consider the following probabilistic experiment. Given a subset A of $\{0,1\}^n$, choose uniformly at random a point x in A. Flip each coordinate in x independently with probability ϵ and check whether the obtained point is also in A. Then, the probability of this event is maximized (up to a sublinear in n factor) if A is a Hamming sphere (ball).

Let us say a few words about the proofs, focusing on the proof of (1), since the applications described above follow from it in a more or less standard manner. We prove (1) in Theorem 1.3

by a comparison argument, showing by induction on the dimension that for a homogeneous polynomial f of degree s, the ratio $\frac{\mathbb{E}|f|^p}{(\mathbb{E}f^2)^{p/2}}$ cannot be much larger than that for the Krawchouk polynomial K_s. The error we obtain in this part of the argument is subexponential in the dimension. It is then reduced to a polynomial error by a tensorization argument (see Subsection 1.1.4 below), applying the claim proved in the first step to tensor powers $f^{\otimes m}$ and passing to the limit as $m \to \infty$. In this limit argument, the behavior of discrete objects such as Krawchouk polynomials is smoothened out, leading to a simply exponential expression in (1), and incurring a polynomial loss (see also the discussion above).

A key element in controlling the growth of $\frac{\mathbb{E}|f|^p}{(\mathbb{E}f^2)^{p/2}}$ with dimension in the induction part of this argument is Hanner's inequality [30]: for $p \geq 2$ and for any two functions g_0, g_1 holds

$$\|g_0+g_1\|_p^p + \|g_0-g_1\|_p^p \;\leq\; (\|g_0\|_p + \|g_1\|_p)^p + \Big|\|g_0\|_p - \|g_1\|_p\Big|^p.$$

An important part of our argument is showing the following intriguing fact: for any fixed $p \geq 2$ and for sufficiently large n and s, Krawchouk polynomials K_{s-1} and K_s on $\{0,1\}^n$ satisfy Hanner's inequality almost with equality. To show this we rely on many known properties of Krawchouk polynomials (see Section 2.2) and also prove some seemingly new ones: In particular, we provide a rather tight estimate for the ℓ_p norms of Krawchouk polynomials; and show their ℓ_2 norm to be attained with only polynomial loss between any two of their roots, and also before their first and after their last roots. An additional implication of our results is that the above mentioned bivariate functions provide an accurate description of the behavior of Krawchouk polynomials K_s for any sufficiently large s (even a large constant s). Previously this seems to have been known mostly for s growing linearly with dimension n (see also [22] where the behavior of the magnitude of $|K_s|$ was analyzed for any s).

Related work

– A special case of (1), for $p = 4$, was shown in [19], where it was also conjectured that the Krawchouk polynomials actually attain the maximum for $\frac{\mathbb{E}f^4}{(\mathbb{E}f^2)^2}$ among all homogeneous polynomials of the same degree. This conjecture has been recently proved in [1], by a short and a very elegant argument (using compression). It seems possible to extend the argument in [1] to work for any even integer p. However, since this argument is essentially combinatorial in nature, it is not immediately obvious how to extend it to general $p > 2$.

– After completing this paper, we have learned [38] that a generalization of Theorem 1.8 and Corollary 1.9 was proved in a concurrent work [39]. More specifically [39] proves the conjecture of [37] (see the discussion following Corollary 1.9 in Section 1.2.4).

– It was shown in [7] that characteristic functions of Hamming spheres (or Hamming balls) of cardinality $2^{n-\alpha(n)}$, where $\alpha(n)$ is a slowly growing function of n, are (almost) stablest with respect to noise among all sets of the same cardinality. In [35] Hamming spheres (or Hamming balls) of any cardinality are shown to be nearly stablest if the noise is very small, and it is conjectured that this should hold for any level of noise.

– The hypercontractive inequality (4) was used in [3] to obtain bounds on the distance components and other parameters of binary codes. We follow the approach of [3] in deriving some of

our results, such as the second of the claims above, but replacing (4) with a (stronger) inequality (9). We remark that the idea of using (4) to study the distance distribution of binary codes was introduced already in [18].

Organization of the paper

The remainder of this paper is organized as follows. We describe the relevant notions and provide some background in the next subsection. Our results are stated formally and discussed in Section 1.2. Somewhat unfortunately, the statements of the results involve certain functions of two variables, which will be defined later on in Section 2.1. This is done in order not to interrupt the flow of presentation.

We define several bivariate functions which play an important role in this paper and describe their pertinent properties in Section 2.1. Some properties of Krawchouk polynomials and Hamming spheres are described in Sections 2.2 - 2.4. These subsections also clarify the relevance of some of the bivariate functions defined in Section 2.1, by showing them to describe aspects of behavior of Krawchouk polynomials or of Hamming spheres.

Theorems 1.5 to 1.10 and some related results are derived from Theorem 1.3 in Section 3. Theorem 1.3 itself is proved in Section 4. This paper contains many auxiliary results describing the behaviour of various univariate and bivariate functions. The proofs of these results are relegated to the Appendix.

Let us suggest that (most of) Section 2 and the Appendix are better viewed as reference sections, written as laundry lists of results, and suitable for lookup, rather than for continuous reading.

1.1 Background, definitions, and notation

We view $\{0,1\}^n$ as a metric space, with the Hamming distance between $x, y \in \{0,1\}^n$ given by $|x-y| = |\{i : x_i \neq y_i\}|$. The *Hamming sphere* of radius r centered at x is the set $S(x,r) = \{y \in \{0,1\}^n : \ |x-y| = r\}$. The *Hamming ball* of radius r centered at x is the set $B(x,r) = \{y \in \{0,1\}^n : \ |x-y| \leq r\}$. Clearly, for any $x \in \{0,1\}^n$ and $0 \leq r \leq n$ holds $|S(x,r)| = \binom{n}{r}$ and $|B(x,r) = \sum_{k=0}^{r} \binom{n}{k}$.

Let $H(t) = t \log_2 \left(\frac{1}{t}\right) + (1-t) \log_2 \left(\frac{1}{1-t}\right)$ be the binary entropy function. We will make repeated use of the following sequence of estimates (the first estimate follows from the Stirling formula, for the second estimate see e.g., Theorem 1.4.5. in [31]): For $x \in \{0,1\}^n$ and $0 < r \leq \frac{n}{2}$ holds

$$\Theta\left(\sqrt{\frac{n}{r(n-r)}}\right) \cdot 2^{H\left(\frac{r}{n}\right) \cdot n} \ = \ |S(x,r)| \ \leq \ |B(x,r)| \ \leq \ 2^{H\left(\frac{r}{n}\right) \cdot n}. \tag{3}$$

The asymptotic notation will always hide absolute constants (unless specifically stated otherwise).

1.1.1 Distance distribution, edge-isoperimetry, binary codes

The distance distribution of a subset A of $\{0,1\}^n$ is given by the vector $(a_0, a_1, ...a_n)$ with $a_i = |\{(x,y) \in A \times A, |x-y| = i\}|$. That is, $a_i = a_i(A)$ counts the pairs of points at distance i in A. The distance distribution captures many important properties of a subset.

Edge-Isoperimetry. For $1 \leq i \leq n$, let G_i be the graph with vertices indexed by $\{0,1\}^n$, in which two vertices are connected by an edge iff the Hamming distance between them is i. In particular, G_1 is the usual graph of the boolean cube. The *edge-isoperimetric problem* (see [6] for a survey on discrete isoperimetry) in a graph G asks for a subset of vertices of a given cardinality, such that the number of edges crossing from this subset to its complement is as small as possible. If G is regular, this is the same as maximizing the number of edges in an induced subgraph of G with a given number of vertices. Note that a subset A of vertices of G_i, this number is given by $a_i(A)$. The edge-isoperimetric problem has been resolved for $i = 1$ [12, 15], in which case the solution to the problem is the initial segment of the lexicographic ordering on the cube. To the best of our knowledge, the problem is still open for any $i > 1$.

Undetected error probability. A *binary symmetric channel* (see e.g., [9]), with crossover probability $0 \leq \epsilon \leq 1/2$ is a communication channel which on input $x \in \{0,1\}^n$ outputs a random vector $y \in \{0,1\}^n$ obtained by flipping each bit of x independently, with probability ϵ. Given a binary code $C \subseteq \{0,1\}^n$, the undetected error probability [21] of C is the average probability (over the codewords) that a codeword transmitted over a binary symmetric channel is distorted in such a way that the received word, though different from the transmitted one, also belongs to the code. It is easy to see that this can be expressed in terms of the distance distribution of C:

$$P_{\mathrm{ue}}(C,\epsilon) \;=\; \frac{1}{|C|} \cdot \sum_{i=1}^{n} a_i(C) \epsilon^i (1-\epsilon)^{n-i}.$$

The *worst asymptotic undetected error exponent* for codes of rate $0 \leq R \leq 1$ and crossover probability ϵ was defined in [3] as

$$P_{\mathrm{ue}}(R,\epsilon) \;=\; \mathrm{limsup}_{n\to\infty} \left(\frac{1}{n} \max_{C} \log_2 \left(P_{\mathrm{ue}}(C,\epsilon) \right) \right),$$

where the maximum is taken over all codes $C \subseteq \{0,1\}^n$ of cardinality at most 2^{Rn}.

Binary error-correcting codes. A binary error-correcting code C of length n and minimal distance d is a subset of $\{0,1\}^n$ such that the Hamming distance between any two distinct points in C is at least d. This is clearly equivalent to $a_1 = ... = a_{d-1} = 0$. The problem of finding the largest possible code with a given minimal distance is open. In [10] a family of linear inequalities holding for the distance distribution vector of any binary code were obtained. These inequalities play a key role in the linear programming relaxation of this problem [10], which led to the best known upper bounds [32] on the cardinality of a code with a given minimal distance.

The *largest asymptotic distance component rate* (see e.g., [2, 3]) of a code with given rate and minimal distance is defined for $0 \leq \mu, \delta \leq \frac{1}{2}$ and $0 \leq R \leq 1$ as

$$b_\mu(R,\delta) \;=\; \mathrm{limsup}_{n\to\infty} \left(\frac{1}{n} \max_{C} \log_2 \left(a_{\lfloor \mu n \rfloor}(C) \right) \right),$$

where the maximum is taken over all codes $C \subseteq \{0,1\}^n$ of cardinality at most 2^{Rn} and minimal distance at least δn.

1.1.2 Fourier analysis, polynomials, noise operators, and spectral projections

We recall some basic notions in Fourier analysis on the boolean cube (see [34]). For $\alpha \in \{0,1\}^n$, define the Walsh-Fourier character W_α on $\{0,1\}^n$ by setting $W_\alpha(y) = (-1)^{\sum \alpha_i y_i}$, for all $y \in \{0,1\}^n$. The *weight* of the character W_α is the Hamming weight $|\alpha|$ of α. The characters $\{W_\alpha\}_{\alpha \in \{0,1\}^n}$ form an orthonormal basis in the space of real-valued functions on $\{0,1\}^n$, under the inner product $\langle f, g \rangle = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x)g(x)$. The expansion $f = \sum_{\alpha \in \{0,1\}^n} \widehat{f}(\alpha) W_\alpha$ defines the Fourier transform $\widehat{f}$ of f. We also have the Parseval identity, $\|f\|_2^2 = \sum_{\alpha \in \{0,1\}^n} \widehat{f}^2(\alpha)$. One additional simple fact will be used several times in this paper: Let $g(x) = (-1)^{|x|} f(x)$. Then, writing $\bar{\alpha} = \alpha \oplus 1$ for the complement of a vector $\alpha \in \{0,1\}^n$, for all $\alpha \in \{0,1\}^n$ holds $\widehat{f}(\alpha) = \widehat{g}\,(\bar{\alpha})$.

Polynomials on $\{0,1\}^n$. A function f on $\{0,1\}^n$ is a called a polynomial of degree s, for some $0 \leq s \leq n$, if f belongs to the span of Walsh-Fourier characters of weight at most s. Alternatively, for $1 \leq i \leq n$, let r_i be the Walsh-Fourier character of weight 1, corresponding to $\alpha = \{i\}$. The functions r_i are known as the *Rademacher functions* on $\{0,1\}^n$. Then f is a polynomial of degree s if and only if f is a multilinear polynomial of degree s in $r_1, ..., r_n$.

A function f is a homogeneous polynomial of degree s on $\{0,1\}^n$ if f is a homogeneous multilinear polynomial of degree s in $r_1, ..., r_n$. Such functions are also called *Rademacher chaos of order s*. Note that if f is a homogeneous polynomial of degree s and $g(x) = (-1)^{|x|} f(x)$, then g is a homogeneous polynomial of degree $n - s$. In particular, the spaces of homogeneous polynomials of degrees s and $n - s$ are isometric for any ℓ_p norm on $\{0,1\}^n$.

Krawchouk polynomials. For $0 \leq s \leq n$, let F_s be the sum of all Walsh-Fourier characters of weight s, that is $F_s = \sum_{|\alpha|=s} W_\alpha$. Note that F_s is the Fourier transform of $2^n \cdot 1_S$, where S is the Hamming sphere of radius s around 0. It is easy to see that $F_s(x)$ depends only on the Hamming weight $|x|$ of x, and it can be viewed as a univariate function on the integer points $0, ..., n$, given by the restriction to $\{0, ..., n\}$ of the univariate polynomial $K_s = \sum_{k=0}^{s} (-1)^k \binom{x}{k} \binom{n-x}{s-k}$ of degree s. That is, $F_s(x) = K_s(|x|)$. The polynomial K_s is the s^{th} *Krawchouk polynomial*. Abusing notation, we will also call F_s the s^{th} Krawchouk polynomial, and write K_s for F_s when the context is clear.

Spectral projections. For $0 \leq k \leq n$ we define Π_k to be the orthogonal projection to the subspace spanned by Walsh-Fourier characters of weight k. (This is the eigenspace of the Laplacian of the discrete cube corresponding to eigenvaleue $2k$.) That is, for a function f on $\{0,1\}^n$, and $0 \leq k \leq n$, we have $\Pi_k f = \sum_{|\alpha|=k} \widehat{f}(\alpha) W_\alpha$. We will also write f_k for $\Pi_k f$ for ease of notation.

The noise operator. Given a noise parameter $0 \leq \epsilon \leq 1/2$, the noise operator T_ϵ is a linear operator acting on functions on the boolean cube as follows: for $f : \{0,1\}^n \to \mathbb{R}$, $T_\epsilon f$ at a point x is the expected value of f at y, where y is "$(1-\epsilon)$-correlated" with x. That is, y is a random binary vector whose i^{th} coordinate is x_i with probability $1 - \epsilon$ and $1 - x_i$ with probability ϵ, independently for different coordinates. In other words, $(T_\epsilon f)(x) = \mathbb{E}_y f(y)$, where y is the output of the binary symmetric channel on input x. Writing this out explicitly,

we have $(T_\epsilon f)(x) = \sum_{y\in\{0,1\}^n} \epsilon^{|y-x|}(1-\epsilon)^{n-|y-x|} f(y)$. The noise operators form a semigroup: $T_{\epsilon_1} T_{\epsilon_2} = T_{\epsilon_1+\epsilon_2-2\epsilon_1\epsilon_2}$. We will also write f_ϵ for $T_\epsilon f$, for brevity. It is easy to see that $\widehat{f_\epsilon}(\alpha) = (1-2\epsilon)^{|\alpha|}\widehat{f}(\alpha)$, which means that $T_\epsilon = \sum_{k=0}^n (1-2\epsilon)^k \Pi_k$.

1.1.3 Hypercontractive inequalities

Hypercontractive inequalities [8, 11, 4] form a family of analytic inequalities for functions on $\{0,1\}^n$, with many applications in discrete mathematics, information theory, and theoretical computer science, see e.g., [17, 33, 25], and also the monograph [34] and the references there.

Let the ℓ_p norm of a real-valued function f on $\{0,1\}^n$ be given by $\|f\|_p = \left(\frac{1}{2^n}\sum_{x\in\{0,1\}^n} |f(x)|^p\right)^{1/p}$. Hypercontractive inequalities assert that applying noise to a function flattens it in a well-defined sense: a higher norm of the noisy function is upperbounded by a lower norm of the original function. A useful special case is the one involving the ℓ_2 norm, (since this norm is easy to work with in applications). The inequality in this case is

$$\|f_\epsilon\|_2 \le \|f\|_{1+(1-2\epsilon)^2}. \tag{4}$$

It is easy to see that if f is a characteristic function of a subset $A \subseteq \{0,1\}^n$, then $\|f_\epsilon\|_2^2 = \frac{1}{2^n}\sum_{i=0}^n a_i(A)\delta^i(1-\delta)^{n-i}$, for $\delta = 2\epsilon(1-\epsilon)$. The relevance of (4) to the study of distance distributions of binary codes has been pointed out in [18]. In [3] this inequality was used to obtain new bounds on the distance distribution, the undetected error probability, and other related parameters of binary codes of a given cardinality and minimal distance.

Stronger hypercontractive inequalities for highly concentrated functions. While (4) is known to be essentially tight for functions which are almost flat to begin with, stronger hypercontractive inequalities were proved in [40] for functions f on $\{0,1\}^n$ for which the ratio $\frac{\|f\|_p}{\|f\|_1}$, for some $p > 1$, is exponentially large in n.

An uncertainty theorem. Strong hypercontractive inequalities for highly concentrated functions were used in [40] to obtain a tight uncertainty-type result for $\{0,1\}^n$. Let a non-zero function f be supported on a set $A \subseteq \{0,1\}^n$, with cardinality of A being at most that of a Hamming ball of radius $\rho_1 n$ (for some $0 < \rho_1 < \frac{1}{2}$). In fact, it suffices to assume, more generally, that the ratio $\frac{\|f\|_2}{\|f\|_1}$ is lower-bounded by $2^{\frac{1-H(\rho_1)}{2}\cdot n}$. Then $\widehat{f}$ attains only an exponentially small fraction of its ℓ_2 norm on any Hamming ball of radius $\rho_2 n$, provided $\rho_2 < \frac{1}{2} - \sqrt{\rho_1(1-\rho_1)}$.

1.1.4 Tensorization

We describe a useful and well-known tool in analysis which will be used several times in this paper. Let f be a function on $\{0,1\}^n$. For an integer $m \ge 1$, the m^{th} tensor power $F_m := f^{\otimes m}$ is a function on nm boolean variables defined for $x_1, ..., x_m \in \{0,1\}^n$ by $F_m(x_1, ..., x_m) = \prod_{i=1}^m f(x_i)$. We recall some useful properties of tensor powers:

– For any $\alpha_1, ..., \alpha_m \in \{0,1\}^n$ holds $\widehat{F_m}(\alpha_1, ..., \alpha_m) = \prod_{i=1}^m \widehat{f}(\alpha_i)$. In particular, if f is a homogeneous polynomial of degree s, then F_m is a homogeneous polynomial of degree sm; and

if f is a (not necessarily homogeneous) polynomial of degree s, then F_m is a polynomial of degree sm;

– For any $q \in \mathbb{R}$ holds $\mathbb{E}\,|F_m|^q = (\mathbb{E}\,|f|^q)^m$.

1.2 Our results

1.2.1 Upper bounds for moments of polynomials

We show that for any $p \geq 2$ and for any $1 \leq s \leq \frac{n}{2}$, the s^{th} Krawchouk polynomial $K_s = \sum_{|\alpha|=s} W_\alpha$ attains, within a relatively small error, the maximal ratio of $\frac{\|f\|_p}{\|f\|_2}$ among all homogeneous polynomials of degree s. Let $\psi(p,x)$ be the function defined in Subsection 2.1.4.

Theorem 1.3: *For any $p \geq 2$, $0 \leq s \leq \frac{n}{2}$, and for any homogeneous polynomial f of degree s on $\{0,1\}^n$ holds*

$$\frac{\mathbb{E}\,|f|^p}{(\mathbb{E}\,f^2)^{\frac{p}{2}}} \;\leq\; 2^{\psi\left(p,\frac{s}{n}\right)\cdot n}. \tag{5}$$

There is an absolute constant $C > 0$ such that for any $p \geq 2$ and $0 \leq s \leq \frac{n}{2}$ holds

$$2^{\psi\left(p,\frac{s}{n}\right)\cdot n} \;\leq\; n \cdot C^p \cdot s^{\frac{p}{4}} \cdot \frac{\mathbb{E}\,|K_s|^p}{(\mathbb{E}\,K_s^2)^{\frac{p}{2}}}.$$

Discussion.

– The assumption that f is homogeneous is not necessary. In fact, we have, as a simple corollary of (5):

Corollary 1.4: *The upper bound (5) holds for general polynomials of degree as most s.*

– As mentioned above, if p is an even integer, it seems possible to extend the argument given in [1] for $p = 4$ and to show that Krawchouk polynomials actually attain the maximum for $\frac{\mathbb{E}\,|f|^p}{(\mathbb{E}\,f^2)^{\frac{p}{2}}}$ among all homogeneous polynomials of the same degree.

– The inequality (5) is a *Khintchine-type inequality*. Recall that Khintchine-type inequalities establish an upper bound on the ratio of two ℓ_p norms for functions coming from a certain restricted domain, typically a space of multivariate polynomials of a specified degree over a given product space. In particular, the prototypical Khintchine inequality [20] states that the ratio of ℓ_2 and ℓ_1 norms of linear polynomials over the boolean cube $\{0,1\}^n$ is bounded by an absolute constant. See [16] for a recent discussion and references. Viewed in this context, Theorem 1.3 states that for any $p > 2$ the "Khintchine ratio" $\frac{\|f\|_p}{\|f\|_2}$ for polynomials of a given degree on the boolean cube is maximized, up to a small error, on the Krawchouk polynomial of this degree.

– It is easy to see that Theorem 1.3 essentially determines the $\|\cdot\|_{2\to p}$ norm of the spectral projection operator Π_k (see [36] where the norms of these operators are investigated).

1.2.2 Tail bounds for polynomials

We show that Krawchouk polynomials have (almost) the heaviest tails among all polynomials of same degree and ℓ_2 norm. Let τ be the function defined in Subsection 2.1.2. Let H be the binary entropy function.

To make the statement of the second part of the following theorem more legible, recall (see Section 2.2) that the Krawchouk polynomial K_s on $\{0,1\}^n$ has all its roots in the interval $\left[\frac{n}{2} - \sqrt{s(n-s)}, \frac{n}{2} + \sqrt{s(n-s)}\right]$, and that the distance between any two consecutive roots is $o(n)$.

Theorem 1.5:

Let f be a polynomial of degree $s \leq \frac{n}{2}$ on $\{0,1\}^n$. Then for all $0 \leq i \leq \frac{n}{2}$ holds

$$\Pr\left\{|f| \geq \|f\|_2 \cdot 2^{\left(\tau\left(\frac{s}{n},\frac{i}{n}\right) - \frac{1}{2}H\left(\frac{s}{n}\right)\right)\cdot n}\right\} \;\leq\; 2^{\left(H\left(\frac{i}{n}\right)-1\right)\cdot n}. \tag{6}$$

Moreover, for $f = K_s$ we have:

- *For any $0 \leq i \leq \frac{n}{2} - \sqrt{s(n-s)}$ holds*

$$\Pr\left\{|K_s| \geq \sqrt{\frac{\binom{n}{s}}{2^{H\left(\frac{s}{n}\right)\cdot n}}} \cdot \|K_s\|_2 \cdot 2^{\left(\tau\left(\frac{s}{n},\frac{i}{n}\right) - \frac{1}{2}H\left(\frac{s}{n}\right)\right)\cdot n}\right\} \;\geq\; \Omega\left(\frac{1}{\sqrt{i}}\right) \cdot 2^{\left(H\left(\frac{i}{n}\right)-1\right)\cdot n}.$$

- *Between any two consecutive roots of K_s there is a point i for which*

$$\Pr\left\{|K_s| \geq \Omega\left(\frac{1}{n^{5/2}}\right) \cdot \|K_s\|_2 \cdot 2^{\left(\tau\left(\frac{s}{n},\frac{i}{n}\right) - \frac{1}{2}H\left(\frac{s}{n}\right)\right)\cdot n}\right\} \;\geq\; \Omega\left(\frac{1}{\sqrt{i}}\right) \cdot 2^{\left(H\left(\frac{i}{n}\right)-1\right)\cdot n}.$$

Discussion

– Note that, by (3), the correction factor $\sqrt{\frac{\binom{n}{s}}{2^{H\left(\frac{s}{n}\right)\cdot n}}}$ is $\Theta\left(s^{\frac{1}{4}}\right)$.

– A polynomial f of degree s is a linear combination of $\sum_{k=0}^{s}\binom{n}{k}$ Walsh-Fourier characters, which are orthonormal, and all of which evaluate to 1 at 0. It is easy to see that this implies that $\frac{\|f\|_\infty}{\|f\|_2} \leq \sqrt{\sum_{k=0}^{s}\binom{n}{k}} \leq 2^{\frac{1}{2}H\left(\frac{s}{n}\right)\cdot n}$, with equality attained for $f = \sum_{|\alpha|\leq s} W_s = \sum_{k=0}^{s} K_k$. On the other hand, the function $\tau\left(\frac{s}{n},\frac{i}{n}\right) - \frac{1}{2}H\left(\frac{s}{n}\right)$ decreases from $\frac{1}{2}H\left(\frac{s}{n}\right)$ to 0 as i does from 0 to $\frac{n}{2}$. Hence (6) provides tail estimates for the whole range of values of $|f|$.

– The bound (6) is a pointwise improvement over the estimate $\Pr\{|f| \geq \|f\|_2 \cdot t\} \leq e^2 \cdot e^{-\left(\frac{t}{e}\right)^{\frac{1}{k}}}$, due to [5, 43]. To see this, note that the latter bound was obtained by applying the inequality $\Pr\{|f| \geq T\} \leq \min_{p\geq 2}\left\{\frac{\mathbb{E}|f|^p}{T^p}\right\}$, using (2) to bound the RHS, and choosing a suitable value of p. The proof of (6) uses the same approach, while replacing (2) with a stronger bound (1) and choosing the *optimal* value of p.

– There is a gap between the upper bound (6) and the lower bound provided by Krawchouk polynomials, due to the correction factor $\sqrt{\frac{\binom{n}{s}}{2^{H(\frac{s}{n})\cdot n}}}$ in the value of the threshold. One can ask how accurate (6) is for Krawchouk polynomials, if there is no correction factor. It's not hard to see, using the properties of the function τ, that e.g., for i growing linearly with n, that is for t bounded away from $\frac{\|K_s\|_\infty}{\|K_s\|_2}$, the error in the estimate of (6) for the probability of $|K_s| \geq \|K_s\|_2 \cdot t$ is of order at most $2^{O\left(\frac{\log(s)}{s}\right)\cdot n}$. In particular, if s is an increasing function of n, (6) provides the right constant in the large deviation inequalities for homogeneous polynomials of degree s (that is, Rademacher chaos of order s).

1.2.3 An isoperimetric-type inequality

Recall that G_i is a graph with vertices indexed by $\{0,1\}^n$, in which two vertices are connected by an edge iff the Hamming distance between them is i. We prove an edge-isoperimetric inequality for the graphs $\{G_i\}_i$, and show this inequality to be somewhat tight for the Hamming sphere, or a union of two adjacent spheres, depending on the parity of i.

Let $A \subseteq \{0,1\}^n$. Recall that $a_i(A) = |\{(x,y) \in A \times A, |x-y| = i\}|$ is the i^{th} distance component of A.

Theorem 1.6:

Let $A \subseteq \{0,1\}^n$, with $|A| \leq 2^{H(\sigma)n}$, for some $0 \leq \sigma \leq \frac{1}{2}$. Then for $1 \leq i \leq 2\sigma(1-\sigma)n$ holds

$$a_i(A) \quad \leq \quad |A| \cdot 2^{\left(\sigma H\left(\frac{i}{2\sigma n}\right)+(1-\sigma)H\left(\frac{i}{2(n-\sigma n)}\right)\right)\cdot n}. \tag{7}$$

Let $s = \sigma n$, and assume s to be integer. If i is even, this inequality is tight, up to a factor of $O(i)$, if A a Hamming sphere of radius s. For an arbitrary i, this is tight, up to a factor of $O\left(\sqrt{\frac{n-s}{s}} \cdot i\right)$, if A is the union of two adjacent spheres of dimension $n-1$ and radii $s-1$ and s.

Discussion

– For $2\sigma(1-\sigma)n \leq i \leq \frac{n}{2}$, the distance component $a_i(A)$ could be (essentially) as large as $|A|^2$. For $i > \frac{n}{2}$, the bounds on a_i reduce to these on a_{n-i}. See Remark 3.3.

– The proof of (7) follows the argument in [3], replacing the hypercontractive inequality (4) used in [3] with a stronger inequality proved in Corollary 3.2 (which is a special case of (9)).

– Choosing $i = 2$ in (7), we get the following claim:

Corollary 1.7: *For $A \subseteq \{0,1\}^n$ with $|A| \leq \binom{n}{s}$ holds $a_2(A) \leq e^2 s(n-s) \cdot |A|$.*

This is tight, within a factor of e^2, if A is a Hamming sphere of radius s.

Let us also consider this bound in the context of the Kleitman-West problem (see e.g., [13]). This is the edge-isoperimetric problem for the Hamming sphere (see Section 1.1.1). One way to

pose this problem is as follows. Given the dimension n and the radius $0 \leq r \leq \frac{n}{2}$ of the sphere $S = S(0,r) \subseteq \{0,1\}^n$, determine how large can $a_2(A)$ be for a subset A of S of a given size.

Let $s \leq r$ and let $|A| = \binom{n}{s}$. Let A be an $\big(n-(r-s)\big)$-dimensional Hamming sphere of radius s embedded in S, by concatenating $r-s$ coordinates to points in A and setting them to be 1. Then $a_2(A) = s(n-r) \cdot |A|$, so the bound in Corollary 1.7 is tight for A within a factor of $2e^2$. For $s < \frac{r}{C}$, where C is a sufficiently large constant, this bound seems to improve the best known upper bounds on $a_2(A)$ for $A \subseteq S$ (which, to the best of our knowledge, come from the logarithmic Sobolev inequality for the Hamming sphere [27]).

– Recall (see Section 1.1.1) that an undetected error probability of a binary code $C \subseteq \{0,1\}^n$ is given by $P_{\text{ue}}(C,\epsilon) = \frac{1}{|C|} \cdot \sum_{i=1}^{n} a_i(C)\epsilon^i(1-\epsilon)^{n-i}$. Theorem 1.6 implies (as shown in Section 3.2) that a union of two adjacent spheres of dimension $n-1$ and radii $s-1$ and s maximizes the undetected error probability for all codes of cardinality at most $2^{H\left(\frac{s}{n}\right)\cdot n}$, up to at most a polynomial in n factor. A simple consequence of this fact is the following expression for the worst asymptotic undetected error exponent: For all $0 < R \leq 1$ and $0 < \epsilon \leq \frac{1}{2}$ holds

$$P_{\text{ue}}(R,\epsilon) \;=\; \sigma H\left(\frac{x}{\sigma}\right) + (1-\sigma)H\left(\frac{x}{1-\sigma}\right) + 2x\log_2(\epsilon) + (1-2x)\log_2(1-\epsilon), \tag{8}$$

where $\sigma = H^{-1}(R)$ and $x = x(\sigma,\epsilon) = \frac{-\epsilon^2+\epsilon\sqrt{\epsilon^2+4(1-2\epsilon)\sigma(1-\sigma)}}{2(1-2\epsilon)}$.

1.2.4 A hypercontractive inequality

We prove a nearly tight hypercontractive inequality for functions on $\{0,1\}^n$, which takes into account the distribution of a function, specifically the ratio between its ℓ_p and ℓ_1 norms. (See [40] for a different family of hypercontractive inequalities taking into account the ratio between ℓ_p and ℓ_1 norms of a function.)

Let η be the function defined in Section 2.1.7. Recall that $\eta(x,\epsilon) < 0$ for all $0 < \epsilon < \frac{1}{2}$ and $0 < x \leq \frac{(1-2\epsilon)^2}{1+(1-2\epsilon)^2}$. Moreover, η is concave and decreasing in x for any $0 \leq \epsilon \leq \frac{1}{2}$.

For a function f on $\{0,1\}^n$ and $1 \leq p \leq \infty$ let $r(p) = r_f(p) = \frac{1}{n}\log_2\left(\frac{\|f\|_p}{\|f\|_1}\right)$. Note that $0 \leq r(p) \leq \frac{p-1}{p}$.

Theorem 1.8:

Let f be a function on $\{0,1\}^n$, and let $0 \leq \epsilon \leq 1/2$. Then

$$\|f_\epsilon\|_2 \quad\leq\quad 2^{\eta\left(r\left(1+(1-2\epsilon)^2\right),\,\epsilon\right)\cdot n} \cdot \|f\|_{1+(1-2\epsilon)^2}. \tag{9}$$

This is tight up to a factor of $O\left(s^{3/4}\right)$ if f is proportional to a characteristic function of a Hamming sphere of radius s.

Discussion.

– (9) is a strengthening of the hypercontractive inequality (4). However, since (as is easily seen) $\frac{\partial \eta(x,\epsilon)}{\partial x}_{|x=0} = 0$, this improvement is significant only if $\frac{\|f\|_{1+(1-2\epsilon)^2}}{\|f\|_1} \geq 2^{\Omega(\sqrt{n})}$.

– We have the following corollary of (9), extending it to other ℓ_p norms. For $p \geq 1$, let η_p be the function defined in Section2.1.7.

Corollary 1.9: *Let f be a function on $\{0,1\}^n$, let $0 \leq \epsilon \leq 1/2$ and let $p \geq 1 + (1-2\epsilon)^2$. Then*

$$\|f_\epsilon\|_2 \quad \leq \quad 2^{\eta_p(r(p),\epsilon)\cdot n} \cdot \|f\|_p.$$

As will be seen in the proof of this result, the RHS of the inequality above increases in p, so it is in general weaker than (9). However, it is still tight up to a factor of $O\left(s^{3/4}\right)$ if f is proportional to the characteristic function of a Hamming sphere of radius s. To see this, note that the RHS of this inequality does not depend on p, if f is a characteristic function of a set. Hence, this corollary can be rephrased as follows: For any $0 \leq \delta \leq \frac{1}{2}$ and $p \geq 2 - 2\delta$, a characteristic function of a Hamming sphere of radius s maximizes, within a factor of $O\left(s^{3/4}\right)$, the inner product $\langle f_\delta, f\rangle$ among all functions with the same ℓ_1 and ℓ_p norms[4].

– Let us mention a more general conjecture [37]: For any $q > 1$ and a threshold value $t = t(q,\epsilon)$ there exists $p_0 = p_0(q,\epsilon,t) \leq 1 + (1-2\epsilon)^2(q-1)$ such that for any $p \geq p_0$ the maximum of the ratio $\frac{\|f_\epsilon\|_q}{\|f\|_p}$ over all functions f on $\{0,1\}^n$ with $r(p) \geq t$ is essentially attained at the characteristic function of a Hamming sphere of an appropriate radius.

– Upper bounds on the asymptotic distance component rates $b_\mu(R,\delta)$ of binary codes with given rate and minimal distance (see Section 1.1.1) were obtained in [3] using the hypercontractive inequality (4). These bounds can be improved by using (9) instead of (4), similarly to the improvement obtained in (7) over the bounds of [3] for $b_\mu(R,0)$. We do not go into details, since the bounds, both in [3] and here, are not explicit, but rather given, for each R and δ, as the minimal value of a certain explicit function in a constant number of variables (three in [3] and two in our case) over its domain.

1.2.5 An uncertainty theorem

We give an extension of an uncertainty-type result from [40] (see Section 1.1.3). Let π be the function defined in Section 2.1.5. Recall that for a function f on $\{0,1\}^n$ and $0 \leq k \leq n$ we write f_k for the orthogonal projection of f on the space of Walsh-Fourier characters of weight k. We write $x \wedge y$ for the minimum of x and y and, as above, given a function f on $\{0,1\}^n$, write $r(p)$ for $\frac{1}{n}\log_2\left(\frac{\|f\|_p}{\|f\|_1}\right)$.

Theorem 1.10:

Let f be a function on $\{0,1\}^n$. Then for any $p \geq 2$ and $0 \leq k \leq n$ holds

$$\|f_k\|_2 \quad \leq \quad 2^{\left(\pi\left(\frac{k}{n}\wedge\frac{n-k}{n}, H^{-1}\left(1-\frac{p}{p-1}\cdot r(p)\right)\right)-\frac{p-2}{2p-2}\cdot r(p)\right)\cdot n} \cdot \|f\|_p. \tag{10}$$

[4]Recall that $\|f_\epsilon\|_2^2 = \langle f_\delta, f\rangle$, for $\delta = 2\epsilon(1-\epsilon)$.

Let f be proportional to a characteristic function of a Hamming sphere of radius s. Then this inequality is tight up to a factor of $O\left((ks)^{1/4}\right)$ for $0 \leq k \leq \frac{n}{2} - \sqrt{s(n-s)}$ and $\frac{n}{2} + \sqrt{s(n-s)} \leq k \leq n$. In addition, between any two consecutive roots of the Krawchouk polynomial K_s there is a point k for which this is tight up to a factor of $O\left(n^{5/2}\right)$.

Discussion.

– An alternative (somewhat imprecise) way to phrase this result is as follows: For any $p \geq 2$, characteristic functions of Hamming spheres have (almost) the largest spectral projections among all functions with the same ℓ_1 and ℓ_p norms.

– Let $\frac{\|f\|_2}{\|f\|_1} = 2^{\frac{1-H(\rho)}{2} \cdot n}$, for some $0 < \rho < \frac{1}{2}$. Then the theorem with $p = 2$ implies $\|f_k\|_2 \leq 2^{\pi\left(\rho, \frac{k}{n}\right) \cdot n} \cdot \|f\|_2$. In particular, for $\frac{k}{n}$ bounded away from below from $\frac{1}{2} - \sqrt{\rho\,(1-\rho)}$, this implies that $\|f_k\|_2$ is exponentially smaller than $\|f\|_2$ (since $\pi\left(\rho, \frac{k}{n}\right)$ is negative in this range, see Lemma 2.8), recovering the result in [40]. Furthermore, we get a quantitative upper bound on the exponent of the ratio $\frac{\|f_k\|_2}{\|f\|_2}$ for $0 < \frac{k}{n} \leq \frac{1}{2} - \sqrt{\rho\,(1-\rho)}$.

2 Bivariate functions, Krawchouk polynomials, and Hamming spheres

2.1 Some bivariate functions

Section 1.2 describes some functional inequalities on the Hamming cube. These inequalities involve certain functions of two variables. A good way to come to terms with these functions is to realize that they describe various aspects of the behavior of Hamming spheres or of Krawchouk polynomials (see Sections 2.2 and 2.4). In this subsection we define these functions and list their relevant properties.

2.1.1 The function I

For $0 \leq x \leq \frac{1}{2}$ and $0 < y \leq \frac{1}{2} - \sqrt{x(1-x)}$, let

$$I(x,y) \;=\; \log_2(1-x) + \frac{a}{2}\log_2\left(1-2x-b\right) + x\log_2\left(\frac{a+b}{2(1-x)}\right) - \frac{1}{2}\log_2\left(2(1-x) - a^2 - ab\right),$$

where $a = 1-2y$ and $b = \sqrt{a^2 - 4x(1-x)}$. Extend this by continuity to $y = 0$ by setting $I(x,0) = -1$ for all $0 \leq x \leq \frac{1}{2}$.

Let $r(x,y) = \frac{(1-2x)+\sqrt{(1-2x)^2 - 4y(1-y)}}{2(1-y)}$ for For $0 \leq x \leq \frac{1}{2}$ and $0 \leq y \leq \frac{1}{2} - \sqrt{x(1-x)}$. Then ([18], with a correction in [23]) I is an indefinite integral of r, that is $\int_0^y \log_2\left(r(x,z)\right) dz = I\left(y,x\right) - I\left(0,x\right)$.

Lemma 2.1: *For a fixed $0 \leq x \leq \frac{1}{2}$, the function $r(x,y)$ decreases in y. In particular, for any $y \geq 0$ holds $r(x,y) \leq r(x,0) = 1-2x$.*

Corollary 2.2: *For a fixed $0 \leq x \leq \frac{1}{2}$, the function $I(y,x)$ is decreasing and concave in y.*

2.1.2 The function τ

For $0 \leq x, y \leq \frac{1}{2}$, let

$$\tau(x,y) \;=\; \begin{cases} H(x) + I(y,x) - I(0,x) & \text{if} \quad y \leq \frac{1}{2} - \sqrt{x(1-x)} \\ \frac{1+H(x)-H(y)}{2} & \text{otherwise} \end{cases}$$

It is easy to verify that τ is continuous in both variables. Using the results in Subsection 2.1.1, we see that $\frac{\partial \tau(x,y)}{\partial y} = \begin{cases} \log_2(r(x,y)) & \text{if} \quad y < \frac{1}{2} - \sqrt{x(1-x)} \\ \frac{1}{2}\log_2\left(\frac{y}{1-y}\right) & \text{if} \quad \frac{1}{2} - \sqrt{x(1-x)} < y < \frac{1}{2} \end{cases}$. This means that $\tau(x,y)$ is decreasing and concave in y on $0 \leq y \leq \frac{1}{2} - \sqrt{x(1-x)}$, and is decreasing and convex in y afterwards.

Lemma 2.3:

For all $0 \leq x, y \leq \frac{1}{2}$ *holds*

$$H(y) + \tau(x,y) \;=\; H(x) + \tau(y,x).$$

2.1.3 The function h

For $2 \leq p < \infty$ and $0 \leq x \leq \frac{1}{2}$, let

$$h(p,x) \;=\; x^{\frac{1}{p}}(1-x)^{\frac{p-1}{p}} + x^{\frac{p-1}{p}}(1-x)^{\frac{1}{p}}.$$

Lemma 2.4:

1. *For any* $p \geq 2$ *he function* $h(p,x)$ *increases from* 0 *to* 1*, as* x *goes from* 0 *to* $\frac{1}{2}$*.*
2. *For any* $0 < x < \frac{1}{2}$*, the function* $h(p,x)$ *increases from* $2\sqrt{x(1-x)}$ *to* 1*, as* p *goes from* 2 *to* ∞*.*
3. *Let* $0 < x < \frac{1}{2}$*. If* $p > 2$ *and* $h(p,y) = 1 - 2x$*, then* $y < \frac{1}{2} - \sqrt{x(1-x)}$*.*

2.1.4 The function ψ

For $2 \leq p < \infty$ and $0 \leq x \leq \frac{1}{2}$, let

$$\psi(p,x) \;=\; H(y) - 1 + p\tau(x,y) - \frac{p}{2}H(x),$$

where y is determined by $h(p,y) = 1 - 2x$.

The function ψ has another useful representation. First, we state an auxiliary lemma.

Lemma 2.5: *For any $p \geq 2$ the function $a(p,\delta) = \left(\frac{1}{2} - \delta\right) \cdot \frac{(1-\delta)^{p-1} - \delta^{p-1}}{(1-\delta)^p + \delta^p}$ decreases from $\frac{1}{2}$ to 0, as δ goes from 0 to $\frac{1}{2}$.*

Proposition 2.6: *We also have*

$$\psi(p,x) \;=\; (p-1) + \log_2\left((1-\delta)^p + \delta^p\right) - \frac{p}{2}H(x) - px\log_2(1-2\delta),$$

where δ is determined by $x = a(p,\delta)$. (Note that δ is well-defined by Lemma 2.5.)

Proposition 2.7:

1. *For any $p \geq 2$ and for any $x \geq 0$ holds $\psi(p,0) = \psi(2,x) = 0$.*
2. *The function $\psi(p,x)$ is increasing and (strongly) convex in p for any $x > 0$.*
3. *The function $\psi(p,x)$ is increasing and (strongly) concave in x for any $p > 2$. We also have $\frac{\partial\psi(p,x)}{\partial x}_{|x=0} = \frac{p\log_2(p-1)}{2}$.*

2.1.5 The function π

For $0 \leq x, y \leq \frac{1}{2}$, let

$$\pi(x,y) \quad = \quad \begin{cases} I(y,x) - I(0,x) + \frac{H(x)+H(y)-1}{2} & \text{if} \quad y \leq \frac{1}{2} - \sqrt{x(1-x)} \\ 0 & \text{otherwise} \end{cases}$$

Note that $\pi(x,y) = \tau(x,y) - \frac{1+H(x)-H(y)}{2}$. In particular, π is continuous in both variables.

Lemma 2.8: *The function π is symmetric, that is $\pi(x,y) = \pi(y,x)$ for all $0 \leq x, y \leq \frac{1}{2}$. It is strictly negative for $y < \frac{1}{2} - \sqrt{x(1-x)}$ and increasing in both arguments.*

Lemma 2.9: *For any $0 \leq \kappa, \sigma \leq \frac{1}{2}$ holds*

$$\pi(\sigma,\kappa) \;=\; \frac{1}{2}\min_{0\leq\delta\leq\frac{1}{2}}\left\{\sigma H\left(\frac{x}{\sigma}\right) + (1-\sigma)\,H\left(\frac{x}{1-\sigma}\right) + 2x\log_2(\delta) + (1-2x)\log_2(1-\delta) - \kappa\log_2(1-2\delta)\right\},$$

where $x = x(\sigma,\delta) = \frac{-\delta^2 + \delta\sqrt{\delta^2 + 4(1-2\delta)\sigma(1-\sigma)}}{2(1-2\delta)}$.

2.1.6 The functions ϕ and $\tilde{\phi}$

For $0 \leq \sigma, \epsilon \leq \frac{1}{2}$, let

$$\alpha_{\sigma,\epsilon}(x) \;=\; \sigma H\left(\frac{x}{\sigma}\right) + (1-\sigma)H\left(\frac{x}{1-\sigma}\right) + 2x\log_2(\epsilon) + (1-2x)\log_2(1-\epsilon).$$

Let

$$\phi(\sigma,\epsilon) \;=\; H(\sigma) - 1 + \max_{0\leq x\leq\sigma} \alpha_{\sigma,\epsilon}(x).$$

The value of x for which the maximum is attained is $x = x(\sigma,\epsilon) = \frac{-\epsilon^2+\epsilon\sqrt{\epsilon^2+4(1-2\epsilon)\sigma(1-\sigma)}}{2(1-2\epsilon)}$ (see e.g., [3])

For $0 \leq y \leq 1$ and $0 \leq \epsilon \leq \frac{1}{2}$, let

$$\tilde{\phi}(y,\epsilon) \;=\; \phi\left(H^{-1}(y),\epsilon\right).$$

We list some relevant properties of ϕ.

Lemma 2.10:

For all $0 \leq \sigma, \epsilon \leq \frac{1}{2}$ holds

$$\phi(\sigma,\epsilon) \;=\; \max_{0\leq y\leq\frac{1}{2}} \Big\{ y\log_2(1-2\epsilon) + H(y) + 2\tau(\sigma,y) \Big\} - 2.$$

Lemma 2.11: *Let $0 \leq \sigma \leq \frac{1}{2}$ and let $0 \leq y \leq 2\sigma(1-\sigma)$. Then*

$$\min_{0<\epsilon\leq\frac{1}{2}} \Big\{ \phi(\sigma,\epsilon) + 1 - H(\sigma) - y\log_2(\epsilon) - (1-y)\log_2(1-\epsilon) \Big\} = \sigma H\left(\frac{y}{2\sigma}\right) + (1-\sigma)H\left(\frac{y}{2(1-\sigma)}\right).$$

Lemma 2.12: *Let $\sigma = \frac{s}{n}$, s an integer between 0 and $\frac{n}{2}$. Let $0 \leq \epsilon \leq \frac{1}{2}$. Let $A = \max_{0\leq x\leq\sigma} \alpha_{\sigma,\epsilon}(x)$ and $B = \max_{0\leq i\leq s} \alpha_{\sigma,\epsilon}\left(\frac{i}{n}\right)$. Then $B \leq A \leq B + O\left(\frac{1}{n}\right)$, where the constant in the asymptotic notation is absolute.*

We also list some relevant properties of $\tilde{\phi}$.

Lemma 2.13:

The function $\tilde{\phi}(y,\epsilon)$ is (strictly) increasing and concave in y, for any fixed $0 \leq \epsilon \leq \frac{1}{2}$. Moreover, $\tilde{\phi}(1,\epsilon) = 0$, and the one-sided derivatives with respect to y of $\tilde{\phi}$ at the endpoints of the interval are $\frac{\partial\tilde{\phi}}{\partial y}(0,\epsilon) = 2$ and $\frac{\partial\tilde{\phi}}{\partial y}(1,\epsilon) = \frac{1}{1-\epsilon}$.

2.1.7 The functions η and η_p

For $0 \le \epsilon \le \frac{1}{2}$, and for $0 \le x \le \frac{(1-2\epsilon)^2}{1+(1-2\epsilon)^2}$, let

$$\eta(x,\epsilon) \;=\; \frac{1}{2}\tilde{\phi}\left(1 - \frac{1+(1-2\epsilon)^2}{(1-2\epsilon)^2}\cdot x,\; 2\epsilon(1-\epsilon)\right) + \frac{1}{(1-2\epsilon)^2}\cdot x,$$

where the function $\tilde{\phi}$ is defined in Section 2.1.6.

More generally, for $p > 1$, for $0 \le \epsilon \le \frac{1}{2}$, and for $0 \le x \le \frac{p-1}{p}$, let

$$\eta_p(x,\epsilon) \;=\; \frac{1}{2}\tilde{\phi}\left(1 - \frac{p}{p-1}\cdot x,\; 2\epsilon(1-\epsilon)\right) + \frac{1}{p-1}\cdot x.$$

Note that $\eta = \eta_{1+(1-2\epsilon)^2}$.

Lemma 2.14: *The function $\eta_p(x,\epsilon)$ is concave and decreasing in x for fixed p and ϵ satisfying $p \ge 1 + (1-2\epsilon)^2$. Moreover, if we also assume $0 < \epsilon < \frac{1}{2}$ then it is strictly negative for $0 < x \le \frac{p-1}{p}$.*

2.2 Krawchouk polynomials

Krawchouk polynomials were defined in Section 1.1.2. In this subsection we list some of their properties. We refer to [22, 24, 28, 29] for many of the facts stated below. Some of the properties we describe, in particular Proposition 2.15 and Corollary 2.16 seem to be new and might be of independent interest.

Notation: Here and below we will write $a \in b \pm \epsilon$ as a shorthand for $b - \epsilon \le a \le b + \epsilon$.

1. *Value at* 0. For all $0 \le s \le n$ holds $K_s(0) = \binom{n}{s}$.
2. *Symmetry.* For all $0 \le i, s \le n$ holds $K_s(i) = (-1)^s \cdot K_s(n-i)$.
3. *Reciprocity.* For all $0 \le i, s \le n$ holds $\binom{n}{i}K_s(i) = \binom{n}{s}K_i(s)$.
4. *ℓ_2 norm.* Viewing K_s as a function on $\{0,1\}^n$ or, equivalently, as a univariate real polynomial, endowing $\mathbb{R}$ with the binomial measure $\mu(i) = \frac{\binom{n}{i}}{2^n}$, for $0 \le i \le n$, we have $\|K_s\|_2 = \sqrt{\binom{n}{s}}$.
5. *Roots.* The polynomial K_s (viewed as a univariate polynomial) has s distinct real roots, which lie in the interval $\frac{n}{2} \pm \sqrt{s(n-s)}$. For $1 \le s \le \frac{n}{2}$, the distance between any two consecutive roots is at least 2 and at most $o(n)$.
6. *Magnitude outside the root region.* Let I and τ be the functions defined in Section 2.1. As shown in [22] (a more precise version of a result in [18]), for any $0 \le s \le \frac{n}{2}$ and $0 \le i \le \frac{n}{2} - \sqrt{s(n-s)}$ holds

$$\frac{\binom{n}{s}}{2^{H\left(\frac{s}{n}\right)\cdot n}} \cdot 2^{\tau\left(\frac{s}{n},\frac{i}{n}\right)\cdot n} \;=\; \binom{n}{s}\cdot 2^{\left(I\left(\frac{i}{n},\frac{s}{n}\right)-I\left(0,\frac{s}{n}\right)\right)\cdot n} \;\le\; K_s(i) \;\le\; 2^{\left(\tau\left(\frac{s}{n},\frac{i}{n}\right)+o(1)\right)\cdot n}. \quad (11)$$

In particular, using (3),

$$2^{\left(\tau\left(\frac{s}{n},\frac{i}{n}\right)-o(1)\right)\cdot n} \;\le\; K_s(i) \;\le\; 2^{\left(\tau\left(\frac{s}{n},\frac{i}{n}\right)+o(1)\right)\cdot n}.$$

7. *Magnitude in the root region.* By Corollary 2.16, the polynomial K_s, for $1 \le s \le \frac{n}{2}$, attains its ℓ_2 norm, up to an error of $O\left(n^{5/2}\right)$, between any two consecutive roots. That is, there are at least $s-1$ points i between the minimal and the maximal roots of K_s so that $\Omega\left(\frac{\binom{n}{s}}{n^5}\right) \le \frac{\binom{n}{i}}{2^n}K_s^2(i) \le \binom{n}{s}$. By (3), and by the definition of τ, for any such i holds

$$|K_s(i)| \;\in\; 2^{\left(\tau\left(\frac{s}{n},\frac{i}{n}\right)\pm o(1)\right)\cdot n}. \tag{12}$$

8. *Higher norms.* Let h be the function defined in Section 2.1.3. The following estimate seems to be new[5]. Let $p > 2$ be fixed. Let n be sufficiently large. Let $\frac{n}{\ln n} < s < \frac{n}{2} - \frac{n}{\ln n}$. Let $0 < y < \frac{1}{2}$ be such that $h(p,y) = 1 - \frac{2s}{n}$. By Proposition 4.16 the ℓ_p norm of K_s is attained, up to a small factor, in the vicinity of yn. More precisely, for some (in fact for all) $i \in yn \pm o(n)$ holds $\|K_s\|_p^p \le 2^{o(n)} \cdot \frac{\binom{n}{i}}{2^n} \cdot |K_s(i)|^p$.

   By Lemma 2.4, $yn < \frac{n}{2} - \sqrt{s(n-s)}$. Hence, using (11), we get the following estimate for $\|K_s\|_p^p$:

$$\|K_s\|_p^p \;\in\; 2^{\left(H(y)-1+p\tau\left(\frac{s}{n},y\right)\pm o(1)\right)\cdot n}.$$

   Since the ℓ_2 norm of K_s is $\sqrt{\binom{n}{s}}$ this implies that for $\frac{n}{\ln n} < s < \frac{n}{2} - \frac{n}{\ln n}$ holds

$$\frac{||K_s||_p^p}{||K_s||_2^p} \;\in\; 2^{\left(\psi\left(p,\frac{s}{n}\right)\pm o(1)\right)\cdot n},$$

   where ψ is the function defined in Section 2.1.4.

   We remark that Theorem 1.3 will imply that these estimates are valid for all $0 \le s \le n$.

2.3 Attaining norms between consecutive roots

The goal of this subsection is to show that Krawchouk polynomials attain their ℓ_2 norm, within a polynomial factor, between any two of their consecutive roots, and also in the intervals below their first and above their last roots. We prove this property, in a somewhat higher degree of generality, for any family of polynomials orthogonal with respect to a discrete measure supported on $\{0,...,n\}$.

Let μ be a positive measure on $\{0,...,n\}$, and let $\{P_0,...,P_n\}$ be the family of polynomials orthogonal with respect to the inner product $\langle f,g\rangle = \sum_{i=0}^n \mu(i)f(i)g(i)$ induced by μ, and normalized[6] so that $P_s(0) = 1$ for all $0 \le s \le n$.

[5] It might be known to experts, but we are unaware of its appearance in the literature.

[6] Note that we need to choose a normalization to make the polynomials $\{P_s\}$ well-defined, however the specific choice of a normalization is immaterial for the discussion below.

We will need two properties of orthogonal polynomials. First (see e.g., [42]), for any $0 \leq s \leq n$, the roots of P_s are real and distinct, and lie in the interval $(0, n)$; and second that its ℓ_2 norm $\|P_s\|_2 = \sqrt{\sum_{i=0}^{n} \mu(i) P_s^2(i)}$ is minimal among all polynomials of degree s with the same leading coefficient. This is a simple and a well-known fact, but we provide an argument for completeness. Let P be a polynomial of degree s with the same leading coefficient as P_s. Then $Q = P - P_s$ is a polynomial of a smaller degree and hence is orthogonal to P_s. This means that $||P||_2^2 = ||P_s + Q||_2^2 = ||P_s||_2^2 + ||Q||_2^2 \geq ||P_s||_2^2$.

We can now state our claim.

Proposition 2.15: *Let $s > 0$. Let the roots of P_s be $y_1 < y_2 < ... < y_s$. Assume that $y_1 \geq 1$ and that $y_s \leq n - 1$, and that the distance between any two consecutive roots is at least 2. Assume also that the ratios $\frac{\mu(j)}{\mu(j+1)}$ and their inverses are uniformly bounded by some $R > 0$.*

1. *The ℓ_2 norm of P_s is attained on the intervals $[0, y_1]$ and $[y_s, n]$ up to a factor of at most $O\left(\sqrt{Rn}\right)$.*
2. *For any $1 \leq k \leq s - 1$ the ℓ_2 norm of P_s is attained between y_k and y_{k+1}, up to a factor of at most $O\left(\sqrt{R}n^2\right)$.*

Proof:

We start with the first claim. We will prove it for the interval $[0, y_1]$, the proof for $[y_s, n]$ is similar. We will assume that the claim does not hold, and reach contradiction by constructing a polynomial P as above with $||P||_2^2 < ||P_s||_2^2$. There are two cases to consider: y_1 is non-integer, and y_1 is an integer. We will deal only with the first case, the second case is similar (and easier).

Assume then that the claim does not hold and that y_1 is not integer. Let $i_m = \lfloor y_1 \rfloor$. Let a_s be the leading coefficient of P_s. That is $P_s(y) = a_s \cdot \prod_{j=1}^{s} (y - y_j)$. For a (small) parameter τ define the polynomial P_τ as follows: $P_\tau(y) = a_s \cdot (y - y_1 - \tau) \cdot \prod_{j=2}^{s} (y - y_j)$. Note that the roots of P_τ, except for the first root, are those of P_s, and the first root is shifted downwards by τ. In particular, $P_0 = P_s$. We will show that $\frac{d}{d\tau}_{|\tau=0} ||P_\tau||_2^2 < 0$. This would mean that for some $\tau > 0$ we have $\|P_\tau\|_2^2 < \|P\|_2^2$, reaching a contradiction.

A simple computation shows that $\frac{d}{d\tau}_{|\tau=0} ||P_\tau||_2^2$ is proportional to $\sum_{i=0}^{n} \frac{\mu(i) P_s^2(i)}{y_1 - i}$. We write this expression as follows:

$$\sum_{i=0}^{i_m - 1} \frac{\mu(i) P_s^2(i)}{y_1 - i} \; + \; \frac{\mu(i_m) P_s^2(i_m)}{y_1 - i_m} \; + \; \sum_{i > y_1} \frac{\mu(i) P_s^2(i)}{y_1 - i}. \tag{13}$$

Since the denominators in the summands in the first sum are at least one, the first sum is bounded from above by $\sum_{i=0}^{i_m - 1} \mu(i) P_s^2(i)$, which, by assumption, is at most $\frac{\|P_s\|_2^2}{CRn}$, for some constant C which we may assume to be large. The last sum is negative. Since the denominators in its summands are at most n in absolute value, its absolute value is bounded from below by $\frac{1}{n} \cdot \sum_{i > y_1} \mu(i) P_s^2(i)$, which, by assumption, is at least $\frac{1}{n} \cdot \left(1 - \frac{1}{CRn}\right) \cdot \|P_s\|_2^2$.

Finally, we need to bound the second summand. Note that both $P_s(i_m-1)$ and $P_s(i_m)$ are positive, since P_s is positive at 0, and both points lie below the first root of P_s. Note also that $\frac{P_s(i_m-1)}{P_s(i_m)} = \prod_{i=1}^{n} \frac{y_i-(i_m-1)}{y_i-i_m} \geq \frac{y_1-i_m+1}{y_1-i_m} > \frac{1}{y_1-i_m}$. Hence

$$\frac{\mu(i_m) P_s^2(i_m)}{y_1-i_m} < R\cdot(y_1-i_m)\,\mu(i_m-1)\,P_s^2(i_m-1) < R\cdot\mu(i_m-1)\,P_s^2(i_m-1) \leq \frac{\|P_s\|_2^2}{Cn},$$

Summing up, we see that for a sufficiently large constant C, the derivative $\frac{d}{d\tau}_{|\tau=0}||P_\tau||_2^2$ is negative, proving the claim.

We pass to the second claim of the proposition, proceeding via a similar line of argument. We will assume that both y_k, y_{k+1} are non-integer. The other cases are similar (and simpler).

For a parameter τ define $P_\tau(y) = a_s \cdot \prod_{j\neq k,k+1}(y-y_j)\cdot(y-y_k+\tau)(y-y_{k+1}-\tau)$. That is, we move the two roots in question *outwards* by τ. A simple computation gives that $\frac{\partial}{\partial\tau}_{|\tau=0}\|P_\tau\|_2^2 = (y_k-y_{k+1})\cdot\sum_{i=0}^{n}\frac{\mu(i)P_s^2(i)}{(i-y_k)(i-y_{k+1})}$. Hence the contribution of all integer points outside the region between the two roots is negative, and inside positive. We want to argue that if the norm inside is smaller than the total ℓ_2 norm by a factor of more than CRn^4, for some sufficiently large constant C, then $\frac{d}{d\tau}_{|\tau=0}||P_\tau||_2^2$ is negative, reaching a contradiction.

Dividing out by $|y_k-y_{k+1}|$, the outside contributes in absolute value at least $\frac{1}{n^2}\cdot\left(1-\frac{1}{CR^2n^2}\right)\cdot\|P_s\|_2^2$. All the terms on the inside, for which the distance from both roots is at least $\frac{1}{2R}$, contribute together (note that the larger of these two distances is always at least 1) at most $\frac{2}{Cn^4}\cdot\|P_s\|_2^2$. It remains to deal with the inside terms which are close to one of the roots. Since the distance between the roots is at least 2, there could be only one such term at the most. Say, i is close to y_k from the inside. But then the contribution of $i+1$ would be at least $\frac{R}{4n^2}$ that of i, by an argument similar to the argument above. Since $i+1$ contributes at most $\frac{2}{Cn^4}\cdot\|P_s\|_2^2$, we have that i contributes at most $\frac{8}{CRn^2}\cdot\|P_s\|_2^2$, and the total contribution of the inside is bounded by $\frac{10}{Cn^2}\cdot\|P_s\|_2^2$. This means that for a sufficiently large constant C, the derivative $\frac{d}{d\tau}_{|\tau=0}||P_\tau||_2^2$ is negative, proving the second claim, and completing the proof of the proposition. ∎

Corollary 2.16: *Let $1\leq s\leq\frac{n}{2}$, and let x_s be the minimal root of the Krawchouk polynomial K_s. Then K_s attains its ℓ_2 norm within a factor of $O(n)$ on $[0,x_s]$ and within a factor of $O\left(n^{5/2}\right)$ between any two consecutive roots.*

Proof: Recall that for $1\leq s\leq\frac{n}{2}$ the distance between any two consecutive roots of K_s is at least 2. We also use one additional facts about the Krawchouk polynomials: for $1\leq s\leq\frac{n}{2}$, the first root of K_s is at least 1 (see [29]). Hence we may apply the previous proposition with μ being the binomial measure on $\{0,...,n\}$. Note that the value of R in this case is $n=\frac{\mu(1)}{\mu(0)}$. The claim of the corollary follows. ∎

2.4 Hamming spheres

Let $f=1_S$, where S is the Hamming sphere of radius $s\leq\frac{n}{2}$ around zero. Let ϕ be the function

defined in Subsection 2.1.6. Let $\sigma = \frac{s}{n}$. Then (see e.g., [3]):

$$\langle T_\epsilon f, f\rangle \;=\; \frac{1}{2^n}\binom{n}{s}\sum_{i=0}^{s}\binom{s}{i}\binom{n-s}{i}\epsilon^{2i}(1-\epsilon)^{n-2i} \;\in\; 2^{(\phi(\sigma,\epsilon)\pm o(1))\cdot n}.$$

The second step is by (3) and Lemma 2.12.

Since $\widehat{f} = \frac{1}{2^n}K_s$, and since $T_\epsilon = \sum_{k=0}^{n}(1-2\epsilon)^k\Pi_k$, we have, by Parseval's identity, that $\langle T_\epsilon f, f\rangle = \sum_{k=0}^{n}(1-2\epsilon)^k\langle f_k, f_k\rangle = \frac{1}{2^{2n}}\sum_{k=0}^{n}(1-2\epsilon)^k\binom{n}{k}K_s^2(k)$. Using (11) and (12) it can bee seen that the last expression is in $2^{\left(\max_{0\le y\le\frac{1}{2}}\{y\log_2(1-2\epsilon)+H(y)+2\tau(\sigma,y)\}-2\pm o(1)\right)\cdot n}$.

Comparing the two expressions for $\langle T_\epsilon f, f\rangle$, the following identity should hold:

$$\phi(\sigma,\epsilon) \;=\; \max_{0\le y\le\frac{1}{2}}\Big\{y\log_2(1-2\epsilon)+H(y)+2\tau(\sigma,y)\Big\}-2. \tag{14}$$

This is verified directly in Lemma 2.10. We remark that this identity, which shows that ϕ is, in an appropriate sense, a transform of τ, might be considered as a step towards understanding of the somewhat 'arbitrary looking' functions τ and I.

3 Some Proofs

In this section we prove theorems 1.5, 1.6, 1.8 and 1.10 and some related results.

Note that the arguments in this and in the following sections will rely, without further justification, on the properties of the bivariate functions detailed in Section 2.1.

3.1 Proof of Theorem 1.5

We start with the proof of (6), distinguishing two cases: $0 \le i < \frac{n}{2} - \sqrt{s(n-s)}$, and $\frac{n}{2} - \sqrt{s(n-s)} \le i \le \frac{n}{2}$.

Consider first the case $\frac{n}{2} - \sqrt{s(n-s)} \le i \le \frac{n}{2}$. By the definition of τ, for i in this range we have $\tau\left(\frac{s}{n},\frac{i}{n}\right) = \frac{1+H\left(\frac{s}{n}\right)-H\left(\frac{i}{n}\right)}{2}$. Therefore (6) reduces to

$$\Pr\left\{|f| \ge \|f\|_2\cdot 2^{\frac{1-H\left(\frac{i}{n}\right)}{2}\cdot n}\right\} \;\le\; 2^{\left(H\left(\frac{i}{n}\right)-1\right)\cdot n}.$$

Set $t = \|f\|_2\cdot 2^{\frac{1-H\left(\frac{i}{n}\right)}{2}\cdot n}$. Then, by Markov's inequality, $\Pr\{|f|\ge t\} \le \frac{\mathbb{E}f^2}{t^2} = 2^{\left(H\left(\frac{i}{n}\right)-1\right)\cdot n}$, completing the argument in this case.

For $0 \le i < \frac{n}{2} - \sqrt{s(n-s)}$, set $t = \|f\|_2\cdot 2^{\left(\tau\left(\frac{s}{n},\frac{i}{n}\right)-\frac{1}{2}H\left(\frac{s}{n}\right)\right)\cdot n}$ and use Markov's inequality and Theorem 1.3 to obtain

$$\Pr\{|f|\ge t\} \;\le\; \min_{p\ge 2}\left\{\frac{\mathbb{E}\,|f|^p}{t^p}\right\} \;=\; \min_{p\ge 2}\left\{\frac{\mathbb{E}\,|f|^p}{(\mathbb{E}\,f^2)^{\frac{p}{2}}}\cdot 2^{\left(-p\tau\left(\frac{s}{n},\frac{i}{n}\right)+\frac{p}{2}H\left(\frac{s}{n}\right)\right)\cdot n}\right\} \;\le$$

$2^{n\cdot\min_{p\geq 2}\left\{\psi\left(p,\frac{s}{n}\right)-p\tau\left(\frac{s}{n},\frac{i}{n}\right)+\frac{p}{2}H\left(\frac{s}{n}\right)\right\}}$.

Consider the function $g(p)=\psi\left(p,\frac{s}{n}\right)-p\tau\left(\frac{s}{n},\frac{i}{n}\right)+\frac{p}{2}H\left(\frac{s}{n}\right)$. We claim that its minimum on $[2,\infty)$ is attained at p^*, which is defined by $h\left(p^*,\frac{i}{n}\right)=1-\frac{2s}{n}$. (Note that p^* is well-defined in this range of i.) Since $\psi\left(p,x\right)$ is convex in p, the function g is convex, and it will suffice to verify that $g'\left(p^*\right)=0$. Proceeding as in the proof of Proposition 2.7 below, we have that

$$g'\left(p^*\right) \;=\; \frac{\partial\psi\left(p,\frac{s}{n}\right)}{\partial p}_{|p=p^*}-\tau\left(\frac{s}{n},\frac{i}{n}\right)+\frac{1}{2}H\left(\frac{s}{n}\right) \;=\; \tau\left(\frac{s}{n},y\right)-\tau\left(\frac{s}{n},\frac{i}{n}\right),$$

where y is determined by $h\left(p^*,y\right)=1-\frac{2s}{n}$. By the definition of p^* we have $y=\frac{i}{n}$, and therefore $g'\left(p^*\right)=0$, as claimed.

We now compute $g\left(p^*\right)$. Recall that $\psi(p,x)=H(y)-1+p\tau(x,y)-\frac{p}{2}H(x)$, where y is determined by $h(p,y)=1-2x$. In our case $x=\frac{s}{n}$ and $y=\frac{i}{n}$. Substituting, we get $g\left(p^*\right)=H\left(\frac{i}{n}\right)-1$, completing the proof of (6).

We proceed to the second part of the theorem. Let $0\leq i\leq\frac{n}{2}-\sqrt{s(n-s)}$, and let $0\leq j\leq i$. Recall that by (11) we have for $0\leq j\leq\frac{n}{2}-\sqrt{s(n-s)}$ that $K_s(j)\geq\binom{n}{s}\cdot 2^{\left(I\left(\frac{j}{n},\frac{s}{n}\right)-I\left(0,\frac{s}{n}\right)\right)\cdot n}=\frac{\binom{n}{s}}{2^{H\left(\frac{s}{n}\right)\cdot n}}\cdot 2^{\tau\left(\frac{s}{n},\frac{j}{n}\right)}\geq\frac{\binom{n}{s}}{2^{H\left(\frac{s}{n}\right)\cdot n}}\cdot 2^{\tau\left(\frac{s}{n},\frac{i}{n}\right)}$. Recall also that $\|K_s\|_2=\sqrt{\binom{n}{s}}$. Hence, using (3) in the last inequality,

$$\Pr\left\{|K_s|\geq\|K_s\|_2\cdot\sqrt{\frac{\binom{n}{s}}{2^{H\left(\frac{s}{n}\right)\cdot n}}}\cdot 2^{\left(\tau\left(\frac{s}{n},\frac{i}{n}\right)-\frac{1}{2}H\left(\frac{s}{n}\right)\right)\cdot n}\right\} \;\geq\; \Pr\left\{K_s\geq\frac{\binom{n}{s}}{2^{H\left(\frac{s}{n}\right)\cdot n}}\cdot 2^{\tau\left(\frac{s}{n},\frac{i}{n}\right)\cdot n}\right\} \;\geq$$

$$\frac{1}{2^n}\cdot\sum_{j=0}^{i}\binom{n}{j} \;\geq\; \Omega\left(\frac{1}{\sqrt{i}}\right)\cdot 2^{\left(H\left(\frac{i}{n}\right)-1\right)\cdot n}.$$

Next, by Corollary 2.16, between any two consecutive roots of K_s there is a point i for which $\frac{\binom{n}{i}}{2^n}\cdot K_s^2(i)\geq\Omega\left(\frac{1}{n^5}\right)\cdot\|K_s\|_2^2$. This means that

$$\Pr\left\{|K_s|\geq\Omega\left(\frac{1}{n^{5/2}}\right)\cdot\|K_s\|_2\cdot\sqrt{\frac{2^n}{\binom{n}{i}}}\right\} \;\geq\; \frac{\binom{n}{i}}{2^n} \;\geq\; \Omega\left(\frac{1}{\sqrt{i}}\right)\cdot 2^{\left(H\left(\frac{i}{n}\right)-1\right)\cdot n}.$$

Since in this range of i we have $\sqrt{\frac{2^n}{\binom{n}{i}}}\geq 2^{\frac{1-H\left(\frac{i}{n}\right)}{2}\cdot n}=2^{\left(\tau\left(\frac{s}{n},\frac{i}{n}\right)-\frac{1}{2}H\left(\frac{s}{n}\right)\right)\cdot n}$, this proves the last claim of the theorem.

∎

3.2 Two auxiliary claims

The following claim provides a key to all the remaining results in this section. Note that this is a special case of (10). Let π be the function defined in Section 2.1.5. We write $x\wedge y$ for the minimum of x and y.

Proposition 3.1: *Let $0 \leq \sigma \leq \frac{1}{2}$. Let f be a function on $\{0,1\}^n$ supported on a set of cardinality at most $2^{H(\sigma)n}$. Then, for any $0 \leq k \leq n$ holds*

$$\|f_k\|_2 \quad \leq \quad 2^{\pi\left(\sigma, \frac{k}{n} \wedge \frac{n-k}{n}\right) \cdot n} \cdot \|f\|_2.$$

Proof:

Given a function f, let g be defined by $g(x) = (-1)^{|x|} f(x)$. Then f and g have supports of the same cardinality and (see Section 1.1.2) for any $0 \leq k \leq n$ holds $g_k = f_{n-k}$. Hence we may and will assume in the following argument that $0 \leq k \leq \frac{n}{2}$. Next, recall that $\pi\left(\sigma, \frac{k}{n}\right) = 0$ for $\frac{k}{n} \geq \frac{1}{2} - \sqrt{\sigma(1-\sigma)}$, reducing the claim of the proposition to the trivial inequality $\|f_k\|_2 \leq \|f\|_2$. So, we may assume $0 \leq \frac{k}{n} < \frac{1}{2} - \sqrt{\sigma(1-\sigma)}$.

Let f be a function on $\{0,1\}^n$, supported on a subset $A \subseteq \{0,1\}^n$. Let $0 \leq k \leq \frac{n}{2}$. Using the fact that f_k is an orthogonal projection of f in the first step and the Cauchy-Schwarz inequality in the last step, we have

$$\langle f_k, f_k \rangle \;=\; \langle f, f_k \rangle \;=\; \langle f \cdot 1_A, f_k \rangle \;=\; \langle f, f_k \cdot 1_A \rangle \;\leq\; \|f\|_2 \cdot \|f_k \cdot 1_A\|_2,$$

implying that $\frac{\langle f_k, f_k \rangle}{\|f\|_2} \leq \|f_k \cdot 1_A\|_2$. On the other hand, for $p \geq 2$, we can apply Hölder's inequality to obtain $\|f_k \cdot 1_A\|_2^2 = \left\langle f_k^2, 1_A \right\rangle \leq \|1_A\|_{\frac{p}{p-2}} \cdot \|f_k^2\|_{\frac{p}{2}} = \left(\frac{|A|}{2^n}\right)^{\frac{p-2}{p}} \cdot \|f_k\|_p^2$. Combining the two estimates gives

$$\frac{\|f_k\|_2}{\|f\|_2} \;=\; \frac{1}{\|f_k\|_2} \cdot \frac{\langle f_k, f_k \rangle}{\|f\|_2} \;\leq\; \left(\frac{|A|}{2^n}\right)^{\frac{p-2}{2p}} \cdot \frac{\|f_k\|_p}{\|f_k\|_2} \;\leq\; \left(\frac{|A|}{2^n}\right)^{\frac{p-2}{2p}} \cdot 2^{\frac{1}{p}\psi(p,k/n)\cdot n},$$

where in the last step we have applied Theorem 1.3, using the fact that f_k is a homogeneous polynomial of degree k.

Since, by assumption, $|A| \leq 2^{H(\sigma)n}$, we have that for any $p \geq 2$ holds

$$\|f_k\|_2 \quad \leq \quad 2^{\left(\frac{p-2}{2p} \cdot (H(\sigma)-1) + \frac{1}{p}\psi\left(p, \frac{k}{n}\right)\right) \cdot n} \cdot \|f\|_2.$$

Since $\frac{k}{n} < \frac{1}{2} - \sqrt{\sigma(1-\sigma)}$, there is a unique $p > 2$ such that $h(p,\sigma) = 1 - 2\frac{k}{n}$. Fix this p. We then have, by the first definition of ψ, that $\psi\left(p, \frac{k}{n}\right) = H(\sigma) - 1 + p\tau\left(\frac{k}{n}, \sigma\right) - \frac{p}{2} H\left(\frac{k}{n}\right)$. And hence $\frac{p-2}{2p} \cdot (H(\sigma)-1) + \frac{1}{p}\psi\left(p, \frac{k}{n}\right) = \tau\left(\frac{k}{n}, \sigma\right) - \frac{1 + H\left(\frac{k}{n}\right) - H(\sigma)}{2} = \pi\left(\frac{k}{n}, \sigma\right) = \pi\left(\sigma, \frac{k}{n}\right)$, completing the proof of the proposition. ∎

As a corollary we prove the following special case of (9). Let T_ϵ be the noise operator corresponding to a noise parameter $0 \leq \epsilon \leq \frac{1}{2}$ (see Section 1.1.2). Let ϕ be the function defined in Section 2.1.6.

Corollary 3.2:

Let $0 \leq \sigma \leq \frac{1}{2}$. Let f be a function on $\{0,1\}^n$ supported on a set of cardinality at most $2^{H(\sigma)n}$. Then, for any $0 \leq \epsilon \leq \frac{1}{2}$ holds

$$\langle T_\epsilon f, f \rangle \quad \leq \quad 2^{(\phi(\sigma,\epsilon) + 1 - H(\sigma)) \cdot n} \cdot \|f\|_2^2.$$

Proof:

We have, using Proposition 3.1 in the first inequality, that

$$\langle T_\epsilon f, f\rangle \quad = \quad \sum_{k=0}^{n}(1-2\epsilon)^k \langle f_k, f_k\rangle \quad \leq \quad \|f\|_2^2 \cdot \sum_{k=0}^{n}(1-2\epsilon)^k 2^{2\pi\left(\sigma,\frac{k}{n}\wedge\frac{n-k}{n}\right)\cdot n} \quad \leq$$

$$n\|f\|_2^2 \cdot \max_{0\leq k\leq\frac{n}{2}}\left\{(1-2\epsilon)^k 2^{2\pi\left(\sigma,\frac{k}{n}\right)\cdot n}\right\} \quad \leq \quad n\|f\|_2^2 \cdot 2^{\left(\max_{0\leq y\leq\frac{1}{2}}\{y\log_2(1-2\epsilon)+2\pi(\sigma,y)\}\right)\cdot n}.$$

Since $\pi(x,y) = \tau(x,y) - \frac{1+H(x)-H(y)}{2}$ we have

$$\max_{0\leq y\leq\frac{1}{2}}\{y\log_2(1-2\epsilon)+2\pi(\sigma,y)\} \quad = \quad \max_{0\leq y\leq\frac{1}{2}}\{y\log_2(1-2\epsilon)+H(y)+2\tau(\sigma,y)\}-(1+H(\sigma)) \quad =$$

$$\phi(\sigma,\epsilon)+1-H(\sigma),$$

where the last equality is by Lemma 2.10. So, $\langle T_\epsilon f, f\rangle \leq n\cdot 2^{(\phi(\sigma,\epsilon)+1-H(\sigma))\cdot n}\cdot\|f\|_2^2$.

Finally, we remove the extra n-factor by a tensorization argument (see Section 1.1.4). For an integer $m\geq 1$, let $F_m = f^{\otimes m}$. Observe that F_m is supported on a subset of $\{0,1\}^{nm}$ of cardinality at most $2^{H(\sigma)nm}$. In addition, $\langle F_m, F_m\rangle = \langle f, f\rangle^m$ and $\langle T_\epsilon F_m, F_m\rangle = \langle T_\epsilon f, f\rangle^m$. Hence, by the above argument, we have

$$\langle T_\epsilon f, f\rangle \quad = \quad \langle T_\epsilon F_m, F_m\rangle^{\frac{1}{m}} \quad \leq \quad (nm)^{\frac{1}{m}}\cdot\left(2^{(\phi(\sigma,\epsilon)+1-H(\sigma))\cdot nm}\right)^{\frac{1}{m}}\cdot\left(\|F_m\|_2^2\right)^{\frac{1}{m}} \quad =$$

$$(nm)^{\frac{1}{m}}\cdot 2^{(\phi(\sigma,\epsilon)+1-H(\sigma))\cdot n}\cdot\|f\|_2^2.$$

Taking m to infinity, gives the claim of the corollary. ∎

We can now prove Theorems 1.6 to 1.10.

Proof of Theorem 1.6 and related statements

Remark 3.3: Let us first briefly explain why $1\leq i\leq 2\sigma(1-\sigma)n$ is the relevant range of parameters. (See also the discussion following the proof of Theorem 5 in [3].) Let $s=\sigma n$ and assume s to be integer, and $i = 2\sigma(1-\sigma)n = \frac{2s(n-s)}{n}$ to be an even integer. If A is the Hamming sphere of radius s around zero then i is the expected distance between two points chosen uniformly at random from A, and it is easy to see that, up to at most an $O\left(\frac{1}{\sqrt{n}}\right)$-factor, we have $a_i(A) = |A|^2$. Hence in this case we cannot expect a non-trivial upper bound on $a_i(A)$.

For a larger i, that is $2\sigma(1-\sigma)n\leq i\leq\frac{n}{2}$, write $i=\frac{2t(n-t)}{n}$, and, assuming $t>s$ to be integer, choose A to be a random subset of cardinality $\binom{n}{s}$ of the sphere of radius t around zero, to a similar effect. For $\frac{n}{2}<i\leq n$, let $j=n-i$, and let $B=A\cup\bar{A}$, where $\bar{A}$ is the shift of A by an all-1 vector. Then $0\leq j<\frac{n}{2}$, $|A|\leq|B|\leq 2|A|$, $a_i(B)=a_j(B)$ and $a_i(A)\leq a_j(B)\leq 2\cdot(a_i(A)+a_j(A))$. ∎

Proof of (7)

We follow the argument in the proof of Theorem 5 from [3] replacing the hypercontractive inequality (4) used in [3] with Corollary 3.2.

Let ϵ be between 0 and $\frac{1}{2}$. Note that for any function f on $\{0,1\}^n$ holds $\langle f_\epsilon, f\rangle = \frac{1}{2^n}\sum_{x,y} f(x)f(y)\epsilon^{|x+y|}(1-\epsilon)^{n-|x+y|}$. Substituting $f = 1_A$ gives $\langle f_\epsilon, f\rangle = \frac{1}{2^n}\sum_{i=0}^{n} a_i(A)\epsilon^i(1-\epsilon)^{n-i}$. On the other hand, Corollary 3.2, gives $\langle f_\epsilon, f\rangle \leq 2^{(\phi(\sigma,\epsilon)+1-H(\sigma))\cdot n}\cdot\|f\|_2^2 = |A|\cdot 2^{(\phi(\sigma,\epsilon)-H(\sigma))\cdot n}$.

Hence, for any $0 \leq i \leq n$ and for any $0 < \epsilon \leq \frac{1}{2}$ we have

$$a_i(A) \;\leq\; \frac{\sum_{k=0}^{n} a_k(A)\epsilon^k(1-\epsilon)^{n-k}}{\epsilon^i(1-\epsilon)^{n-i}} \;\leq\; |A|\cdot\frac{2^{(\phi(\sigma,\epsilon)+1-H(\sigma))\cdot n}}{\epsilon^i(1-\epsilon)^{n-i}}.$$

Minimizing over ϵ gives

$$a_i(A) \;\leq\; |A|\cdot 2^{\min_{0<\epsilon\leq\frac{1}{2}}\left\{\phi(\sigma,\epsilon)+1-H(\sigma)-\frac{i}{n}\log_2(\epsilon)-\left(1-\frac{i}{n}\right)\log_2(1-\epsilon)\right\}\cdot n} \;\leq$$

$$|A|\cdot 2^{\left(\sigma H\left(\frac{i}{2\sigma n}\right)+(1-\sigma)H\left(\frac{i}{2(n-\sigma n)}\right)\right)\cdot n},$$

where the last step is via Lemma 2.11. ∎

Proof of near tightness of (7) for spheres or unions of spheres

For an even $i = 2j$, let A be a sphere of radius s. Then $a_i(A) = |A|\cdot\binom{s}{j}\binom{n-s}{j}$, while (7) gives $a_i(A) \leq |A|\cdot 2^{sH\left(\frac{j}{s}\right)+(n-s)H\left(\frac{j}{n-s}\right)}$. By (3), the upper bound provided by (7) is larger than $a_i(A)$ by at most a factor of $\Theta\left(j\cdot\sqrt{\frac{(s-j)(n-s-j)}{s(n-s)}}\right) \leq O(i)$.

For an odd $i = 2j-1$, let A be the union of two adjacent spheres of dimension $n-1$ and radii $s-1$ and s. Note that $|A| = \binom{n}{s}$, and that $a_i(A) = 2|A|\binom{s}{j}\binom{n-s-1}{j-1}$. As above, this loses a factor of at most $O(i)$ to $|A|\cdot 2^{sH\left(\frac{j}{s}\right)+(n-s-1)H\left(\frac{j-1}{n-s-1}\right)}$. Next, a simple analysis, which we omit, shows that $sH\left(\frac{i}{2s}\right)+(n-s)H\left(\frac{i}{2(n-s)}\right) = sH\left(\frac{2j-1}{2s}\right)+(n-s)H\left(\frac{2j-1}{2(n-s)}\right)$ is larger than $sH\left(\frac{j}{s}\right)+(n-s-1)H\left(\frac{j-1}{n-s-1}\right)$ by at most $\frac{1}{2}\log_2\left(\frac{n-s}{s}\right)+O(1)$, and hence the upper bound on $a_i(A)$ provided by (7) is tight, up to a factor of $O\left(\sqrt{\frac{n-s}{s}}\cdot i\right)$. ∎

Proof of Corollary 1.7

First, note that for any $t \geq 1$ holds $tH\left(\frac{1}{t}\right) = \log_2(t)+(t-1)\log_2\left(\frac{t}{t-1}\right) \leq \log_2(t)+\frac{1}{\ln 2}$, where in the last step we have used the fact that $\ln(1+x) \leq x$ for any $x > -1$.

Let $1 \leq s \leq \frac{n}{2}$ be an integer. Using the observation above and (7) with $\sigma = \frac{s}{n}$ and $i = 2$, gives

$$a_2(A) \;\leq\; |A|\cdot 2^{sH\left(\frac{1}{s}\right)+(n-s)H\left(\frac{1}{n-s}\right)} \;\leq\; e^2 s(n-s)\cdot|A|.$$

∎

Proof of (8)

If $R = 1$, the claim becomes $P_{\text{ue}}(1, \epsilon) = 0$, which is correct, since in this case the undetected error probability is $1 - (1-\epsilon)^n \approx 1$. So, we may assume $0 < R < 1$. Let $\sigma = H^{-1}(R)$. Let n be a large integer. We will assume, somewhat unaccurately, that $s = \sigma n$ is integer, whenever required (to do this with full accuracy we would sandwich $P_{\text{ue}}(R, \epsilon)$ between $P_{\text{ue}}\left(H\left(\frac{\lfloor \sigma n \rfloor}{n}\right), \epsilon\right)$ and $P_{\text{ue}}\left(H\left(\frac{\lceil \sigma n \rceil}{n}\right), \epsilon\right)$ and proceed similarly).

Let $A(\sigma, n)$ be the union spheres of dimension $n-1$ and radii $s-1$ and s around zero. We will show below that A maximizes the undetected error probability for all codes in $\{0,1\}^n$ of cardinality at most $2^{H(\sigma)n} = 2^{Rn}$, up to a factor of at most $O\left(s^{\frac{1}{2}} n^{\frac{3}{2}}\right)$. This will imply that

$$P_{\text{ue}}(R, \epsilon) \;=\; \limsup_{n\to\infty} \left(\frac{1}{n} \log_2 \left(P_{\text{ue}}\left(A(\sigma, n), \epsilon\right) \right) \right).$$

Let $A = A(\sigma, n)$, and let $f = 1_A$. Then $f = g + h$, where g is the characteristic function of the sphere of radius s around zero and h is the characteristic function of the sphere of radius $s-1$ around zero. Note that $P_{\text{ue}}(A, \epsilon) = \frac{2^n}{|A|} \cdot \langle f_\epsilon, f \rangle - (1-\epsilon)^n$. We claim that for $\epsilon > 0$ the first of these terms is exponentially in n larger than the second one, and consequently $P_{\text{ue}}(A, \epsilon) \approx \frac{2^n}{|A|} \cdot \langle f_\epsilon, f \rangle$. Indeed, we have

$$\frac{1}{n} \log_2 \left(\frac{2^n}{|A|} \cdot \langle f_\epsilon, f \rangle \right) \;\geq\; \frac{1}{n} \log_2 \left(\frac{2^n}{|A|} \cdot \langle g_\epsilon, g \rangle \right) \;\in\; \phi(\sigma, \epsilon) + 1 - H(\sigma) \pm o_n(1) \;=$$

$$\max_{0 \leq x \leq \sigma} \left\{ \alpha_{\sigma,\epsilon}(x) \right\} \pm o_n(1).$$

For the second step, see Section 2.4, and for the third step Section 2.1.6. As stated in Section 2.1.6, the value of x for which the maximum is attained is $x^* = x(\sigma, \epsilon) = \frac{-\epsilon^2 + \epsilon\sqrt{\epsilon^2 + 4(1-2\epsilon)\sigma(1-\sigma)}}{2(1-2\epsilon)}$, which is strictly positive for $\sigma, \epsilon > 0$. And it is easy to see that $\alpha_{\sigma,\epsilon}(x^*) > \alpha_{\sigma,\epsilon}(0) = \log_2(1-\epsilon)$.

Next, note that $\langle g_\epsilon, g \rangle \leq \langle f_\epsilon, f \rangle = \langle g_\epsilon + h_\epsilon, g + h \rangle \leq 2 \cdot (\langle g_\epsilon, g \rangle + \langle h_\epsilon, h \rangle)$. For the last step recall that the noise operator is a positive semidefinite linear operator, and hence we have the Cauchy-Schwarz inequality $\langle g_\epsilon, h \rangle \leq \langle g_\epsilon, g \rangle^{\frac{1}{2}} \cdot \langle h_\epsilon, h \rangle^{\frac{1}{2}}$, and similarly for $\langle g, h_\epsilon \rangle$. Since both $\frac{1}{n} \log_2 \left(\frac{2^n}{|A|} \cdot \langle g_\epsilon, g \rangle \right)$ and $\frac{1}{n} \log_2 \left(\frac{2^n}{|A|} \cdot \langle h_\epsilon, h \rangle \right)$ are in $\alpha_{\sigma,\epsilon}(x^*) \pm o_n(1)$, this implies that also $\frac{1}{n} \log_2 \left(P_{\text{ue}}(A, \epsilon) \right) \in \alpha_{\sigma,\epsilon}(x^*) \pm o_n(1)$. Hence

$$P_{\text{ue}}(R, \epsilon) \;=\; \limsup_{n\to\infty} \left(\frac{1}{n} \log_2 \left(P_{\text{ue}}\left(A(\sigma, n), \epsilon\right) \right) \right) \;=\; \alpha_{\sigma,\epsilon}(x^*),$$

proving (8).

To conclude the proof it remains to show that A maximizes the undetected error probability up to a polynomial factor. Let $t = \frac{2s(n-s)}{n}$. We first claim that the number of pairs of points in A at distances between 1 and t from each other is not negligible. More precisely, $\sum_{i=1}^{t} a_i(A) \geq \Omega\left(n^{-\frac{3}{2}}\right) \cdot |A|^2$. This can be shown by observing that distance distributions

inside a Hamming sphere and between two distinct Hamming spheres are closely related to hypergeometric distributions with appropriate parameters, and by applying straightforward first moment estimates for hypergeometric distributions. We omit the details.

Let now $C \subseteq \{0,1\}^n$ with $|C| \leq 2^{H(\sigma)n}$. On one hand we have by Theorem 1.6 that

$$\frac{1}{|C|} \cdot \sum_{i=1}^{t} a_i(C)\epsilon^i(1-\epsilon)^{n-i} \;\leq\; O\left(s^{-\frac{1}{2}} n^{\frac{3}{2}}\right) \frac{1}{|A|} \cdot \sum_{i=1}^{t} a_i(A)\epsilon^i(1-\epsilon)^{n-i} \;\leq\; O\left(s^{-\frac{1}{2}} n^{\frac{3}{2}}\right) \cdot P_{\mathrm{ue}}(A,\epsilon).$$

On the other hand, we have

$$\frac{1}{|C|} \cdot \sum_{i=t+1}^{n} a_i(C)\epsilon^i(1-\epsilon)^{n-i} \;\leq\; \epsilon^{t+1}(1-\epsilon)^{n-t-1} \cdot \frac{1}{|C|} \sum_{i=t+1}^{n} a_i(C) \;\leq\; \epsilon^{t+1}(1-\epsilon)^{n-t-1} \cdot |C| \;\leq$$

$$\epsilon^t(1-\epsilon)^{n-t} \cdot O\left(s^{\frac{1}{2}}\right)|A| \;\leq\; O\left(s^{\frac{1}{2}} n^{\frac{3}{2}}\right) \cdot \frac{1}{|A|} \sum_{i=1}^{t} a_i(A)\epsilon^i(1-\epsilon)^{n-i} \;\leq\; O\left(s^{\frac{1}{2}} n^{\frac{3}{2}}\right) \cdot P_{\mathrm{ue}}(A,\epsilon).$$

We use (3) in the third step, and the inequality $\sum_{i=1}^{t} a_i(A) \geq \Omega\left(n^{-\frac{3}{2}}\right) \cdot |A|^2$ in the fourth step.

Combining these two inequalities gives

$$P_{\mathrm{ue}}(C,\epsilon) \;=\; \frac{1}{|C|} \cdot \sum_{i=1}^{n} a_i(C)\epsilon^i(1-\epsilon)^{n-i} \;\leq\; O\left(s^{\frac{1}{2}} n^{\frac{3}{2}}\right) \cdot P_{\mathrm{ue}}(A,\epsilon).$$

∎

Proof of Theorem 1.8

Proof of (9)

Note that it would suffice to show a somewhat weaker statement:

$$\|f_\epsilon\|_2 \;\leq\; 2^{o(n)} \cdot 2^{\eta\left(\frac{1}{n}\log_2\left(\frac{\|f\|_{1+(1-2\epsilon)^2}}{\|f\|_1}\right),\ \epsilon\right)\cdot n} \cdot \|f\|_{1+(1-2\epsilon)^2}, \tag{15}$$

since the $2^{o(n)}$ term by can be removed by a tensorization argument, like in the proof of Corollary 3.2. We proceed to show (15), with an error term which is polynomial in n.

We may assume that $f \geq 0$, since replacing f with $|f|$ increases the LHS of (15) and does not change the RHS. We may also assume, by homogeneilty, that $\|f\|_1 = 1$. This means that $\|f\|_\infty \leq 2^n$, and that the points at which $f < 2^{-n}$, say, contribute little to both sides ot the inequality, so we may ignore them for the sake of discussion (that is, we may and will assume that f vanishes on these points). All the remaining points can be partitioned into $O(n)$ level sets $A_1, ...A_r$ such that f varies by a factor of 2 at most in each level set. Let $\alpha_i = \frac{1}{n}\log_2\left(\frac{|A_i|}{2^n}\right)$,

and let $\nu_i = \frac{1}{n}\log_2(v_i)$, where v_i is, say, the median value of f on A_i. Then, up to an additive error term of $O\left(\frac{\log(n)}{n}\right)$, we have, for any $p \geq 1$, that

$$\frac{1}{n}\log_2\|f\|_p \approx \frac{1}{n}\log_2\left(\left(\sum_{i=1}^{r}\frac{|A_i|}{2^n}\cdot v_i^p\right)^{1/p}\right) \approx \max_{1\leq i\leq r}\left\{\frac{\alpha_i-1}{p}+\nu_i\right\}.$$

Here we use the approximate equality sign "$\approx$" to register that the equality holds up to a negligible error.

Next, we estimate the LHS of (15) in terms of $\{\alpha_i\}$ and $\{\nu_i\}$. Let f_i be the restriction of f to A_i. Then $f_\epsilon = \sum_{i=1}^{r}(f_i)_\epsilon$, and we have that,up to an additive error term of $O\left(\frac{\log(n)}{n}\right)$,

$$\frac{1}{n}\log_2\langle f_\epsilon, f_\epsilon\rangle \approx \frac{1}{n}\log_2\left(\sum_{i=1}^{r}\langle (f_i)_\epsilon, (f_i)_\epsilon\rangle\right) = \frac{1}{n}\log_2\left(\sum_{i=1}^{r}\left\langle (f_i)_{2\epsilon(1-\epsilon)}, f_i\right\rangle\right) \approx$$

$$\max_{1\leq i\leq r}\left\{\frac{1}{n}\log_2\left\langle (f_i)_{2\epsilon(1-\epsilon)}, f_i\right\rangle\right\} \approx \max_{1\leq i\leq r}\left\{\frac{1}{n}\log_2\left(v_i^2\cdot\left\langle (1_{A_i})_{2\epsilon(1-\epsilon)}, 1_{A_i}\right\rangle\right)\right\} \leq$$

$$\max_{1\leq i\leq r}\left\{\tilde{\phi}\left(\alpha_i, 2\epsilon(1-\epsilon)\right)+2\nu_i\right\},$$

The first step follows from the Cauchy-Schwarz inequality, the second step uses the semigroup property of noise operators, and the last step follows from Corollary 3.2, and the definition of $\tilde{\phi}$.

We will show (15) to be a simple corollary of the following lemma. Given $0 \leq \alpha_1, ...\alpha_r \leq 1$ and $0 \leq \nu_1, ...\nu_r$, we write $N(p)$ for $\max_{1\leq i\leq r}\left\{\frac{\alpha_i-1}{p}+\nu_i\right\}$.

Lemma 3.4: *Let $0 \leq \delta \leq 1/2$, and let $p \geq 2-2\delta$. Then, for any $0 \leq \alpha_1, ...\alpha_r \leq 1$ and $0 \leq \nu_1, ...\nu_r$ holds*

$$\max_{1\leq i\leq r}\left\{\tilde{\phi}\left(\alpha_i, \delta\right)+2\nu_i\right\} \leq \tilde{\phi}\left(\alpha^*, \delta\right)+2\nu^*,$$

where $\alpha^* = 1-\frac{p}{p-1}\cdot(N(p)-N(1))$, and $\nu^* = \frac{p}{p-1}\cdot N(p)-\frac{1}{p-1}\cdot N(1)$.

Proof: We need to show that for all $1 \leq i \leq r$ holds $\tilde{\phi}\left(\alpha_i, \delta\right)+2\nu_i \leq \phi\left(\alpha^*, \delta\right)+2\nu^*$. Fix i. Recall that $\frac{1}{1-\delta} \leq \frac{\partial\tilde{\phi}(\alpha,\delta)}{\partial\alpha} \leq 2$. There are two cases to consider.

1. $\nu_i \leq \nu^*$.

   If also $\alpha^* \geq \alpha_i$, the claim follows from the monotonicity of $\tilde{\phi}$ in α. If, on the other hand, $\alpha^* < \alpha_i$, then $\tilde{\phi}\left(\alpha_i, \delta\right)-\tilde{\phi}\left(\alpha^*, \delta\right) \leq 2\alpha_i - 2\alpha^*$, and hence it only remains to verify that $\alpha_i+\nu_i \leq \alpha^*+\nu^*$. To see that note that

   $$\alpha_i+\nu_i \leq 1+N(1) = \alpha^*+\nu^*.$$

2. $\nu_i > \nu^*$.

   Note that $\frac{\alpha_i}{p} + \nu_i \leq \frac{1}{p} + N(p) = \frac{\alpha^*}{p} + \nu^*$. In particular, we have that $\alpha^* > \alpha_i$. And hence

$$\tilde{\phi}(\alpha^*, \delta) - \tilde{\phi}(\alpha_i, \delta) \quad \geq \quad \frac{1}{1-\delta} \cdot (\alpha^* - \alpha_i) \quad \geq \quad 2 \cdot \frac{\alpha^* - \alpha_i}{p} \quad \geq \quad 2\nu_i - 2\nu^*.$$

∎

We can now conclude the proof of (15). Given a function f, define, as above, the partition of $\{0,1\}^n$ into level sets $A_1, ..., A_r$ of f, and define $\{\alpha_i\}$ and $\{\nu_i\}$ correspondingly. Apply the lemma with $p = 1 + (1-2\epsilon)^2$ and $\delta = 2\epsilon(1-\epsilon)$. By the discussion above, and bearing in mind that $\left|\frac{\partial\tilde{\phi}(\alpha,\epsilon)}{\partial\alpha}\right|$ is bounded by a constant (in fact by 2), we have, up to an additive error term of $O\left(\frac{\log(n)}{n}\right)$, that

$$\frac{1}{n}\log_2 \|f_\epsilon\|_2^2 \;=\; \frac{1}{n}\log_2 \langle f_\epsilon, f_\epsilon \rangle \;\lesssim\; \max_{1\leq i\leq r}\left\{\tilde{\phi}(\alpha_i, 2\epsilon(1-\epsilon)) + 2\nu_i\right\} \;\leq\; \tilde{\phi}(\alpha^*, 2\epsilon(1-\epsilon)) + 2\nu^* \;=$$

$$\tilde{\phi}\left(1 - \frac{p}{p-1}\cdot(N(p) - N(1)),\; 2\epsilon(1-\epsilon)\right) + 2\cdot\frac{N(p) - N(1)}{p-1} + 2N(p) \;\approx$$

$$\tilde{\phi}\left(1 - \frac{p}{p-1}\cdot\frac{1}{n}\log_2\left(\frac{\|f\|_p}{\|f\|_1}\right),\; 2\epsilon(1-\epsilon)\right) + 2\cdot\frac{\frac{1}{n}\log_2\left(\frac{\|f\|_p}{\|f\|_1}\right)}{p-1} + \frac{2}{n}\log_2\left(\|f\|_p\right) \;=$$

$$2\eta_p\left(\frac{1}{n}\log_2\left(\frac{\|f\|_p}{\|f\|_1}\right),\; \epsilon\right) + \frac{2}{n}\log_2\left(\|f\|_p\right) \;=\; 2\eta\left(\frac{1}{n}\log_2\left(\frac{\|f\|_{1+(1-2\epsilon)^2}}{\|f\|_1}\right),\; \epsilon\right) + \frac{2}{n}\log_2\left(\|f\|_{1+(1-2\epsilon)^2}\right).$$

∎

Proof of near tightness of (9) for spheres

Proposition 3.5: *Let f be proportional to 1_S, where S is a Hamming sphere of radius s. Then (9) is tight for f up to a factor of $O\left(s^{3/4}\right)$.*

Proof:

We write p for $1+(1-2\epsilon)^2$. We may and will assume, for simplicity, that $\|f\|_1 = 1$ (which means $f = \frac{2^n}{\binom{n}{s}} \cdot 1_S$). This reduces (9), after taking binary logarithms of both sides, and expanding the definition of η, into

$$\frac{1}{n}\log_2\left(\|f_\epsilon\|_2\right) \quad \leq \quad \frac{1}{2}\tilde{\phi}\left(1 - \frac{p}{p-1}\cdot\frac{1}{n}\log_2\left(\|f\|_p\right),\; 2\epsilon(1-\epsilon)\right) + \frac{p}{p-1}\cdot\frac{1}{n}\log_2\left(\|f\|_p\right).$$

We want to show this is nearly tight for f. Let $\sigma = \frac{s}{n}$. We proceed by comparing both sides of this inequality to $M = \frac{1}{2}\phi(\sigma, 2\epsilon(1-\epsilon)) + 1 - H(\sigma)$.

First, consider the RHS. Let $g(x) = \frac{1}{2}\tilde{\phi}\left(1 - \frac{p}{p-1}\cdot x,\; 2\epsilon(1-\epsilon)\right) + \frac{p}{p-1}\cdot x$. Then the RHS is $g\left(\frac{1}{n}\log_2\left(\|f\|_p\right)\right)$. Observe that $M = g\left(\frac{p-1}{p}(1 - H(\sigma))\right)$. Note also that $\frac{1}{n}\log_2\left(\|f\|_p\right) =$

$\frac{p-1}{p}\left(1-\frac{1}{n}\log_2\left(\binom{n}{s}\right)\right)$. Let $x=\frac{p-1}{p}(1-H(\sigma))$ and let $y=\frac{p-1}{p}\left(1-\frac{1}{n}\log_2\left(\binom{n}{s}\right)\right)$. By (3), we have that $y\geq x$ and that $y-x\leq\frac{p-1}{p}\cdot\left(\frac{1}{2n}\log_2(s)+O\left(\frac{1}{n}\right)\right)$, where the asymptotic notation hides absolute constants. Since $\tilde{\phi}$ increases and $2\geq\tilde{\phi}'\geq\frac{1}{1-2\epsilon(1-\epsilon)}\geq 1$, we have that g increases and $g'\leq\frac{p}{2p-2}$. This implies that $g(y)\geq g(x)$ and $|g(y)-g(x)|\leq\frac{1}{2}\cdot\frac{1}{n}\log_2(s)+O\left(\frac{1}{n}\right)$. In other words, $|\mathrm{RHS}-M|\leq\frac{1}{2n}\log_2(s)+O\left(\frac{1}{n}\right)$.

Next, consider the LHS. We have that it equals (writing δ for $2\epsilon(1-\epsilon)$):

$$\frac{1}{2n}\log_2\left(\langle f_\epsilon,f_\epsilon\rangle\right)\;=\;\frac{1}{2n}\log_2\left(\langle f_\delta,f\rangle\right)\;=\;\frac{1}{2n}\log_2\left(\frac{2^n}{\binom{n}{s}}\sum_{i=0}^{s}\binom{s}{i}\binom{n-s}{i}\delta^{2i}(1-\delta)^{n-2i}\right)\;\geq$$

$$\frac{1}{2n}\log_2\left(\frac{2^n}{\binom{n}{s}}\right)+\frac{1}{2n}\max_{0\leq i\leq s}\log_2\left(\binom{s}{i}\binom{n-s}{i}\delta^{2i}(1-\delta)^{n-2i}\right).$$

In the first step we have used the semigroup property of noise operators. For the second step, see Section 2.4.

By (3) we have $\frac{1}{2n}\log_2\left(\frac{2^n}{\binom{n}{s}}\right)\geq\frac{1-H(\sigma)}{2}+\frac{1}{4n}\log_2(s)-O\left(\frac{1}{n}\right)$. Similarly, by (3):

$$\frac{1}{2n}\max_{0\leq i\leq s}\log_2\left(\binom{s}{i}\binom{n-s}{i}\delta^{2i}(1-\delta)^{n-2i}\right)\;\geq$$

$$\frac{1}{2}\max_{0\leq i\leq s}\left\{\sigma H\left(\frac{i/n}{\sigma}\right)+(1-\sigma)H\left(\frac{i/n}{1-\sigma}\right)+2\frac{i}{n}\log_2(\delta)+\left(1-2\frac{i}{n}\right)\log_2(1-\delta)\right\}-\frac{1}{2n}\log_2(s)-O\left(\frac{1}{n}\right)$$

$$\geq\;\frac{1}{2}\max_{0\leq x\leq\sigma}\left\{\sigma H\left(\frac{x}{\sigma}\right)+(1-\sigma)H\left(\frac{x}{1-\sigma}\right)+2x\log_2(\delta)+(1-2x)\log_2(1-\delta)\right\}-\frac{1}{2n}\log_2(s)-O\left(\frac{1}{n}\right).$$

The second inequality is by Lemma 2.12. Summing up, we have that the LHS is bounded from below by

$$\frac{1}{2}\max_{0\leq x\leq\sigma}\left\{\sigma H\left(\frac{x}{\sigma}\right)+(1-\sigma)H\left(\frac{x}{1-\sigma}\right)+2x\log_2(\delta)+(1-2x)\log_2(1-\delta)\right\}+$$

$$\frac{1-H(\sigma)}{2}-\frac{1}{4n}\log_2(s)-O\left(\frac{1}{n}\right)\;=$$

$$=\;\frac{1}{2}\phi\left(\sigma,\;2\epsilon(1-\epsilon)\right)+1-H(\sigma)-\frac{1}{4n}\log_2(s)-O\left(\frac{1}{n}\right)\;=\;M-\frac{1}{4n}\log_2(s)-O\left(\frac{1}{n}\right).$$

The first step follows from the definition of ϕ. Wrapping everything up, we have that LHS $-$ RHS $\leq\frac{3}{4n}\log_2(s)+O\left(\frac{1}{n}\right)$ and hence the hypercontractive inequality is tight for f up to a multiplicative factor of $O\left(s^{3/4}\right)$.

∎

3.2.1 Proof of Corollary 1.9

We may and will assume, by homogeneity, that $\|f\|_1 = 1$. Let $F(p)$ be the (normalized) binary logarithm of the RHS of the inequality in the claim of the corollary. That is $F(p) = \eta_p\left(\frac{1}{n}\log_2\left(\|f\|_p\right),\ \epsilon\right) + \frac{1}{n}\log_2\left(\|f\|_p\right)$. We will show that $F(p)$ increases in p, and hence the claim of the corollary for $p \geq 1 + (1-2\epsilon)^2$ follows from the claim for $p = 1 + (1-2\epsilon)^2$, proved in (9). Expanding the definition of η_p, we have, as in the proof of Proposition 3.5, that

$$F(p) \ = \ \frac{1}{2}\tilde{\phi}\left(1 - \frac{p}{p-1}\cdot\frac{1}{n}\log_2\left(\|f\|_p\right),\ 2\epsilon(1-\epsilon)\right) + \frac{p}{p-1}\cdot\frac{1}{n}\log_2\left(\|f\|_p\right).$$

Since the derivative of $\tilde{\phi}$ with respect to its first argument is bounded from above by 2, it suffices to show that $\frac{p}{p-1}\cdot\log_2\left(\|f\|_p\right)$ is increasing in p to infer that F is increasing. Let $G(t) = \log_2\left(\|f\|_{1/t}\right)$, for $0 < t \leq 1$. The function G is decreasing and convex (this is a consequence of Hölder's inequality, see [14], theorems 196-197). Moreover, $G(1) = 0$, since $\|f\|_1 = 1$. It is easy to see that this implies that $\frac{p}{p-1}\cdot\log_2\left(\|f\|_p\right) = \frac{G\left(\frac{1}{p}\right)-1}{1-\frac{1}{p}}$ is increasing in p. ∎

Proof of Theorem 1.10

Proof of (10)

First, as is observed at the beginning of the proof of Proposition 3.1, it suffices to consider the case $0 \leq k \leq \frac{n}{2}$.

Let $0 \leq \epsilon < \frac{1}{2}$ and consider the action of the noise operator T_ϵ on f. Since $T_\epsilon = \sum_{k=0}^n (1-2\epsilon)^k \Pi_k$ (see Section 1.1.2), we have that $f_\epsilon = \sum_{k=0}^n (1-2\epsilon)^k f_k$, and therefore $\langle f_\epsilon, f_\epsilon\rangle = \sum_{k=0}^n (1-2\epsilon)^{2k}\langle f_k, f_k\rangle$. This implies that $\langle f_k, f_k\rangle$ is upperbounded by $\frac{\langle f_\epsilon, f_\epsilon\rangle}{(1-2\epsilon)^{2k}}$. Taking logarithms of both sides of this inequality, and using Corollary 1.9 in the second step, we get that

$$\frac{1}{n}\log_2\left(\|f_k\|_2\right) \ \leq \ \frac{1}{n}\log_2\left(\|f_\epsilon\|_2\right) - \frac{k}{n}\log_2(1-2\epsilon) \ \leq \ \eta_p\left(r(p),\epsilon\right) + \frac{1}{n}\log_2\left(\|f\|_p\right) - \frac{k}{n}\log_2(1-2\epsilon).$$

Hence (10) would follow if we verify the identity

$$\min_{0\leq\epsilon\leq\frac{1}{2}}\left\{\eta_p\left(r(p),\epsilon\right) - \frac{k}{n}\log_2(1-2\epsilon)\right\} \ = \ \pi\left(\frac{k}{n}\wedge\frac{n-k}{n}, H^{-1}\left(1 - \frac{p}{p-1}\cdot r(p)\right)\right) - \frac{p-2}{2p-2}\cdot r(p).$$

Writing σ for $H^{-1}\left(1-\frac{p}{p-1}\cdot r(p)\right)$, κ for $\frac{k}{n}$, δ for $2\epsilon(1-\epsilon)$, and expanding all the definitions, this reduces to verifying that for all $0 \leq \kappa, \sigma, \delta \leq \frac{1}{2}$ holds, writing $x = x(\sigma,\delta)$ for $\frac{-\delta^2+\delta\sqrt{\delta^2+4(1-2\delta)\sigma(1-\sigma)}}{2(1-2\delta)}$, that $\pi(\kappa,\sigma) = \pi(\sigma,\kappa)$ is given by

$$\frac{1}{2}\min_{0\leq\delta\leq\frac{1}{2}}\left\{\sigma H\left(\frac{x}{\sigma}\right) + (1-\sigma)\,H\left(\frac{x}{1-\sigma}\right) + 2x\log_2(\delta) + (1-2x)\log_2(1-\delta) - \kappa\log_2(1-2\delta)\right\}.$$

This is shown in Lemma 2.9, and (10) follows.

Proof of near tightness of (10) for spheres

Proposition 3.6: *Let f be proportional to 1_S, where S is a Hamming sphere of radius s. Then (10) is tight for f in the following sense: If $k \leq \frac{n}{2} - \sqrt{s(n-s)}$, then (10) is tight for f up to a factor of $O\left(k^{1/4}\right)$. Moreover, (10) is tight for f up to a factor of $O(t)$, provided k is a point at which the ℓ_2 norm of K_s is attained, up to a factor of t.*

Proof:

Let S be the Hamming sphere of radius s around 0. Let $f = 1_S$. Recall (see Section 1.1.2) that $\widehat{f} = \frac{1}{2^n} \cdot K_s$, and therefore for any $0 \leq k \leq n$ holds $f_k = \frac{1}{2^n} K_s(k) \cdot K_k$. Hence $\|f_k\|_2 = \frac{\sqrt{\binom{n}{k}}}{2^n} |K_s(k)|$. In particular, $\|f_k\|_2 = \|f_{n-k}\|_2$, and in the following argument it suffices to consider the case $0 \leq k \leq \frac{n}{2}$.

Let R be the RHS of (10). Then $\frac{1}{n}\log_2\left(\|f\|_p\right) = 1 - \frac{1}{n}\log_2\left(\binom{n}{s}\right)$ and $r(p) = \frac{p-1}{p} \cdot \left(1 - \frac{1}{n}\log_2\left(\binom{n}{s}\right)\right)$. Substituting, we get $\frac{1}{n}\log_2(R) = \pi\left(\frac{k}{n}, H^{-1}\left(\frac{1}{n}\log_2\left(\binom{n}{s}\right)\right)\right) - \frac{1}{2}\left(1 - \frac{1}{n}\log_2\left(\binom{n}{s}\right)\right)$.

We consider two cases.

1. $0 \leq k \leq \frac{n}{2} - \sqrt{s(n-s)}$.

   In this case, see (11), we have $K_s(k) \geq \frac{\binom{n}{s}}{2^{H\left(\frac{s}{n}\right) \cdot n}} \cdot 2^{\tau\left(\frac{s}{n}, \frac{k}{n}\right) \cdot n}$. Hence, recalling that $\tau(x,y) = \pi(y,x) - \frac{H(y) - H(x) - 1}{2}$, we have, after some rearrangement, that

   $$\frac{1}{n}\log_2(\|f_k\|_2) \;\geq\; \pi\left(\frac{k}{n}, \frac{s}{n}\right) - \frac{1}{2} \cdot \left(1 - \frac{1}{n}\log_2\left(\binom{n}{s}\right)\right) + \frac{1}{2} \cdot \left(\frac{1}{n}\log_2\left(\binom{n}{s}\right) - H\left(\frac{s}{n}\right)\right) + \frac{1}{2} \cdot \left(\frac{1}{n}\log_2\left(\binom{n}{k}\right) - H\left(\frac{k}{n}\right)\right).$$

   By the monotonicity of π we have that $\pi\left(\frac{k}{n}, H^{-1}\left(\frac{1}{n}\log_2\left(\binom{n}{s}\right)\right)\right) \leq \pi\left(\frac{k}{n}, \frac{s}{n}\right)$ and hence

   $$\frac{1}{n}\log_2(R) - \frac{1}{n}\log_2(\|f_k\|_2) \;\leq\; \frac{1}{2} \cdot \left(H\left(\frac{s}{n}\right) - \frac{1}{n}\log_2\left(\binom{n}{s}\right)\right) + \frac{1}{2} \cdot \left(H\left(\frac{k}{n}\right) - \frac{1}{n}\log_2\left(\binom{n}{k}\right)\right) \;\leq\; \frac{1}{4}\log_2(ks) + O(1),$$

   where in the last inequality we have used (3). This proves the first part of the proposition.

2. K_s attains its ℓ_2 norm on k up to a factor of t.

   This means that $\frac{1}{2^n}\binom{n}{k} K_s^2(k) \geq \Omega\left(\frac{1}{t^2}\right) \cdot \|K_s\|_2^2 = \Omega\left(\frac{1}{t^2}\right) \cdot \binom{n}{s}$, which implies $|K_s(k)| \geq \Omega\left(\frac{1}{t}\right) \cdot \sqrt{\frac{\binom{n}{s} 2^n}{\binom{n}{k}}}$. Hence $\|f_k\|_2 \geq \Omega\left(\frac{1}{t}\right) \cdot \sqrt{\frac{\binom{n}{s}}{2^n}}$ and, recalling that π is non-positive,

   $$\frac{1}{n}\log_2(R) - \frac{1}{n}\log_2(\|f_k\|_2) \;\leq\; \pi\left(\frac{k}{n}, \frac{1}{n}\log_2\left(\binom{n}{s}\right)\right) - \log_2(t) + O(1) \;\leq\; \log_2(t) + O(1).$$

   This proves the second part of the proposition.

∎

To complete the proof of the tightness of (10) for spheres recall that, by Corollary 2.16, between any two consecutive roots of K_s there is a point on which K_s attains its ℓ_2 norm up to a factor of $O\left(n^{5/2}\right)$.

∎

4 Proof of Theorem 1.3 and related results

We first deduce Corollary 1.4 from Theorem 1.3.

Proof: (of Corollary 1.4):

Let g be a polynomial of degree s. Write $g = \sum_{r=0}^{s} a_r f_r$, where f_r is a homogeneous polynomial of degree r, $0 \leq r \leq s$. By the triangle inequality for the ℓ_p norm we have $\|f\|_p \leq \sum_{r=0}^{s} |a_r| \|f_r\|_p$. On the other hand, the Parseval identity gives $\|f\|_2 = \sqrt{\sum_{r=0}^{s} a_r^2 \|f\|_r^2} \geq \max_{0 \leq r \leq s} |a_r| \|f_r\|_2$. Note also that (5) is equivalent to $\frac{\|f\|_p}{\|f\|_2} \leq 2^{\frac{\psi(p,\frac{s}{n})}{p} \cdot n}$.

Taking all of the above into account, and recalling that $\psi(p,x)$ increases in x, we have that

$$\frac{\|g\|_p}{\|g\|_2} \;\leq\; \frac{\sum_{r=0}^{s} |a_r| \|f_r\|_p}{\max_{0 \leq r \leq s} |a_r| \|f_r\|_2} \;\leq\; \sum_{r=0}^{s} \frac{\|f_r\|_p}{\|f_r\|_2} \;\leq\; \sum_{r=0}^{s} 2^{\frac{\psi(p,\frac{r}{n})}{p} \cdot n} \;\leq\; n \cdot 2^{\frac{\psi(p,\frac{s}{n})}{p} \cdot n}.$$

We proceed with a tensorization argument. For an integer $m \geq 1$, let $G_m = g^{\otimes m}$. Note that G_m is a polynomial of degree at most sm on $\{0,1\}^{nm}$. By the above,

$$\frac{\|g\|_p}{\|g\|_2} \;=\; \left(\frac{\|G_m\|_p}{\|G_m\|_2}\right)^{\frac{1}{m}} \;\leq\; \left(nm \cdot 2^{\frac{\psi(p,\frac{s}{n})}{p} \cdot nm}\right)^{\frac{1}{m}} \;=\; (nm)^{\frac{1}{m}} \cdot 2^{\frac{\psi(p,\frac{s}{n})}{p} \cdot n}.$$

Taking m to infinity gives $\frac{\|g\|_p}{\|g\|_2} \leq 2^{\frac{\psi(p,\frac{s}{n})}{p} \cdot n}$.

∎

We proceed with the proof of Theorem 1.3. First, we introduce some notation. Let $R(n,s,p)$ be the maximum of the ratio $\frac{||f||_p^p}{||f||_2^p}$ over all homogeneous polynomials of degree s on $\{0,1\}^n$. Let $r(n,s,p) = \frac{||K_s||_p^p}{||K_s||_2^p}$. Then (5) becomes

$$R(n,s,p) \;\leq\; 2^{\psi\left(p,\frac{s}{n}\right) \cdot n}.$$

The key step required to show this is the following claim.

Theorem 4.1: *Let $p \geq 2$ be fixed. Then, for all $0 \leq s \leq n/2$ holds*

$$R(n,s,p) \;\leq\; 2^{O\left(\frac{n}{\log(n)}\right)} \cdot r(n,s,p).$$

Here the constant in the asymptotic notation may depend on p.

The inequality (5) will follow from Theorem 4.1 and the following limit estimate.

Lemma 4.2: *For any integers $n \geq 1$ and $0 \leq s \leq \frac{n}{2}$, and for any $p \geq 2$ holds*

$$\lim_{m\to\infty} \Big(r(nm, sm, p)\Big)^{\frac{1}{m}} \;=\; 2^{\psi\left(p,\frac{s}{n}\right)\cdot n}.$$

In fact, assume Theorem 4.1 and Lemma 4.2 to hold. Let f be a homogeneous polynomial of degree s on $\{0,1\}^n$, such that $R(n,s,p) = \frac{\mathbb{E}\, f^p}{\mathbb{E}^{p/2}\, f^2}$. We proceed with a tensorization argument. For an integer $m \geq 1$, let $F_m = f^{\otimes m}$. Then F_m is a homogeneous polynomial of degree sm on $\{0,1\}^{nm}$. Hence,

$$R(n,s,p) \;=\; \frac{\mathbb{E}\, f^p}{\mathbb{E}^{p/2}\, f^2} \;=\; \left(\frac{\mathbb{E}\, F_m^p}{\mathbb{E}^{p/2}\, F_m^2}\right)^{\frac{1}{m}} \;\leq\; R(nm, sm, p)^{\frac{1}{m}} \;\leq\; \left(2^{O\left(\frac{nm}{\log(nm)}\right)} \cdot r(nm, sm, p)\right)^{\frac{1}{m}},$$

where the second inequality follows from Theorem 4.1. Taking m to infinity, and using Lemma 4.2, gives $R(n,s,p) \leq 2^{\psi\left(p,\frac{s}{n}\right)\cdot n}$, establishing the first claim of the theorem.

The second claim of Theorem 1.3 will be dealt with in the following proposition.

Proposition 4.3: *There is an absolute constant C such that for any integers $n \geq 1$ and $0 \leq s \leq \frac{n}{2}$, and for any $p \geq 2$ holds*

$$2^{n\cdot\psi\left(p,\frac{s}{n}\right)} \;\leq\; n \cdot C^p \cdot s^{\frac{p}{4}} \cdot r(n,s,p).$$

In the remainder of this section we prove Theorem 4.1. Lemma 4.2 and Proposition 4.3 will be proved in Section 4.5.

Notation: For the duration of this section let $s_0 = s_0(n) = \frac{n}{\ln n}$. Let $\epsilon = \epsilon(n) = \frac{n^{\frac{11}{2}} \ln^{\frac{1}{2}} n}{s_0^6}$. Note that $\epsilon(n)$ behaves like $\frac{1}{\sqrt{n}}$, up to polylogarithmic factors. The proof of Theorem 4.1 will rely on the following four claims.

Lemma 4.4: *Theorem 4.1 holds for all $s \leq s_0$.*

Proof: We use (2). Since $r(n,s,p) \geq 1$, we have

$$R(n,s,p) \;\leq\; (p-1)^{\frac{ps}{2}} \;\leq\; (p-1)^{\frac{ps_0}{2}} \;\leq\; 2^{\frac{p\log_2(p-1)}{2}\cdot\frac{n}{\ln n}} \;\leq\; 2^{\frac{p\log_2(p-1)}{2}\cdot\frac{n}{\ln n}} \cdot r(n,s,p).$$

∎

Lemma 4.5: *Theorem 4.1 holds for all $\frac{n}{2} - s_0 \leq s \leq \frac{n}{2}$.*

Proof: Let δ_0 be the characteristic function of 0. Clearly,

$$R(n,s,p) \;\leq\; \frac{\mathbb{E}\,\delta_0^p}{\mathbb{E}^{p/2}\,\delta_0^2} \;=\; 2^{\left(\frac{p}{2}-1\right)n}.$$

On the other hand, recall that $K_s(0) = ||K_s||_2^2 = \binom{n}{s}$. Note also that, by (3), we have $\binom{n}{s} \geq \binom{n}{\frac{n}{2}-s_0} \geq \Omega\left(\frac{1}{\sqrt{n}}\right)\cdot 2^{H\left(\frac{1}{2}-\frac{s_0}{n}\right)\cdot n} \geq 2^{n-O\left(\frac{s_0^2}{n}\right)}$. Hence

$$r(n,s,p) \;\geq\; \frac{\frac{1}{2^n}\cdot K_s^p(0)}{||K_s||_2^p} \;=\; \frac{\binom{n}{s}^{p/2}}{2^n} \;\geq\; \frac{2^{\left(\frac{p}{2}-1\right)n}}{2^{O\left(\frac{ps_0^2}{n}\right)}} \;\geq\; 2^{-O\left(\frac{n}{\ln^2 n}\right)}\cdot R(n,s,p).$$

∎

The proofs of the next two claims are harder. We will first state the claims and show how to deduce Theorem 4.1 from the preceding two lemmas and these two claims and then prove the claims.

Proposition 4.6: *There exists an explicitly defined (see (16)) function $F = F_p$ of two nonnegative variables x and y such that*

1. *The function F is increasing in both x and y is 1-homogeneous.*
2. *For any $1 \leq s \leq (n+1)/2$ the following inductive relation holds*

$$R(n+1,s,p) \quad\leq\quad F\left(R(n,s,p), R(n,s-1,p)\right).$$

Proposition 4.7: *There exists a sufficiently large constant n_0 such that for all $n \geq n_0$ and for all $s_0(n+1) \leq s \leq (n+1)/2 - s_0(n+1)$ holds*

$$r(n+1,s,p) \quad\in\quad (1 \pm O(\epsilon))^{2p}\cdot F\Big(r(n,s,p), r(n,s-1,p)\Big).$$

We now prove Theorem 4.1, assuming the four claims above to hold, and proceeding similarly to [19]. We will show by induction on n that for all n and for all $1 \leq s \leq \frac{n}{2}$ holds $R(n,s,p) \leq 2^{c\frac{n}{\log(n)}}\cdot r(n,s,p)$, for some constant c which may depend on p.

For any fixed n_0, we may assume, by choosing c to be sufficiently large, that the claim holds for $n \leq n_0$, which takes care of the base step. We pass to the induction step. Assume the claim holds for n and we will show that it holds for $n+1$ as well. Let $1 \leq s \leq (n+1)/2$ be given. We may and will assume that $n \geq n_0$, for a sufficiently large n_0. By Lemmas 4.4 and 4.5 the claim holds for $s \leq s_0 = s_0(n+1)$ and for $s \geq (n+1)/2 - s_0$. So we may assume $s_0 < s < (n+1)/2 - s_0$.

Let $R_0 = R(n,s,p)$ and $R_1 = R(n,s-1,p)$. By Proposition 4.6 $R(n+1,s,p) \leq F\left(R_0,R_1\right)$. Let $\rho = \max\left\{\frac{R_0}{r(n,s,p)}, \frac{R_1}{r(n,s-1,p)}\right\}$. By the induction hypothesis $\rho \leq 2^{c\frac{n}{\log(n)}}$. By the monotonicity and 1-homogeneity of F given in Proposition 4.6, and by Proposition 4.7, we have that

$$F\left(R_0,R_1\right) \;\leq\; F\Big(\rho\cdot r(n,s,p), \rho\cdot r(n,s-1,p)\Big) \;=\; \rho\cdot F\Big(r(n,s,p), r(n,s-1,p)\Big) \;\leq$$

$$2^{c\frac{n}{\log(n)}} \cdot F\Big(r(n,s,p), r(n,s-1,p)\Big) \;\leq$$

$$2^{c\frac{n}{\log(n)}} \cdot (1+O(\epsilon))^{2p} \cdot r(n+1,s,p) \;\leq\; 2^{c\frac{n+1}{\log(n+1)}} \cdot r(n+1,s,p),$$

completing the proof of Theorem 4.1. Note that the last inequality holds, for a sufficiently large n, since $\epsilon = \tilde{O}\left(\frac{1}{\sqrt{n}}\right)$.

∎

4.1 Proof of Proposition 4.6

Let p be given. We start with defining the function $F = F_p$ at a point (x,y) where $x, y \geq 0$. If $y = 0$, let $F(x,y) = x$. If $y \neq 0$, let $\rho = \left(\frac{x}{y}\right)^{2/p}$. Let $P(z) = \frac{\left(\sqrt{z}+1\right)^p + \left|\sqrt{z}-1\right|^p}{2}$. We define $F(x,y)$ by

$$F(x,y) \;\;=\;\; y \cdot \sup_{\beta\in[0,\infty)} \frac{P(\rho\beta)}{(\beta+1)^{p/2}}. \tag{16}$$

By definition, F is clearly 1-homogeneous. Since, as is easy to see, P increases in z, for $z \geq 0$, we also have that F is increasing in x. To see that F increases in y, substitute $\alpha = \rho\beta$ and note that $F(x,y) = x \cdot \sup_{\alpha\in[0,\infty)} \frac{P(\alpha)}{(\alpha+\rho)^{p/2}}$.

We now proceed similarly to the proof of Proposition 4.5 in [19].

Let f be a homogeneous polynomial of degree s over $\{0,1\}^{n+1}$, such that $\frac{\mathbb{E}\, f^p}{\mathbb{E}^{p/2} f^2} = R(n+1,s,p)$. For $i = 0,1$ let f_i be the restriction of f to the n-dimensional subcube $\{x : x_{n+1} = i\}$. We view both of these subcubes as isomorphic to $\{0,1\}^n$. Note that there is a homogeneous polynomial g_0 of degree s over $\{0,1\}^n$ and a homogeneous polynomial g_1 of degree $s-1$ over $\{0,1\}^n$, such that $f_0 = g_0 + g_1$ and $f_1 = g_0 - g_1$. We sum up the above by writing $f \leftrightarrow (g_0+g_1, g_0-g_1)$.

Let us first deal with the case in which one of the functions g_i vanishes. If $g_1 = 0$ then $f \leftrightarrow (g_0, g_0)$, and hence $R(n+1,s,p) = \frac{\mathbb{E}\, f^p}{\mathbb{E}^{p/2} f^2} = \frac{\mathbb{E}\, g_0^p}{\mathbb{E}^{p/2} g_0^2} \leq R(n,s,p)$. To see that the claim of the proposition holds it remains to verify that $x \leq F(x,y)$. This however is true, since $F(x,0) = y$ and F increases in y. Similarly, if $g_0 = 0$, we have $R(n+1,s,p) \leq R(n,s-1,p)$. In this case we need to verify $y \leq F(x,y)$. This is true, since $F(x,y) \geq yP(0) = y$.

From now on we assume that both g_i do not vanish. Let $R_0 = \frac{\mathbb{E}\, g_0^p}{\mathbb{E}^{p/2} g_0^2}$, and let $R_1 = \frac{\mathbb{E}\, g_1^p}{\mathbb{E}^{p/2} g_1^2}$. Note that $R_0 \leq R(n,s,p)$ and $R_1 \leq R(n,s-1,p)$. We use Hanner's inequality [30]: For $p \geq 2$ holds

$$\|g_0+g_1\|_p^p + \|g_0-g_1\|_p^p \;\leq\; (\|g_0\|_p + \|g_1\|_p)^p + \Big|\|g_0\|_p - \|g_1\|_p\Big|^p.$$

This implies that

$$R(n+1,s,p) \;=\; \frac{\mathbb{E}\, f^p}{\mathbb{E}^{p/2} f^2} = \frac{\frac{1}{2}\cdot(\|g_0+g_1\|_p^p + \|g_0-g_1\|_p^p)}{\left(\mathbb{E}\, g_0^2 + \mathbb{E}\, g_1^2\right)^{p/2}} \;\leq$$

$$\frac{\frac{1}{2}\cdot\left((\|g_0\|_p+\|g_1\|_p)^p+\left|\|g_0\|_p-\|g_1\|_p\right|^p\right)}{\left(\mathbb{E}\,g_0^2+\mathbb{E}\,g_1^2\right)^{p/2}} \;=\; R_1\cdot\frac{\frac{1}{2}\cdot\left(\left(\frac{\|g_0\|_p}{\|g_1\|_p}+1\right)^p+\left|\frac{\|g_0\|_p}{\|g_1\|_p}-1\right|^p\right)}{\left(\frac{\mathbb{E}\,g_0^2}{\mathbb{E}\,g_1^2}+1\right)^{p/2}}.$$

Let $\rho=\left(\frac{R_0}{R_1}\right)^{2/p}$, and $\beta=\frac{\mathbb{E}\,g_0^2}{\mathbb{E}\,g_1^2}$. Then $\frac{\|g_0\|_p}{\|g_1\|_p}=\sqrt{\rho\beta}$, and the last expression can be written as

$$R_1\cdot\frac{\frac{1}{2}\cdot\left(\left(\sqrt{\rho\beta}+1\right)^p+\left|\sqrt{\rho\beta}-1\right|^p\right)}{(\beta+1)^{p/2}} \;=\; R_1\cdot\frac{P(\rho\beta)}{(\beta+1)^{p/2}} \;\le\; F\left(R_0,R_1\right),$$

where the last inequality follows from the definition of F.

∎

4.2 Proof of Proposition 4.7

There are two functions on $[2,\infty]\times\left[0,\frac{1}{2}\right]$ which will play an important role in the following argument. The first of these functions is the function $h(p,x)$ defined in Section 2.1.3. We define the second function to be $g(p,x)=x^{\frac{1}{p}}(1-x)^{\frac{p-1}{p}}-x^{\frac{p-1}{p}}(1-x)^{\frac{1}{p}}$. Note that g is nonnegative. For fixed p we will frequently omit the first variable and view h and g as functions of x only.

Given n, $s\le n/2$, and p, we define $i_0=i_0(n,s,p)$ to be the unique real number in the interval $[0,n/2]$ satisfying

$$1-\frac{2s}{n} \;=\; h\left(p,\frac{i_0}{n}\right). \tag{17}$$

We now define several quantities depending on n,s,p and i_0. Assume $s>0$. Let $t=t(n,s,p)=\frac{(n-2i_0)+\sqrt{(n-2i_0)^2-4s(n-s)}}{2(n-s)}$. Let

$$\rho(n,s,p) \;=\; \frac{n-2i_0}{s}\cdot t-1, \tag{18}$$

and let

$$\Phi(n,s,p) \;=\; \frac{n}{2(n-i_0)}\cdot\left(\frac{s}{n}\right)^{p/2}\cdot\left(1+\frac{n-s}{s}\cdot t\right)^p. \tag{19}$$

The claim of Proposition 4.7 will be based on the following two claims.

Proposition 4.8: *Let F be the function defined in (16). Then, assuming $0<s<n/2$, we have*

$$\Phi(n,s,p) \;=\; F\left(\rho^{p/2}(n,s,p),1\right)$$

Proposition 4.9: *There exists a sufficiently large constant n_0 such that for all $n \geq n_0$ and for all $s_0 \leq s \leq \frac{n}{2} - s_0$ holds*

1.
$$\left(\frac{r(n,s,p)}{r(n,s-1,p)}\right)^{2/p} \in (1 \pm O(\epsilon))^2 \cdot \rho(n,s,p).$$

2.
$$\frac{r(n+1,s,p)}{r(n,s-1,p)} \in (1 \pm O(\epsilon))^p \cdot \Phi(n,s,p).$$

We first derive Proposition 4.7 from these two claims and then prove the claims. By 1-homogeneity of F, the claim of the proposition is equivalent to

$$\frac{r(n+1,s,p)}{r(n,s-1,p)} \in (1 \pm O(\epsilon))^{2p} \cdot F\left(\frac{r(n,s,p)}{r(n,s-1,p)}, 1\right)$$

Assume Propositions 4.8 and 4.9 to hold. By the monotonicity of F in both coordinates and by its 1-homogeneity, we have that

$$\frac{r(n+1,s,p)}{r(n,s-1,p)} \in (1 \pm O(\epsilon))^p \cdot \Phi(n,s,p) = (1 \pm O(\epsilon))^p \cdot F\left(\rho^{p/2}(n,s,p), 1\right) \subseteq$$

$$(1 \pm O(\epsilon))^{2p} \cdot F\left(\frac{r(n,s,p)}{r(n,s-1,p)}, 1\right).$$

∎

4.3 Proof of Proposition 4.8

First, we observe that $\rho(n,s,p)$ lies between 1 and $p-1$. This will be the contents of the following lemma.

Lemma 4.10: *For all $0 < s < n/2$ holds*

$$1 < \rho(n,s,p) < p-1.$$

Proof: Let $x = \frac{i_0}{n}$. Then by (17) we have $0 < x < 1/2$, and $h(x) = 1 - \frac{s}{n}$. Hence $\frac{s}{n} = \frac{1-h(x)}{2}$. In particular, s is a function of x, and hence so is t. In fact, $t = \frac{1-2x+g(x)}{1+h(x)}$. To see this, observe that a simple calculation gives

$$t = \frac{(n-2i_0) + \sqrt{(n-2i_0)^2 - 4s(n-s)}}{2(n-s)} = \frac{(1-2x) + \sqrt{(1-2x)^2 - (1-h^2(x))}}{1+h(x)}.$$

Note that $h^2(x) - g^2(x) = 4x(1-x)$, and hence the last expression is indeed $\frac{1-2x+g(x)}{1+h(x)}$.

Next, we write $\rho = \rho(n,s,p)$ as a function of x as well: $\rho = \frac{1-2x+g(x)}{1-2x-g(x)}$. This can be verified by a simple calculation, using again the identity $h^2(x) = g^2(x) + 4x(1-x)$:

$$\rho \;=\; \frac{n-2i}{s}\cdot t - 1 \;=\; \frac{2(1-2x)(1-2x+g(x))}{1-h^2(x)} - 1 \;=$$

$$\frac{\Big((1-2x)+g(x)\Big)^2}{(1-2x)^2-g^2(x)} \;=\; \frac{1-2x+g(x)}{1-2x-g(x)}.$$

Since $g > 0$ for $0 < x < 1/2$, this implies that $\rho > 1$.

Next, we argue that $\rho < p-1$. This is equivalent to $g < \frac{p-2}{p}\cdot(1-2x)$. Since both sides of this putative inequality vanish at $1/2$, it suffices to show that $g' > -\frac{2p-4}{p}$. Computing g' and rearranging, we have that

$$g' \;=\; \frac{1}{p}\cdot\left(\left(\frac{1-x}{x}\right)^{\frac{p-1}{p}} + \left(\frac{x}{1-x}\right)^{\frac{p-1}{p}}\right) - \frac{p-1}{p}\cdot\left(\left(\frac{1-x}{x}\right)^{\frac{1}{p}} + \left(\frac{x}{1-x}\right)^{\frac{1}{p}}\right).$$

Let $\gamma = \frac{1}{2}\cdot\left(\left(\frac{1-x}{x}\right)^{\frac{1}{p}} + \left(\frac{x}{1-x}\right)^{\frac{1}{p}}\right)$. Then $\gamma > 1$ and by convexity of the function $t \to t^{p-1}$ we have $g' \geq \frac{2}{p}\cdot\gamma^{p-1} - \frac{2p-2}{p}\cdot\gamma$. Using this, it remains to verify the inequality $\gamma^{p-1} - (p-1)\gamma > -(p-2)$, for $\gamma > 1$, and this is true, since it holds with equality for $\gamma = 1$, and the derivative of the LHS is positive for $\gamma > 1$.

∎

Next, we consider $F\left(\rho^{p/2}(n,s,p), 1\right)$. Let $x = \rho^{p/2}(n,s,p)$ and $y = 1$. By the preceding lemma $\rho = \rho(n,s,p) = \left(\frac{x}{y}\right)^{2/p}$ lies between 1 and $p-1$. Recall that $F(x,y) = y\cdot\sup_{\beta\in[0,\infty)}\frac{P(\rho\beta)}{(\beta+1)^{p/2}}$. Let $f(\beta) = \frac{P(\rho\beta)}{(\beta+1)^{p/2}}$. The following lemma describes the behavior of f when $1 < \rho < p-1$.

Lemma 4.11: *Assume $1 < \rho < p-1$. Then f increases from 0 to some point $\frac{1}{\rho} < \beta^* < \infty$ and decreases from β^* on. In particular,* $\sup_{\beta\in[0,\infty)}\frac{P(\rho\beta)}{(\beta+1)^{p/2}} \;=\; \frac{P(\rho\beta^*)}{(\beta^*+1)^{p/2}}$.

We will prove the lemma below. Here we proceed assuming that it holds. By the lemma, we may restrict our attention to the behavior of f on $\left(\frac{1}{\rho},\infty\right]$. In this interval $P(z) = \frac{(\sqrt{z}+1)^p + (\sqrt{z}-1)^p}{2}$. It will be convenient to make the one-to-one substitution $u = \frac{1}{\rho(\beta+1)}$. Then $0 < u < \frac{1}{\rho+1}$ and $f(\beta) = \rho^{p/2}\cdot Q(u)$, where $Q(u) = \frac{1}{2}\cdot\left(\left(\sqrt{1-\rho u}+\sqrt{u}\right)^p + \left(\sqrt{1-\rho u}-\sqrt{u}\right)^p\right)$. In particular, $F(x,y) = y\cdot\rho^{p/2}\cdot Q\left(u^*\right) = \rho^{p/2}\cdot Q\left(u^*\right)$, where u^* is the unique point in the interval $\left(0,\frac{1}{2}\right)$ in which the derivative Q' vanishes. A simple calculation gives that u^* is implicitly given by the following identity (writing u for u^* for simplicity):

$$\left(\sqrt{1-\rho u}+\sqrt{u}\right)^{2k-1}\cdot\left(\sqrt{1-\rho u}-\rho\sqrt{u}\right) \quad=\quad \left(\sqrt{1-\rho u}-\sqrt{u}\right)^{2k-1}\cdot\left(\sqrt{1-\rho u}+\rho\sqrt{u}\right).$$

Given this, the claim of the proposition is immediately implied by the following two lemmas.

Lemma 4.12: *For $\rho = \rho(n,s,p)$ holds*

$$u^* \quad = \quad \frac{s}{\rho n}.$$

And

Lemma 4.13:

$$\Phi(n,s,p) \quad = \quad \rho^{p/2} \cdot Q\left(u^*\right)$$

It remains to prove Lemmas 4.11 - 4.13.

Proof of Lemma 4.11.

We partition $[0,\infty]$ into two subintervals $\left[0,\frac{1}{\rho}\right]$ and $\left[\frac{1}{\rho},\infty\right)$. The claim of the lemma is implied by the following two claims.

1. On $\frac{1}{\rho} \leq \beta < \infty$ the function f increases up to some point $\frac{1}{\rho} < \beta^* < \infty$ and decreases from β^* on.
2. On $0 \leq \beta \leq \frac{1}{\rho}$ the unction f increases.

The case $\frac{1}{\rho} \leq \beta < \infty$: On this interval $f(\beta) = \frac{\left(\sqrt{\rho\beta}+1\right)^p + \left(\sqrt{\rho\beta}-1\right)^p}{(2\beta+1)^{p/2}}$ and, after some rearrangement and simplification, f' is proportional to

$$\frac{\beta+1}{\beta} \cdot \sqrt{\rho\beta} \cdot \left(\left(\sqrt{\rho\beta}+1\right)^{p-1} + \left(\sqrt{\rho\beta}-1\right)^{p-1}\right) - \left(\left(\sqrt{\rho\beta}+1\right)^{p} + \left(\sqrt{\rho\beta}-1\right)^{p}\right).$$

Set $z = \sqrt{\rho\beta}$. Then $z \geq 1$ and the above is proportional to $\frac{z^2+\rho}{z} - \frac{(z+1)^p+(z-1)^p}{(z+1)^{p-1}+(z-1)^{p-1}}$.

Let

$$t(z) \quad = \quad \frac{(z+1)^p + (z-1)^p}{(z+1)^{p-1} + (z-1)^{p-1}} \cdot z - z^2 \quad = \quad \frac{z \cdot \left((z+1)^{p-1} - (z-1)^{p-1}\right)}{(z+1)^{p-1} + (z-1)^{p-1}}.$$

Note that the sign of f' is the same of that of $\rho - t(z)$. Hence, recalling that $1 < \rho < p-1$, the claim will follow if we show that the function $t(z)$ strictly increases from 1 to $p-1$ on $[1,\infty)$.

First, it is easy to see that $t(1) = 1$ and that $t(z) \to_{z\to\infty} p-1$. Next, we claim that $t' > 0$, which is the same as

$$\left(\left((z+1)^{p-1} - (z-1)^{p-1}\right) + (p-1)z \cdot \left((z+1)^{p-2} - (z-1)^{p-2}\right)\right) \cdot \left((z+1)^{p-1} + (z-1)^{p-1}\right) >$$

$$(p-1)z\cdot\left(\left((z+1)^{p-1}-(z-1)^{p-1}\right)\cdot\left((z+1)^{p-2}+(z-1)^{p-2}\right)\right).$$

Rearranging and simplifying, this is the same as

$$(z+1)^{2p-2}-(z-1)^{2p-2} \;>\; 4(p-1)z\cdot\left(z^2-1\right)^{p-2}.$$

Consider the function $r(z)=z^{2p-2}$. Since $2p-2\geq 2$ we have that $r'''\geq 0$ and hence (by developing r into Taylor series around z, up to the second term) that $r(z+1)-r(z-1)\geq 2r'(z)=4(p-1)z^{2p-3}$. Hence it suffices to show

$$4(p-1)z^{2p-3} \;>\; 4(p-1)z\cdot\left(z^2-1\right)^{p-2},$$

which is evidently true for $z\geq 1$.

The case $0\leq\beta\leq\frac{1}{\rho}$: On this interval $f(\beta)=\frac{\left(\sqrt{\rho\beta}+1\right)^p+\left(1-\sqrt{\rho\beta}\right)^p}{(2\beta+1)^{p/2}}$. We have that $f'(0)=\infty$. Next, we consider $f'(\beta)$ for $\beta>0$. We have, similarly to the above, that f' is proportional to

$$\frac{\beta+1}{\beta}\cdot\sqrt{\rho\beta}\cdot\left(\left(\sqrt{\rho\beta}+1\right)^{p-1}-\left(\sqrt{1-\rho\beta}\right)^{p-1}\right)-\left(\left(\sqrt{\rho\beta}+1\right)^{p}+\left(\sqrt{1-\rho\beta}\right)^{p}\right).$$

Set $z=\sqrt{\rho\beta}$. Then $0<z\leq 1$ and the above is proportional to $\frac{z^2+\rho}{z}-\frac{(z+1)^p+(1-z)^p}{(z+1)^{p-1}-(1-z)^{p-1}}$.

Let

$$t(z) \;=\; \frac{(z+1)^p+(1-z)^p}{(z+1)^{p-1}-(1-z)^{p-1}}\cdot z-z^2 \;=\; \frac{z\cdot\left((z+1)^{p-1}+(1-z)^{p-1}\right)}{(z+1)^{p-1}-(1-z)^{p-1}}.$$

The sign of f' is the same of that of $\rho-t(z)$. Hence the claim will follow if we show that the function $t(z)$ strictly increases from $\frac{1}{p-1}$ to 1 on $\left[0,\frac{1}{\rho}\right]$.

First, it is easy to see that $\lim_{z\to 0}t(0)=\frac{1}{p-1}$ and that $t(1)=1$.

Next, we claim that $t'>0$, which, similarly to the discussion above is the same as

$$(z+1)^{2p-2}-(1-z)^{2p-2} \;>\; 4(p-1)z\cdot\left(z^2-1\right)^{p-2}.$$

Consider the function $r(u)=u^{2p-2}$. Since $2p-2\geq 2$ we have that $r'''\geq 0$ and hence that $r(1+z)-r(1-z)\geq 2zr'(1)=4(p-1)z$, which is evidently larger then the RHS above.

∎

Proof of Lemma 4.12.

Let $u=\frac{s}{\rho n}$. Clearly $0<u\leq\frac{1}{2\rho}<\frac{1}{\rho+1}$. So, it remains to verify that u satisfies (20). As in the proof of Lemma 4.10, we will write everything as a function of $x=\frac{i_0}{n}$, reducing to an identity involving the functions $g=g(x)$ and $h=h(x)$, which we then proceed to verify.

First, we observe that $u=\frac{1}{\rho}\cdot\frac{s}{n}=\frac{1-2x-g}{1-2x+g}\cdot\frac{1-h}{2}$. We also have $1-\rho u=1-\frac{s}{n}=\frac{1+h}{2}$, and that $\rho^2u=\rho\cdot\frac{s}{n}=\frac{1-2x+g}{1-2x-g}\cdot\frac{1-h}{2}$.

Using this, substituting in (20) and simplifying, we need to show

$$\Big(\sqrt{(1+h)(1-2x+g)}+\sqrt{(1-h)(1-2x-g)}\Big)^{p-1}\cdot\Big(\sqrt{(1+h)(1-2x-g)}-\sqrt{(1-h)(1-2x+g)}\Big)=$$

$$\Big(\sqrt{(1+h)(1-2x+g)}-\sqrt{(1-h)(1-2x-g)}\Big)^{p-1}\cdot\Big(\sqrt{(1+h)(1-2x-g)}+\sqrt{(1-h)(1-2x+g)}\Big).$$

Next, we multiply both sides by

$$\Big(\sqrt{(1+h)(1-2x+g)}+\sqrt{(1-h)(1-2x-g)}\Big)^{p-1}\cdot\Big(\sqrt{(1+h)(1-2x-g)}+\sqrt{(1-h)(1-2x+g)}\Big).$$

We observe that

$$(1+h)(1-2x+g)\cdot(1-h)(1-2x-g)=(1+h)(1-2x-g)\cdot(1-h)(1-2x+g)=\big(1-h^2\big)\cdot\big((1-2x)^2-g^2\big)=\big(1-h^2\big)^2$$

and hence, after some simplification,

$$\Big(\sqrt{(1+h)(1-2x+g)}+\sqrt{(1-h)(1-2x-g)}\Big)^2 \;=\; 2\cdot\Big((1-2x)+gh+\big(1-h^2\big)\Big),$$

and

$$\Big(\sqrt{(1+h)(1-2x-g)}+\sqrt{(1-h)(1-2x+g)}\Big)^2 \;=\; 2\cdot\Big((1-2x)-gh+\big(1-h^2\big)\Big).$$

In addition, after some simplification, we have

$$(1+h)(1-2x+g)-(1-h)(1-2x-g)=2\cdot\Big((1-2x)h+g\Big)$$

and

$$(1+h)(1-2x-g)-(1-h)(1-2x+g)=2\cdot\Big((1-2x)h-g\Big).$$

Taking all this into account, we need to show that

$$\Big((1-2x)+gh+\big(1-h^2\big)\Big)^{p-1}\cdot\Big((1-2x)h-g\Big)=\Big((1-2x)-gh+\big(1-h^2\big)\Big)\cdot\Big((1-2x)h+g\Big)^{p-1} \quad (21)$$

It's not hard to verify that

$$(1-2z)+gh+\big(1-h^2\big) \;=\; 2\cdot\Big((1-x)^2-x^{\frac{2p-2}{p}}(1-x)^{\frac{2}{p}}\Big),$$

that

$$(1-2z)-gh+\big(1-h^2\big) \;=\; 2\cdot\Big((1-x)^2-x^{\frac{2}{p}}(1-x)^{\frac{2p-2}{p}}\Big),$$

that

$$(1-2z)h+g \;=\; 2x(1-x)\cdot\left(\left(\frac{1-x}{x}\right)^{\frac{p-1}{p}}-\left(\frac{x}{1-x}\right)^{\frac{p-1}{p}}\right),$$

and that

$$(1-2z)h-g \quad = \quad 2x(1-x)\cdot\left(\left(\frac{1-x}{x}\right)^{\frac{1}{p}}-\left(\frac{x}{1-x}\right)^{\frac{1}{p}}\right).$$

We can now complete the proof of the fact that (20) holds. In fact, substituting the above and simplifying, it is not hard to see that both sides of (21) are equal to

$$2^{p/2}\cdot x^{\frac{2p-2}{p}}(1-x)^{\frac{3p-3}{p}}\cdot\left((1-x)^{\frac{2p-2}{p}}-x^{\frac{2p-2}{p}}\right)^{p-1}\cdot\left((1-x)^{\frac{2}{p}}-x^{\frac{2}{p}}\right).$$

This completes the proof of Lemma 4.12

∎

Proof of Lemma 4.13.

We again reduce the claim to an algebraic identity involving the functions $g=g(z)$ and $h=h(z)$, which we proceed to verify. Recalling that $Q(u)=\frac{1}{2}\cdot\left(\left(\sqrt{1-\rho u}+\sqrt{u}\right)^p+\left(\sqrt{1-\rho u}-\sqrt{u}\right)^p\right)$, taking $u=u^*=\frac{s}{\rho n}=\frac{1-2z-g}{1-2z+g}\cdot\frac{1-h}{2}$, and recalling the definition of $\Phi(n,s,p)$, we need to verify the identity

$$\rho^{p/2}\cdot\left(\left(\sqrt{1-\rho u}+\sqrt{u}\right)^p+\left(\sqrt{1-\rho u}-\sqrt{u}\right)^p\right) \quad = \quad \frac{n}{n-i}\cdot\left(\frac{s}{n}\right)^{p/2}\cdot\left(1+\frac{n-s}{s}\cdot t\right)^p. \quad (22)$$

We proceed by expressing everything via the functions g and h, as above. It is not hard to see that the LHS of (22) is equal to

$$\frac{\left(\sqrt{(1+h)(1-2z+g)}+\sqrt{(1-h)(1-2z-g)}\right)^p+\left(\sqrt{(1+h)(1-2z+g)}-\sqrt{(1-h)(1-2z-g)}\right)^p}{2^{p/2}\cdot(1-2z-g)^{p/2}},$$

and the RHS of (22) equals to

$$\frac{1}{1-z}\cdot\left(\frac{1-h}{2}\right)^{p/2}\cdot\left(\frac{2-2z+g-h}{1-h}\right)^p \quad = \quad \frac{1}{1-z}\cdot\frac{1}{2^{p/2}(1-h)^{p/2}}\cdot(2-2z+g-h)^p.$$

Rearranging, we need to show that

$$\left(\sqrt{(1+h)(1-2z+g)}+\sqrt{(1-h)(1-2z-g)}\right)^p+\left(\sqrt{(1+h)(1-2z+g)}-\sqrt{(1-h)(1-2z-g)}\right)^p =$$

$$\frac{1}{1-z}\cdot\left(\frac{1-2z-g}{1-h}\right)^{p/2}\cdot(2-2z+g-h)^p.$$

We start with simplifying the LHS of this putative identity. Recall that

$$\left(\sqrt{(1+h)(1-2z+g)}+\sqrt{(1-h)(1-2z-g)}\right)^2 \;=\; 2\cdot\Big((1-2z)+gh+\left(1-h^2\right)\Big).$$

Similarly, it is easy to see that

$$\left(\sqrt{(1+h)(1-2z+g)}-\sqrt{(1-h)(1-2z-g)}\right)^2 \;=\; 2\cdot\left((1-2z)+gh-\left(1-h^2\right)\right).$$

Hence, the LHS is $2^{p/2}$ times

$$\left((1-2z)+gh+\left(1-h^2\right)\right)^{p/2}+\left((1-2z)+gh-\left(1-h^2\right)\right)^{p/2}.$$

Recall that we have

$$(1-2z)+gh+\left(1-h^2\right) \quad=\quad 2\cdot\left((1-z)^2-z^{\frac{2p-2}{p}}(1-z)^{\frac{2}{p}}\right)$$

Similarly, it is easy to see that

$$(1-2z)+gh-\left(1-h^2\right) \quad=\quad 2\cdot\left(z(1-z)\left(\frac{1-z}{z}\right)^{\frac{p-2}{p}}-z^2\right).$$

Substituting and simplifying, it's not hard to verify that the LHS is

$$2^p\cdot\left((1-z)^{\frac{2p-2}{p}}-z^{\frac{2p-2}{p}}\right)^{p/2}.$$

The RHS is harder to simplify, but we can write it as

$$\frac{1}{1-z}\cdot 2^p\cdot\left(\frac{(1-2z-g)\left(1-z+\frac{g-h}{2}\right)^2}{1-h}\right)^{p/2}.$$

Simplifying and rearranging, veryfying that these two expressions are equal amounts to verifying that

$$\frac{(1-2z-g)\left(1-z+\frac{g-h}{2}\right)^2}{1-h} \quad=\quad (1-z)^2-z^{\frac{2p-2}{p}}(1-z)^{\frac{2}{p}}.$$

Substituting the definitions of g and h, this is equivalent to

$$\left((1-2z)-\left(z^{\frac{1}{p}}(1-z)^{\frac{p-1}{p}}-z^{\frac{p-1}{p}}(1-z)^{\frac{1}{p}}\right)\right)\cdot\left((1-z)-z^{\frac{p-1}{p}}(1-z)^{\frac{1}{p}}\right)^2 \;=$$

$$\left(1-\left(z^{\frac{1}{2k}}(1-z)^{\frac{p-1}{p}}+z^{\frac{p-1}{p}}(1-z)^{\frac{1}{p}}\right)\right)\cdot\left((1-z)^2-z^{\frac{2p-2}{p}}(1-z)^{\frac{2}{p}}\right).$$

Writing $a=1-z$ and $b=z^{\frac{p-1}{p}}(1-z)^{\frac{1}{p}}$, the second term on the LHS is $(a-b)^2$, while the second term on the RHS is a^2-b^2. So we can divide out by $a-b$, and have to show that

$$\left((1-2z)-\left(z^{\frac{1}{p}}(1-z)^{\frac{p-1}{p}}-z^{\frac{p-1}{p}}(1-z)^{\frac{1}{p}}\right)\right)\cdot\left((1-z)-z^{\frac{p-1}{p}}(1-z)^{\frac{1}{p}}\right) \;=$$

$$\Big(1-\Big(z^{\frac{1}{p}}(1-z)^{\frac{p-1}{p}}+z^{\frac{p-1}{p}}(1-z)^{\frac{1}{p}}\Big)\Big)\cdot\Big((1-z)+z^{\frac{p-1}{p}}(1-z)^{\frac{1}{p}}\Big)$$

Write $A := z^{\frac{1}{p}}(1-z)^{\frac{p-1}{p}}$ and $B := z^{\frac{p-1}{p}}(1-z)^{\frac{1}{p}}$. Then the last identity is

$$\Big((1-2z)-(A-B)\Big)\cdot\Big((1-z)-B\Big) \;=\; \Big(1-(A+B)\Big)\cdot\Big((1-z)+B\Big)$$

Simplifying and rearranging, this is the same as $AB = z(1-z)$, which is true.

∎

4.4 Proof of Proposition 4.9

Recall that we assume that n is large and that $s_0 \le s \le \frac{n}{2} - s_0$, where $s_0 = s_0(n) = \frac{n}{\ln n}$.

We first observe that under this assumption, the value of i_0 given by (17) is bounded away from 0 and from $\frac{n}{2}$.

Lemma 4.14: *Let $0 \le i_0 \le \frac{n}{2}$ be given by (17). Then*

$$\Big(\frac{s_0}{n}\Big)^p \quad\le\quad \frac{i_0}{n} \quad\le\quad \Big(\frac{1}{2}-\sqrt{\frac{s}{n}\Big(1-\frac{s}{n}\Big)}\Big) - \Omega\left(\Big(\frac{p-2}{p}\Big)^2\cdot\Big(\frac{s_0}{n}\Big)^4\right)$$

Here the asymptotic notation hides absolute factors.

Proof:

Let $x = \frac{i_0}{n}$, $y = \frac{1}{2} - \sqrt{\frac{s}{n}\left(1-\frac{s}{n}\right)}$. We start with the first inequality, since it is easier. The derivative of h is computed in the proof of Lemma 2.4 below, and it is easy to see that for all $0 < z \le 1/2$ holds $h'(z) \le \frac{p-1}{p}\cdot z^{-\frac{1}{p}} + \frac{1}{p}\cdot z^{-\frac{p-1}{p}} \le z^{-\frac{1}{p}} + z^{-\frac{p-1}{p}}$. Since $h(0) = 0$, it follows that $h(z) \le z^{\frac{p-1}{p}} + z^{\frac{1}{p}} \le 2z^{\frac{1}{p}}$. Hence $h(x) = 1 - \frac{2s}{n} \ge \frac{2s_0}{n}$ implies $x \ge \big(\frac{s_0}{n}\big)^p$, completing the first inequality.

We pass to the second inequality. Recall that $h^2(z) = g^2(z) + 4z(1-z)$. Hence we have, observing that $4y(1-y) = \big(1-\frac{2s}{n}\big)^2$, that

$$h(y) \;=\; \sqrt{4y(1-y)+g^2(y)} \;=\; \sqrt{\Big(1-\frac{2s}{n}\Big)^2+g^2(y)} \;\ge\; \Big(1-\frac{2s}{n}\Big)+\frac{g^2(y)}{4},$$

where the last inequality follows from the following easily verifiable clam: Let $0 \le a, \epsilon \le 1$. Then $\sqrt{a^2+\epsilon} \ge a + \frac{\epsilon}{4}$.

Next, we have

$$g(y) = (y(1-y))^{\frac{1}{p}}\cdot\Big((1-y)^{\frac{p-2}{p}} - y^{\frac{p-2}{p}}\Big) \;\ge\; \left(\frac{1}{4}\cdot\Big(1-\frac{2s}{n}\Big)^2\right)^{\frac{1}{p}}\cdot\frac{2p-4}{p}\cdot\sqrt{\frac{s}{n}\Big(1-\frac{s}{n}\Big)} \;\ge$$

$$\left(\frac{1}{4}\cdot\left(1-\frac{2s}{n}\right)^2\right)^{\frac{1}{2}}\cdot\frac{2p-4}{p}\cdot\sqrt{\frac{s}{n}\left(1-\frac{s}{n}\right)} \;=\; \frac{p-2}{p}\cdot\left(1-\frac{2s}{n}\right)\cdot\sqrt{\frac{s}{n}\left(1-\frac{s}{n}\right)}.$$

To see the first inequality, observe that for $0 \le t \le 1$ holds $\left(t^{\frac{p-2}{p}}\right)' \ge \frac{p-2}{p}$, and hence that $(1-y)^{\frac{p-2}{p}} - y^{\frac{p-2}{p}} \ge \frac{p-2}{p}\cdot(1-2y)$. It follows that

$$h(y) \;\ge\; \left(1-\frac{2s}{n}\right)+\left(\frac{p-2}{2p}\right)^2\cdot\left(1-\frac{2s}{n}\right)^2\cdot\frac{s}{n}\left(1-\frac{s}{n}\right).$$

Let $\delta = \left(\frac{p-2}{2p}\right)^2\cdot\left(1-\frac{2s}{n}\right)^2\cdot\frac{s}{n}\left(1-\frac{s}{n}\right)$. Recall that by definition $h(x) = 1-\frac{2s}{n}$. Hence we get $h(y) \ge h(x)+\delta$. Computing h' as in the proof of Lemma 2.4, it is easy to see that for all $0 < z < 1$ holds $0 < h'(z) < \frac{1}{z}$. From this, $\delta \le h(y)-h(x) \le \frac{y-x}{x}$, which implies to $x \le y - \frac{\delta}{1+\delta}\cdot y \le y - \frac{\delta}{2}\cdot y$ (the last inequality follows since clearly $\delta \le \frac{1}{2}$).

Recalling that $y = \frac{1}{2}-\sqrt{\frac{s}{n}\left(1-\frac{s}{n}\right)}$, and that by assumption $s_0 \le s \le \frac{n}{2}-s_0$, it is easy to check that $\frac{\delta}{2}\cdot y \ge \Omega\left(\left(\frac{p-2}{p}\right)^2\cdot\left(\frac{s_0}{n}\right)^4\right)$, and the claim of the lemma holds.

∎

Remark 4.15: We will assume from now on, to avoid complications in notation arising from replacing i_0 by the nearest integer, that i_0 is integer, whenever it is convenient for us to do so. It is easy to see that the error this introduces is negligible. ∎

The key step in the proof of Proposition 4.9 is the following claim, which may be of independent interest.

Proposition 4.16: *Let $p > 2$ be fixed. Let $0 < i_0 < \frac{n}{2}$ satisfy (17). Then the ℓ_p norm of K_s is attained, up to a small error, in a union of intervals of length $O\left(\sqrt{n\log n}\right)$ around i_0 nad $n-i_0$. More precisely, there is an absolute constant C such that if I is the interval of length $C\sqrt{n\log n}$ around i_0 then, for a sufficiently large n, depending on p, and for any $\sigma \le s \le n/2-\sigma$ holds*

$$\frac{1}{2^n}\sum_{i\in I\cup(n-I)}\binom{n}{i}\left(|K_s(i)|\right)^p \;\ge\; \left(1-O\left(\frac{1}{n^2}\right)\right)\cdot||K_s||_p^p.$$

This proposition and the argument leading towards its proof will have the following corollary as in easy implication. We write a superscript $K_s^{(n)}$ for the Krawchouk polynomial on the n-dimensional cube, when we consider functions on cubes of different dimensions (in the second claim of the corollary).

Corollary 4.17: *Let i_0 be given by (17). Then*

1.

$$\frac{||K_s||_p^p}{||K_{s-1}||_p^p} \in (1 \pm O(\epsilon))^p \cdot \frac{K_s^p(i_0)}{K_{s-1}^p(i_0)}$$

2.

$$\frac{||K_s^{(n+1)}||_p^p}{||K_{s-1}^{(n)}||_p^p} \in (1 \pm O(\epsilon))^p \cdot \frac{\frac{\binom{n+1}{i_0}}{2^{n+1}} \cdot \left(K_s^{(n+1)}(i_0)\right)^p}{\frac{\binom{n}{i_0}}{2^n} \cdot \left(K_{s-1}^{(n)}(i_0)\right)^p}$$

Looking ahead, the first claim of Proposition 4.9 will be a simple consequence of the first claim of this corollary, and the second claim of the proposition will follow easily from the second claim of the corollary. We will prove Proposition 4.16 and Corollary 4.17 and, following this, complete the proof of Proposition 4.9.

We proceed with the proof of Proposition 4.16.

Lemma 4.18: *Let $0 < i_0 < \frac{n}{2}$ be given by (17). Let $n^{\frac{2}{3}} \ll \Delta \ll \frac{n}{2} - \sqrt{s(n-s)} - i_0$. Let $i_1 = \frac{n}{2} - \sqrt{s(n-s)} - \Delta$. Then for any $0 \le i \le i_0$ holds*

$$\frac{\binom{n}{i+1}K_s^p(i+1)}{\binom{n}{i}K_s^p(i)} \ge \frac{i_0}{i+1} \cdot \left(1 - O\left(\frac{s}{\Delta^2}\right)\right)^p,$$

for any $i_0 < i < i_1$ holds

$$\frac{\binom{n}{i+1}K_s^p(i+1)}{\binom{n}{i}K_s^p(i)} \le \frac{i_0}{i+1} \cdot \left(1 + O\left(\frac{s}{\Delta^2}\right)\right)^p,$$

and for any $i_1 \le i < x_s - 1$ holds

$$\frac{\binom{n}{i+1}K_s^p(i+1)}{\binom{n}{i}K_s^p(i)} \le \frac{i_0}{i_1+1} \cdot \left(1 + O\left(\frac{s}{\Delta^2}\right)\right)^p.$$

Proof:

We will need the two following facts: The location of the first root x_s of the Krawchouk polynomial K_s and the behaviour of the values of K_s in the interval $(0, x_s)$. Recall that (see e.g., [24]) we have

$$\frac{n}{2} - \sqrt{s(n-s)} \le x_s \le \frac{n}{2} - \sqrt{s(n-s)} + O\left(n^{\frac{2}{3}}\right). \tag{23}$$

The following fact has been shown in [18, 32] (see also proof of Lemma 5.1 in [41] for a more detailed calculation). Let $0 \le i \le x_s - \Delta$ for some $\Delta \gg \sqrt{s}$. Then

$$\frac{K_s(i+1)}{K_s(i)} \in \left(1, 1 \pm O\left(\frac{s}{\Delta^2}\right)\right) \cdot \frac{(n-2s) + \sqrt{(n-2s)^2 - 4i(n-i)}}{2(n-i)} \tag{24}$$

We proceed with the proof. We will prove the first inequality. The second is similar. The third follows from the second immediately, since the ratio $\frac{\binom{n}{i+1}K_s^p(i+1)}{\binom{n}{i}K_s^p(i)}$ decreases in i on $0 \leq i \leq x_s - 1$ (to see this note that clearly the ratio of the binomial coefficients decreases, and as observed in the proof of the preceding lemma, the ratio $\frac{K_s(i+1)}{K_s(i)}$ decreases as well).

Observe that (17) means that $\frac{(n-2s)+\sqrt{(n-2s)^2-4i_0(n-i_0)}}{2(n-i_0)} = \left(\frac{i_0}{n-i_0}\right)^{\frac{1}{p}}$. To see this, note that, as in the proof of Lemmas 4.12 and 4.13 above, we can rewrite this equaity in terms of the variable $x = \frac{i_0}{n}$ and the functions g and h of this variable. It transforms into $\frac{h(x)+\sqrt{h^2(x)-4x(1-x)}}{2(1-x)} = \left(\frac{x}{1-x}\right)^{\frac{1}{p}}$, which is the same as $h+g = 2x^{\frac{1}{p}}(1-x)^{\frac{p-1}{p}}$, and this follows directly from the definition of g and of h.

Hence, we have (using (24) and Lemma 4.14) that

$$\frac{\binom{n}{i+1}K_s^p(i+1)}{\binom{n}{i}K_s^p(i)} = \frac{n-i}{i+1}\left(\frac{K_s(i+1)}{K_s(i)}\right)^p \geq \frac{n-i}{i+1}\left(\frac{K_s(i_0+1)}{K_s(i_0)}\right)^p \geq$$

$$\left(1-O\left(\frac{s}{\Delta^2}\right)\right)^p \cdot \frac{n-i}{i+1}\cdot\left(\frac{(n-2s)+\sqrt{(n-2s)^2-4i_0(n-i_0)}}{2(n-i_0)}\right)^p =$$

$$\left(1-O\left(\frac{s}{\Delta^2}\right)\right)^p \cdot \frac{n-i}{i+1}\cdot\frac{i_0}{n-i_0} \geq \left(1-O\left(\frac{s}{\Delta^2}\right)\right)^p \cdot \frac{i_0}{i+1}.$$

∎

Lemma 4.19: *Let* $0 \leq i \leq x_s - \Delta$ *for some* $\Delta \gg \frac{n^{\frac{5}{6}}}{s_0^{\frac{1}{6}}}$. *Then*

$$\frac{\binom{n}{i+1}K_s^2(i+1)}{\binom{n}{i}K_s^2(i)} \geq 1+\Omega\left(\frac{\Delta s_0^{\frac{1}{2}}}{n^{\frac{3}{2}}}\right).$$

Proof: We have, by (23) and (24) that

$$\frac{\binom{n}{i+1}K_s^2(i+1)}{\binom{n}{i}K_s^2(i)} = \frac{n-i}{i+1}\cdot\left(\frac{K_s(i+1)}{K_s(i)}\right)^2 \geq$$

$$\left(1-O\left(\frac{s}{\Delta^2}\right)\right)^2 \cdot \frac{n-i}{i+1}\cdot\left(\frac{(n-2s)+\sqrt{(n-2s)^2-4i(n-i)}}{2(n-i)}\right)^2 \geq$$

$$\left(1-O\left(\frac{s}{\Delta^2}\right)\right)^2 \cdot \frac{(n-2s)^2}{4(n-i)(i+1)}$$

The quadratic $Q(x) = 4x(n-x)$ equals $(n-2s)^2$ at $x = \frac{n}{2} - \sqrt{s(n-s)}$. It is easy to see that this means that for $i < \frac{n}{2} - \sqrt{s(n-s)} - \Omega(\Delta)$ (this estimate on i is valid by (23) and by

our assumption on Δ), we have $4(i+1)(n-i) \leq (n-2s)^2 - \Omega\left(\Delta \cdot (n-2i)\right) \leq (n-2s)^2 - \Omega\left(\Delta \cdot \sqrt{s(n-s)}\right)$. Hence

$$\frac{(n-2s)^2}{4(n-i)(i+1)} \;\geq\; 1+\Omega\left(\frac{\Delta\cdot\sqrt{s(n-s)}}{(n-2s)^2}\right) \;\geq\; 1+\Omega\left(\frac{\Delta s_0^{\frac{1}{2}}}{n^{\frac{3}{2}}}\right).$$

∎

We can now complete the proof of Proposition 4.16, using the three auxiliary claims above.

Proof of Proposition 4.16.

Let $\Delta = n^{\frac{4}{5}}$. By Corollary 2.16 and by Lemma 4.19, there exists $x_s - \Delta \leq i' \leq x_s$ such that, say, $\binom{n}{i'}K_s^2(i') \geq \frac{1}{n^3}\cdot 2^n\binom{n}{s}$, which means that $K_s(i') \geq \sqrt{\frac{1}{n^3}\cdot\frac{2^n\binom{n}{s}}{\binom{n}{i'}}}$. Let $i_1 = \lfloor x_s - \Delta \rfloor$. Let $D = i_1 - i_0$. Note that by Lemma 4.14 we have $D \geq \Omega\left(\frac{s_0^4}{n^3}\right)$, and hence by the third claim of Lemma 4.18 we have that

$$\binom{n}{i_1}K_s^p(i_1) \;\geq\; \binom{n}{i'}K_s^p(i') \;\geq\; \frac{1}{n^{3p/2}}\cdot\frac{2^{pn/2}\binom{n}{s}^{p/2}}{\binom{n}{i_1}^{p/2-1}}.$$

By Lemma 4.18,

$$\frac{\binom{n}{i_0}K_s^p(i_0)}{\binom{n}{i_1}K_s^p(i_1)} \;\geq\; \left(1+O\left(\frac{s}{\Delta^2}\right)\right)^{-pD}\cdot\frac{(i_0+1)(i_0+2)...(i_0+D)}{i_0^D} \;\geq$$

$$e^{-\frac{pDn}{\Delta^2}}\cdot e^{\Omega\left(\frac{D^2}{n}\right)} \;\geq\; e^{\Omega\left(\frac{D^2}{n}\right)} \;\geq\; e^{\Omega\left(\frac{s_0^8}{n^7}\right)}.$$

This implies that

$$\binom{n}{i_0}K_s^p(i_0) \;\geq\; e^{\Omega\left(\frac{\sigma^8}{n^7}\right)}\cdot\binom{n}{i_1}K_s^p(i_1) \;\geq$$

$$e^{\Omega\left(\frac{s_0^8}{n^7}\right)}\cdot\frac{1}{n^{3p/2}}\cdot\frac{2^{pn/2}\binom{n}{s}^{p/2}}{\binom{n}{i_1}^{p/2-1}} \;\geq\; e^{\Omega\left(\frac{s_0^8}{n^7}\right)}\cdot\frac{2^{pn/2}\binom{n}{s}^{p/2}}{\binom{n}{i_1}^{p/2-1}}.$$

Next, for any $i_1 \leq i \leq n/2$ holds $\binom{n}{i}K_s^2(i) \leq 2^n||K_s||_2^2 = 2^n\binom{n}{s}$, and hence $\binom{n}{i}\left(|K_s|(i)\right)^p \leq \frac{2^{pn/2}\binom{n}{s}^{p/2}}{\binom{n}{i}^{p/2-1}} \leq \frac{2^{pn/2}\binom{n}{s}^{p/2}}{\binom{n}{i_1}^{p/2-1}}$. This implies that

$$\binom{n}{i_0}K_s^p(i_0) \;\geq\; e^{\Omega\left(\frac{s_0^8}{n^7}\right)}\cdot\left(\sum_{i_1\leq i\leq n/2}\binom{n}{i}\left(|K_s|(i)\right)^p\right).$$

In addition, by Lemma 4.18, similarly to the above, we have that for a sufficiently large constant C holds that if $0 \leq i < i_1$ and $|i-i_0| \geq C\cdot\sqrt{n\log n}$ then $\frac{\binom{n}{i_0}K_s^p(i_0)}{\binom{n}{i}K_s^p(i)} \geq n^3$. Let I be the interval

$\left[i_0 - C \cdot \sqrt{n \log n}, i_0 + C \cdot \sqrt{n \log n}\right]$. Then by the above, we have $\sum_{i \notin I, i \leq n/2} \binom{n}{i} (|K_s|(i))^p \leq O\left(\frac{1}{n^2}\right) \cdot \binom{n}{i_0} K_s^p(i_0)$.

Finally, recall that K_s is symmetric around $n/2$ if s is even, and antisymmetric if s is odd. Taking all of this into account, we have

$$\sum_{i \in I \cup (n-I)} \binom{n}{i} (|K_s(i)|)^p \;\geq\; \left(1 - O\left(\frac{1}{n^2}\right)\right) \cdot ||K_s||_p^p,$$

completing the proof of the proposition.

∎

Proof of Corollary 4.17.

We start with the first part of the corollary. By Proposition 4.16 applied to both K_{s-1} and K_s there is an interval I of length $O\left(\sqrt{n \log n}\right)$ around i_0, where i_0 is defined by (17) such that both the ℓ_p norms of K_{s-1} and K_s are attained, up to a factor of $1 - O\left(\frac{1}{n^2}\right)$ on I and on $n - I$. Taking into account the symmetry (or anti-symmetry) of the Krawchouk polynomials around $\frac{n}{2}$, the claim of the corollary will follow if we show that for all i in I holds $\frac{K_s(i)}{K_{s-1}(i)} \in (1 \pm O(\epsilon)) \cdot \frac{K_s(i_0)}{K_{s-1}(i_0)}$.

Recall that $\binom{n}{i} K_s(i) = \binom{n}{s} K_i(s)$. Hence $\frac{K_s(i)}{K_{s-1}(i)} = \frac{\binom{n}{s}}{\binom{n}{s-1}} \cdot \frac{K_i(s)}{K_i(s-1)}$.

By the above discussion we know that each i in I satisfies $i \leq x_s - \Omega\left(\frac{s_0^4}{n^3}\right)$. An easy calculation[7] using (23) shows that $i \leq x_s - \Omega\left(\frac{\sigma^4}{n^3}\right)$ implies $s \leq x_i - \Omega\left(\frac{s_0^{4.5}}{n^{3.5}}\right)$. Hence we may apply (24), with roles of s and i reversed, to obtain $\frac{K_i(s)}{K_i(s-1)} \in \left(1 \pm O\left(\frac{n^8}{s_0^9}\right)\right) \cdot \frac{(n-2i) + \sqrt{(n-2i)^2 - 4s(n-s)}}{2(n-s)}$.

This means that

$$\frac{K_s(i)}{K_{s-1}(i)} \;/\; \frac{K_s(i_0)}{K_{s-1}(i_0)} \;\in\; \left(1 \pm O\left(\frac{n^8}{s_0^9}\right)\right) \cdot \frac{(n-2i) + \sqrt{(n-2i)^2 - 4s(n-s)}}{(n-2i_0) + \sqrt{(n-2i_0)^2 - 4s(n-s)}}.$$

By Lemma 4.14, both $n - 2i$ and $n - 2i_0$ are lowerbounded by $\Omega\left(\frac{s_0^5}{n^4}\right)$. Hence $\frac{n-2i}{n-2i_0} \in 1 \pm \frac{n^{9/2} \log^{1/2}(n)}{s_0^5}$. In addition, since both i and i_0 are upperbounded by $x_s - \Omega\left(\frac{s_0^4}{n^3}\right) = \frac{n}{2} - \sqrt{s(n-s)} - \Omega\left(\frac{s_0^4}{n^3}\right)$, it is easy to see that $\frac{(n-2i)^2 - 4s(n-s)}{(n-2i_0)^2 - 4s(n-s)} \in 1 \pm \frac{n^4 \log^{1/2}(n)}{s_0^{9/2}}$. Recalling the definition of $\epsilon = \epsilon(n)$ we see that both ratios lie in $1 \pm \epsilon$ and hence

$$\frac{K_s(i)}{K_{s-1}(i)} \;/\; \frac{K_s(i_0)}{K_{s-1}(i_0)} \;\in\; \left(1 \pm O\left(\frac{n^7}{\sigma^8}\right)\right) \cdot (1 \pm O(\epsilon)) \;\subseteq\; 1 \pm O(\epsilon).$$

We pass to the second part of the corollary. Again, we may focus our attention on an interval I of length $O\left(\sqrt{n \log n}\right)$ around i_0 given by (17) in which (half of) the ℓ_p norms of $K_{s-1}^{(n)}$, $K_s^{(n)}$, and $K_s^{(n+1)}$ are attained, up to a factor of $1 - \frac{1}{n^2}$. Moreover, following the argument above, for

[7] We omit the details.

all $i \in I$ the ratio $\frac{K_s^{(n)}(i)}{K_{s-1}^{(n)}(i)}$ is in $(1 \pm O(\epsilon)) \cdot \frac{K_s^{(n)}(i_0)}{K_{s-1}^{(n)}(i_0)}$. Recall the identity (see [28], but also easy to verify directly) $K_s^{(n+1)}(i) = K_s^{(n)}(i) + K_{s-1}^{(n)}(i)$. This means that for all $i \in I$ holds

$$\frac{K_s^{(n+1)}(i)}{K_{s-1}^{(n)}(i)} = \frac{K_s^{(n)}(i)}{K_{s-1}^{(n)}(i)} + 1 \in (1 \pm O(\epsilon)) \cdot \left(\frac{K_s^{(n)}(i_0)}{K_{s-1}^{(n)}(i_0)} + 1 \right) = (1 \pm O(\epsilon)) \cdot \frac{K_s^{(n+1)}(i_0)}{K_{s-1}^{(n)}(i_0)}.$$

∎

Proof of Proposition 4.9

We start with the first claim of the proposition. Let $\rho_1 = \left(\frac{r(n,s,p)}{r(n,s-1,p)} \right)^{2/p} = \frac{||K_s||_p^2}{\mathbb{E}(K_s)^2} \; / \; \frac{||K_{s-1}||_p^2}{\mathbb{E}(K_{s-1})^2}$. We need to show that $\rho_1 \in (1 \pm O(\epsilon)) \cdot \rho$, where $\rho = \frac{n-2i_0}{s} \cdot t - 1$, and $t = \frac{(n-2i_0)+\sqrt{(n-2i_0)^2 - 4s(n-s)}}{2(n-s)}$. Recall that $\mathbb{E}\, K_s^2 = \binom{n}{s}$ and that $\binom{n}{i} K_s(i) = \binom{n}{s} K_i(s)$. Applying the first claim of Corollary 4.17 and using (24) (with roles of s and i reversed) we have that (estimating $\frac{K_{i_0}(s)}{K_{i_0}(s-1)}$ as in the proof of Corollary 4.17):

$$\rho_1 \in (1 \pm O(\epsilon))^2 \cdot \frac{\binom{n}{s-1}}{\binom{n}{s}} \cdot \left(\frac{K_s(i_0)}{K_{s-1}(i_0)} \right)^2 = (1 \pm O(\epsilon))^2 \cdot \frac{\binom{n}{s}}{\binom{n}{s-1}} \cdot \left(\frac{K_{i_0}(s)}{K_{i_0}(s-1)} \right)^2 \subseteq$$

$$(1 \pm O(\epsilon))^2 \cdot \left(1 \pm O\left(\frac{n^7}{\sigma^8} \right) \right) \cdot \frac{n-s+1}{s} \cdot t^2(n,s) \subseteq (1 \pm O(\epsilon)) \cdot \frac{n-s}{s} \cdot t^2(n,s)$$

Finally, recall that $t = t(n,s)$ is a root of the quadratic $(n-s)t^2 - (n-2i_0)\,t + s = 0$. Hence $t^2 = \frac{(n-2i_0)t - s}{n-s}$. Substituting this in the above expression and simplifying gives the first claim of the proposition.

We proceed to the second claim of the proposition. We need to show that $\frac{r(n+1,s,p)}{r(n-1,s-1,p)} \in (1 \pm O(\epsilon))^p \cdot \Phi(n,s,p)$, where $\Phi(n,s,p) = \frac{n}{2(n-i_0)} \cdot \left(\frac{s}{n}\right)^{p/2} \cdot \left(1 + \frac{n-s}{s} \cdot t\right)^p$.

We have that $\frac{r(n+1,s,p)}{r(n-1,s-1,p)} = \frac{||K_s^{(n+1)}||_p^p}{\mathbb{E}^{p/2}\left(K_s^{(n+1)}\right)^2} \; / \; \frac{||K_{s-1}^{(n)}||_p^p}{\mathbb{E}^{p/2}\left(K_{s-1}^{(n)}\right)^2}$. Applying the second claim of Corollary 4.17 (and replacing, within negligible error, $\frac{n+1}{2(n+1-i_0)}$ by $\frac{n}{2(n-i_0)}$) we have that the RHS of this expression is in

$$(1 \pm O(\epsilon))^p \cdot \frac{n}{2(n-i_0)} \left(\frac{\binom{n}{s-1}}{\binom{n+1}{s}} \right)^{p/2} \left(\frac{K_s^{(n+1)}(i_0)}{K_{s-1}^{(n)}(i_0)} \right)^p \subseteq$$

$$(1 \pm O(\epsilon))^p \cdot \frac{n}{2(n-i_0)} \left(\frac{s}{n+1} \right)^{p/2} \left(1 + \frac{K_s^{(n)}(i_0)}{K_{s-1}^{(n)}(i_0)} \right)^p.$$

Recalling that

$$\frac{K_s^{(n)}(i_0)}{K_{s-1}^{(n)}(i_0)} = \frac{\binom{n}{s}}{\binom{n}{s-1}} \cdot \frac{K_{i_0}^{(n)}(s)}{K_{i_0}^{(n)}(s-1)} = \frac{n-s+1}{s} \cdot \frac{K_{i_0}^{(n)}(s)}{K_{i_0}^{(n)}(s-1)} \in \left(1 \pm O\left(\frac{n^7}{s_0^8} \right) \right) \cdot \frac{n-s}{s} \cdot t,$$

and replacing, within negligible error, $\frac{s}{n+1}$ by $\frac{s}{n}$, we obtain the second claim of the proposition.

∎

4.5 Proofs of Lemma 4.2 and Proposition 4.3

Proof of Lemma 4.2

Let n and $p \geq 2$ be fixed. Let $0 \leq s \leq \frac{n}{2}$. Let m be an integer, and let $N = nm$ and $S = sm$. We need to show that

$$\lim_{m\to\infty} \left(r(N,S,p)\right)^{\frac{1}{m}} \;=\; \lim_{m\to\infty} \left(\frac{\mathbb{E}\left(K_S^{(N)}\right)^p}{\mathbb{E}^{p/2}\left(K_S^{(N)}\right)^2} \right)^{\frac{1}{m}} \;=\; 2^{n\cdot\psi\left(p,\frac{s}{n}\right)},$$

For $s = 0$ the claim of the lemma reduces to verifying that $\psi(p,0) = 0$, and for $s = \frac{n}{2}$ to verifying that $\psi\left(p,\frac{1}{2}\right) = \frac{p-2}{2}$. Both these facts follow easily from the definition of ψ. So we may assume from now on that $0 < s < \frac{n}{2}$.

Consider first the denominator. Recalling that $\mathbb{E}\left(K_S^{(N)}\right)^2 = \binom{N}{S}$ and using (3), we have that

$$\lim_{m\to\infty} \left(\mathbb{E}^{p/2}\left(K_S^{(N)}\right)^2\right)^{\frac{1}{m}} \;=\; \lim_{m\to\infty} \binom{N}{S}^{\frac{p}{2m}} \;=\; \lim_{m\to\infty} 2^{\frac{pN}{2m}H\left(\frac{S}{N}\right)} \;=\; 2^{\frac{1}{2}pnH\left(\frac{s}{n}\right)}.$$

Next, consider the numerator. Recall that $s_0 = s_0(N) = \frac{N}{\ln N}$. For a sufficiently large m, $S = sm$ satisfies $s_0 < S < \frac{N}{2} - s_0$. Hence, by Proposition 4.16, we have that, up to a constant factor, the value of $\mathbb{E}\left(K_S^{(N)}\right)^p$ is given by $\frac{1}{2^N}\sum_{i\in I}\binom{N}{i}\left(K_S^{(N)}(i)\right)^p$, where I is an interval of length $O\left(\sqrt{N\log N}\right)$ around i_0, and i_0 is determined by $h\left(p,\frac{i_0}{N}\right) = 1 - \frac{2S}{N} = 1 - \frac{2s}{n}$. Let i_1 be the leftmost integer point of I and let i_2 be the rightmost point. Then for any $i \in I$ holds $\binom{N}{i}\left(K_S^{(N)}(i)\right)^p \leq \binom{N}{i_2}\left(K_S^{(N)}(i_1)\right)^p$, since the binomial coefficients increase as i increases in I, while the value of $K_S^{(N)}(i)$ decreases. Next, by the Lemma 4.19 (see also Remark 4.15), we have that $\binom{N}{i_0}\left(K_S^{(N)}\right)^2(i_0) \geq \binom{N}{i_1}\left(K_S^{(N)}\right)^2(i_1)$. Hence $K_S^{(N)}(i_1) \leq \sqrt{\frac{\binom{N}{i_0}}{\binom{N}{i_1}}}K_S^{(N)}(i_0)$. Altogether, we have,

$$\frac{\binom{N}{i_0}}{2^N}\left(K_S^{(N)}(i_0)\right)^p \;\leq\; \mathbb{E}\left(K_S^{(N)}\right)^p \;\leq\; O\left(|I|\cdot\frac{\binom{N}{i_2}}{2^N}\cdot\left(\frac{\binom{N}{\lceil i_0\rceil}}{\binom{N}{i_1}}\right)^{p/2}\left(K_S^{(N)}(i_0)\right)^p\right).$$

Taking the limit as m goes to infinity, and using the approximation of the binomial coefficient $\binom{b}{a}$ by $2^{bH(a/b)}$, we have

$$\lim_{m\to\infty}\left(\mathbb{E}\left(K_S^{(N)}\right)^p\right)^{\frac{1}{m}} \;=\; \lim_{m\to\infty}\left(\frac{\binom{N}{i_0}}{2^N}\left(K_S^{(N)}(i_0)\right)^p\right)^{\frac{1}{m}} \;=\; 2^{(H(y)-1)n}\cdot\lim_{m\to\infty}\left(K_S^{(N)}(i_0)\right)^{\frac{p}{m}},$$

where we write y for $\frac{i_0}{N}$, remembering that y is determined by $h(p,y) = 1 - \frac{2s}{n}$.

Note that $i_0 \leq \frac{N}{2} - \sqrt{S(N-S)}$, and hence we may apply (11) to obtain $K_S^{(N)}(i_0) \in 2^{\left(\tau\left(\frac{S}{N},y\right)\pm o(1)\right)\cdot N} = 2^{\left(\tau\left(\frac{s}{n},y\right)\pm o_N(1)\right)\cdot N}$. This gives $\lim_{m\to\infty}\left(K_S^{(N)}(i_0)\right)^{\frac{p}{m}} = 2^{\left(p\cdot\tau\left(\frac{s}{n},y\right)\right)\cdot n}$, and $\lim_{m\to\infty}\left(\mathbb{E}\left(K_S^{(N)}\right)^p\right)^{\frac{1}{m}} = 2^{\left(H(y)-1+p\cdot\tau\left(\frac{s}{n},y\right)\right)\cdot n}$.

Summing up, we have

$$\lim_{m\to\infty}\left(r(N,S,p)\right)^{\frac{1}{m}} \;=\; 2^{\left(H(y)-1+p\cdot\tau\left(\frac{s}{n},y\right)-\frac{p}{2}H\left(\frac{s}{n}\right)\right)\cdot n} = 2^{\psi\left(p,\frac{s}{n}\right)\cdot n},$$

where in the last step we use the first definition of ψ in Section 2.1.4.

∎

Proof of Proposition 4.3

We will need a simple technical lemma.

Lemma 4.20:

$$\max_{0\leq i\leq n/2}\frac{2^{nH\left(\frac{i}{n}\right)}}{\binom{n}{\lfloor i\rfloor}} \;\leq\; O(n).$$

Proof:

For $0 \leq i < 1$ we have $\frac{2^{nH\left(\frac{i}{n}\right)}}{\binom{n}{\lfloor i\rfloor}} \leq 2^{nH\left(\frac{1}{n}\right)} \leq 2^{n\cdot\left(\frac{1}{n}\log_2(n)+O\left(\frac{1}{n}\right)\right)} \leq O(n)$.

Using (3) we have that for $1 \leq i \leq n/2$ holds

$$\frac{2^{nH\left(\frac{i}{n}\right)}}{\binom{n}{\lfloor i\rfloor}} \;\leq\; O\left(\sqrt{i}\cdot 2^{n\cdot\left(H\left(\frac{i}{n}\right)-H\left(\frac{\lfloor i\rfloor}{n}\right)\right)}\right) \;\leq\; O\left(\sqrt{i}\cdot 2^{n\cdot\frac{i-\lfloor i\rfloor}{n}\cdot H'\left(\frac{\lfloor i\rfloor}{n}\right)}\right),$$

where the inequality follows from the concavity of H. Recalling that $H'(x) = \log_2\left(\frac{1-x}{x}\right) \leq \log_2\left(\frac{1}{x}\right)$, the last expression is at most $O\left(\sqrt{i}\cdot\frac{n}{\lfloor i\rfloor}\right) \leq O(n)$.

∎

We proceed with the proof of the proposition. We have that $\psi\left(p,\frac{s}{n}\right) = H(y) - 1 + p\cdot\tau\left(\frac{s}{n},y\right) - \frac{p}{2}H\left(\frac{s}{n}\right)$, where y is determined by $h(p,y) = 1 - 2\frac{s}{n}$. Set $i_0 = \lfloor ny\rfloor$ and observe that $i_0 \leq ny \leq \frac{n}{2} - \sqrt{s(n-s)}$. Hence, by (11), we have

$$K_s(i_0) \;\geq\; \frac{\binom{n}{s}}{2^{H\left(\frac{s}{n}\right)\cdot n}}\cdot 2^{\tau\left(\frac{s}{n},\frac{i_0}{n}\right)\cdot n} \;\geq\; \frac{\binom{n}{s}}{2^{H\left(\frac{s}{n}\right)\cdot n}}\cdot 2^{\tau\left(\frac{s}{n},y\right)\cdot n},$$

where in the second step we have used the fact that τ is decreasing in the second variable.

Hence, using Lemma 4.20 in the third inequality and (3) in the last inequality, we have

$$r(n,s,p) \;=\; \frac{\mathbb{E}\left|K_s\right|^p}{\mathbb{E}^{p/2}K_s^2} \;=\; \frac{\mathbb{E}\left|K_s\right|^p}{\binom{n}{s}^{p/2}} \;\geq\; \frac{\frac{1}{2^n}\cdot\binom{n}{i_0}K_s^p(i_0)}{\binom{n}{s}^{p/2}} \;\geq$$

$$\left(\frac{\binom{n}{s}}{2^{H\left(\frac{s}{n}\right)\cdot n}}\right)^{\frac{p}{2}} \cdot \frac{1}{2^n} \cdot \binom{n}{i_0} 2^{\left(p\tau\left(\frac{s}{n},y\right)-\frac{p}{2}H\left(\frac{s}{n}\right)\right)\cdot n} \;\geq$$

$$\Omega\left(\frac{1}{n}\right)\left(\frac{\binom{n}{s}}{2^{H\left(\frac{s}{n}\right)\cdot n}}\right)^{\frac{p}{2}} \cdot 2^{\left(H(y)-1+p\tau\left(\frac{s}{n},y\right)-\frac{p}{2}H\left(\frac{s}{n}\right)\right)\cdot n} \;=\; \Omega\left(\frac{1}{n}\right)\left(\frac{\binom{n}{s}}{2^{H\left(\frac{s}{n}\right)\cdot n}}\right)^{\frac{p}{2}} \cdot 2^{\psi\left(p,\frac{s}{n}\right)\cdot n} \;\geq$$

$$\Omega\left(\frac{1}{n}\right) \cdot C^{-p} \cdot s^{-\frac{p}{4}} \cdot 2^{\psi\left(p,\frac{s}{n}\right)\cdot n}.$$

∎

5 Appendix: Proofs of claims about univariate and bivariate functions

Proof of Lemma 2.1

The derivative $\frac{\partial r}{\partial y}$ is easily seen to be proportional, up to a positive factor, to $(1-2x)^2 + (1-2x)\sqrt{(1-2x)^2 - 4y(1-y)} - 2(1-y)$. For a fixed y, this is maximized at $x = 0$, in which case this is 0. ∎

Proof of Lemma 2.3

For $y \geq \frac{1}{2} - \sqrt{x(1-x)}$ the claim of the lemma follows immediately from the definition of τ. It should also be possible to verify the claim directly for $y < \frac{1}{2} - \sqrt{x(1-x)}$, but we proceed by observing that in this range the claim follows immediately from the reciprocity of Krawchouk polynomials (property 1 in Section 2.2), from (11), from (3), and from the continuity of the function τ. ∎

Proof of Lemma 2.4

For the first claim of the lemma, the values of h at the endpoints of x are easy to verify. And, it is easy to see that for $0 < x < 1/2$ holds

$$\frac{\partial h}{\partial x} \;=\; \frac{p-1}{p} \cdot \left(\left(\frac{1-x}{x}\right)^{\frac{1}{p}} - \left(\frac{x}{1-x}\right)^{\frac{1}{p}}\right) + \frac{1}{p} \cdot \left(\left(\frac{1-x}{x}\right)^{\frac{p-1}{p}} - \left(\frac{x}{1-x}\right)^{\frac{p-1}{p}}\right) \;>\; 0.$$

For the second claim of the lemma, writing $h(p,x) = \sqrt{x(1-x)} \cdot \left(\left(\frac{1-x}{x}\right)^{\frac{p-2}{2p}} + \left(\frac{x}{1-x}\right)^{\frac{p-2}{2p}}\right)$, it is easy to see that for a fixed $0 < x < 1/2$ this is a strongly increasing function in p. It is also easy to see that $h(2,x) = 2\sqrt{x(1-x)}$ and that $h(p,x) \to_{p\to\infty} 1$.

For the third claim of the lemma, let z be such that $h(2,z) = 1-2x$. Since $h(2,z) = 2\sqrt{z(1-z)}$, this is equivalent (after rearranging) to $x = \frac{1}{2} - \sqrt{z(1-z)}$, which is the same as $z = \frac{1}{2} - \sqrt{x(1-x)}$. Since $h(p,z)$ increases in p this means that $h(p,z) > 1-2x$. Since $h(p,u)$ increases in u, and $h(p,y) = 1-2x$, this implies $y < z = \frac{1}{2} - \sqrt{x(1-x)}$. ∎

Proof of Lemma 2.5

We view a as a function of δ for a fixed p. The boundary values of a are easy to verify. It remains to check that a decreases. We will show that $b = 1 - 2a = \delta(1-\delta) \cdot \frac{(1-\delta)^{p-2}+\delta^{p-2}}{(1-\delta)^p+\delta^p}$ increases. Let $g(\delta) = (1-\delta)^{p-2} + \delta^{p-2}$, and $h(\delta) = (1-\delta)^p + \delta^p$. Then $a = \delta(1-\delta) \cdot \frac{g}{h}$, and $a' = \frac{(1-2\delta)gh+\delta(1-\delta)(g'h-gh')}{h^2}$. We will show that the numerator is positive, which will imply $a' > 0$.

Computing and simplifying, we have that $g'h - gh' = (1-2\delta)\big(\delta(1-\delta)\big)^{p-3} \cdot \big(2\delta(1-\delta) + (p-2)\big)$, and that $gh = (1-\delta)^{2p-2} + \delta^{2p-2} + \big((1-\delta)^2 + \delta^2\big)\big(\delta(1-\delta)\big)^{p-2}$. Substituting and simplifying, we get that

$$(1-2\delta)gh + \delta(1-\delta)\big(g'h - gh'\big) \;=\; (1-\delta)^{2p-2} - \delta^{2p-2} + (p-1)(1-2\delta)\big(\delta(1-\delta)\big)^{p-2},$$

which is positive for all $p \geq 2$ and $0 \leq \delta < \frac{1}{2}$.

∎

Proof of Proposition 2.6

We need to show that

$$H(y) - 1 + p\tau(x,y) - \frac{p}{2}H(x) \;=\; (p-1) + \log_2\Big((1-\delta)^p + \delta^p\Big) - \frac{p}{2}H(x) - px\log_2(1-2\delta). \quad (25)$$

We fix p and view both sides as functions of a free variable $0 \leq \delta \leq \frac{1}{2}$. Recall that $x = x(\delta) = \big(\frac{1}{2} - \delta\big) \cdot \frac{(1-\delta)^{p-1}-\delta^{p-1}}{(1-\delta)^p+\delta^p}$ and that $y = y(\delta)$ is determined by $h(p,y) = 1 - 2x$. In fact, we claim that $y(\delta) = \frac{\delta^p}{(1-\delta)^p+\delta^p}$. To see this, one has to verify $h(p,y) = 1-2x$, and this is easy to do.

We start with an auxiliary lemma.

Lemma 5.1: *With our definitions of x and y, we have that*

$$I(0,x) - I(y,x) \;=\; x\log_2(1-2\delta) - 1 + H(x) - y\log_2(\delta) - (1-y)\log_2(1-\delta).$$

Proof:

It is not hard to verify directly that we have

1. $1 - 2x = \delta(1-\delta) \cdot \frac{(1-\delta)^{p-2}+\delta^{p-2}}{(1-\delta)^p+\delta^p}$.
2. $\sqrt{(1-2x)^2 - 4y(1-y)} = \delta(1-\delta) \cdot \frac{(1-\delta)^{p-2}-\delta^{p-2}}{(1-\delta)^p+\delta^p}$.
3. $1 - 2y - \sqrt{(1-2x)^2 - 4y(1-y)} = (1-2\delta) \cdot \frac{(1-\delta)^{p-1}+\delta^{p-1}}{(1-\delta)^p+\delta^p}$.
4. $1 - 2x + \sqrt{(1-2x)^2 - 4y(1-y)} = 2\delta \cdot \frac{(1-\delta)^{p-1}}{(1-\delta)^p+\delta^p} = \frac{\delta}{1-\delta} \cdot (2-2y)$.

   This implies that $\frac{1-2x+\sqrt{(1-2x)^2-4y(1-y)}}{2-2y} = \frac{\delta}{1-\delta}$.

5. $2-2y-(1-2x)^2-(1-2x)\cdot\sqrt{(1-2x)^2-4y(1-y)}=\frac{2-4\delta}{(1-\delta)^2}\cdot(1-y)^2$.

Substituting this in the definition of I leads, after some simplification, to

$$I(0,x)-I(y,x) \;=\; \frac{1-2x}{2}\log_2\left(\frac{(1-\delta)^{p-1}-\delta^{p-1}}{(1-\delta)^{p-1}+\delta^{p-1}}\right)-y\log_2\left(\frac{\delta}{1-\delta}\right)-\frac{1}{2}\log_2\left(8x(1-x)\right)+\frac{1}{2}\log_2\left(\frac{2-4\delta}{(1-\delta)^2}\right).$$

We claim that the RHS of this expression can be further simplified to the RHS in the claim of the lemma. Indeed, simplifying, we need to verify

$$x\log_2\left(1-2\delta\right)+H(x) \;=\; \left(\frac{1}{2}-x\right)\cdot\log_2\left(\frac{(1-\delta)^{p-1}-\delta^{p-1}}{(1-\delta)^{p-1}+\delta^{p-1}}\right)-\frac{1}{2}\log_2\left(x(1-x)\right)+\frac{1}{2}\log_2\left(1-2\delta\right).$$

Expanding the enropy and rearranging, this is the same as

$$\left(\frac{1}{2}-x\right)\cdot\log_2\left((1-2\delta)\cdot\frac{(1-\delta)^{p-1}-\delta^{p-1}}{(1-\delta)^{p-1}+\delta^{p-1}}\right)+\left(x-\frac{1}{2}\right)\log_2(x)+\left(\frac{1}{2}-x\right)\log_2(1-x)=0,$$

which is equivalent to $(1-2\delta)\cdot\frac{(1-\delta)^{p-1}-\delta^{p-1}}{(1-\delta)^{p-1}+\delta^{p-1}}=\frac{x}{1-x}$, and this is easy to verify directly. ∎

We continue with the proof of the proposition. Recalling the definition of τ, and substituting the identity proved in the previous lemma in the LHS of (25), we get that it equals to

$$(p-1)+H(y)-\frac{p}{2}H(x)+p\cdot\Big(y\log_2(\delta)+(1-y)\log_2(1-\delta)-x\log_2(1-2\delta)\Big).$$

Finally, it is easy to see that $H(y)+p\cdot(y\log_2(\delta)+(1-y)\log_2(1-\delta))=\log_2\left((1-\delta)^p+\delta^p\right)$. Substituting this in the last expression gives the RHS of (25).

Proof of Proposition 2.7

We start with the first claim of the proposition. For $x=0$ we use the second definition of ψ. We have that $\delta=\frac{1}{2}$, and hence $\psi(p,0)=p-1+\log_2\left((1-\delta)^p+\delta^p\right)=0$. For $p=2$ we use the first definition of ψ. Since $h\left(2,y\right)=2\sqrt{y(1-y)}$, we have that $x=\frac{1}{2}-\sqrt{y(1-y)}$, which is the same as $y=\frac{1}{2}-\sqrt{x(1-x)}$. Hence, by the definition of τ, we have $\tau(x,y)=\frac{1+H(x)-H(y)}{2}$. Substituting this in the definition of ψ, we have $\psi(2,x)=0$.

We proceed with the second claim of the proposition, using the first definition of ψ. We view x as fixed, and write $g(p)=\psi(p,x)=H(y)-1+p\tau(x,y)-\frac{p}{2}H(x)$, where $y=y(p)$ is determined by $h(p,y)=1-2x$. Note that for $x>0$ holds $y<\frac{1}{2}-\sqrt{x(1-x)}$. And hence $\frac{\partial\tau(x,y)}{\partial y}=\log_2\left(r(x,y)\right)$. Therefore we have that

$$g' \;=\; \frac{\partial y}{\partial p}\cdot\left(\log_2\left(\frac{1-y}{y}\right)+p\frac{\partial\tau}{\partial y}\right)+\tau(x,y)-\frac{1}{2}H(x) \;=$$

$$\frac{\partial y}{\partial p}\cdot\left(\log_2\left(\frac{1-y}{y}\right)+p\log_2\left(r(x,y)\right)\right)+\tau(x,y)-\frac{1}{2}H(x).$$

Next, we claim that the expression in brackets vanishes.

Lemma 5.2: *Let $0 \le x, y \le \frac{1}{2}$ be such that $0 \le x < \frac{1}{2} - \sqrt{y(1-y)}$. Then there is a unique $p > 2$ such that $h(p, y) = 1 - 2x$ and this p is given by*

$$p \;=\; -\frac{\log_2\left(\frac{1-y}{y}\right)}{\log_2\left(r(x,y)\right)} = \frac{\log_2\left(\frac{y}{1-y}\right)}{\log_2\left(r(x,y)\right)},$$

where $r(x,y) = \frac{(1-2x)+\sqrt{(1-2x)^2-4y(1-y)}}{2-2y}$.

Proof: By the properties of the function h, there is a unique $p > 2$ such that $h(p, y) = 1 - 2x$. So it suffices verify the identity $h(p, y) = 1 - 2x$ for p given in the claim of the lemma.

Writing M for $(1-2x) + \sqrt{(1-2x)^2 - 4y(1-y)}$ and $r = \frac{M}{2-2y}$ for $r(x, y)$, we have that

$$h(p,y) \;=\; y^{\frac{1}{p}}(1-y)^{\frac{p-1}{p}} + y^{\frac{p-1}{p}}(1-y)^{\frac{1}{p}} \;=\; y^{\frac{1}{2}}(1-y)^{\frac{1}{2}} \cdot \left(\left(\frac{1-y}{y}\right)^{\frac{p-2}{2p}} + \left(\frac{y}{1-y}\right)^{\frac{p-2}{2p}}\right) \;=$$

$$y^{\frac{1}{2}}(1-y)^{\frac{1}{2}} \cdot \left(\frac{y^{\frac{1}{2}}}{(1-y)^{\frac{1}{2}}r} + \frac{(1-y)^{\frac{1}{2}}r}{y^{\frac{1}{2}}}\right) \;=\; \frac{y + (1-y)r^2}{r} \;=$$

$$\frac{M^2 + 4y(1-y)}{2M} \;=\; 1 - 2x.$$

∎

Using the lemma gives

$$g' \;=\; \tau(x,y) - \frac{1}{2}H(x).$$

We claim that this is positive for any $p > 2$ and hence g is increasing. Recall that $\tau(x, y)$ is decreasing in y. Since $0 \le y < \frac{1}{2} - \sqrt{x(1-x)}$, we have that $\tau(x,y) \ge \tau\left(x, \frac{1}{2} - \sqrt{x(1-x)}\right) = \frac{1+H(x)-H\left(\frac{1}{2}-\sqrt{x(1-x)}\right)}{2} > \frac{1}{2}H(x)$. Hence $g' > 0$.

Next, we have $g'' = \frac{\partial \tau}{\partial y} \cdot \frac{\partial y}{\partial p} > 0$, since both terms in the product are negative (τ decreases in y and h increases in p, while $y(p)$ is determined by $h(p, y) = 1 - 2x$, and x is fixed). This means that g is strongly convex, completing the proof of the second claim of the proposition.

We pass to the third claim of the proposition, using the second definition of ψ. We view p as fixed, and write $g(x) = \psi(p, x) = (p-1) + \log_2\left((1-\delta)^p + \delta^p\right) - \frac{p}{2}H(x) - px\log_2(1-2\delta)$, where δ is determined by $x = \left(\frac{1}{2} - \delta\right) \cdot \frac{(1-\delta)^{p-1} - \delta^{p-1}}{(1-\delta)^p + \delta^p}$. We have

$$g' \;=\; \frac{1}{\ln 2} \cdot \frac{1}{(1-\delta)^p + \delta^p} \cdot \left(p\delta^{p-1} - p(1-\delta)^{p-1}\right) \cdot \delta' - \frac{p}{2}\log_2\left(\frac{1-x}{x}\right) - p\log_2(1-2\delta) + px\frac{1}{\ln 2} \cdot \frac{1}{1-2\delta} \cdot 2\delta'$$

Note that the first and the fourth terms cancel out, by the definition of δ, and hence we get

$$g' \;=\; -\frac{p}{2}\log_2\left(\frac{1-x}{x}\right) - p\log_2(1-2\delta).$$

To show that g is increasing amounts to showing that $\log_2\left(\frac{1-x}{x}\right) + \log_2\left((1-2\delta)^2\right) < 0$, which, recalling $\frac{x}{1-x} = (1-2\delta)\cdot\frac{(1-\delta)^{p-1}-\delta^{p-1}}{(1-\delta)^{p-1}+\delta^{p-1}}$, is easily simplifiable to

$$\frac{(1-\delta)^{p-1}-\delta^{p-1}}{(1-\delta)^{p-1}+\delta^{p-1}} \;>\; 1-2\delta,$$

for any $p > 2$ and $0 < \delta < \frac{1}{2}$. Note that both sides of this inequality coincide for $p = 2$. We claim that the LHS increases with p. Indeed, by a simple calculation, the derivative of the LHS w.r.t. p is proportional to $\ln\left(\frac{1-\delta}{\delta}\right)\cdot\big(\delta(1-\delta)\big)^{p-1}$, which is clearly positive.

Computing the derivative of g at 0 gives, by L'Hospital,

$$g'(0) \;=\; -\frac{p}{2}\cdot\log_2\left(\lim_{x\to 0}\frac{(1-x)(1-2\delta)^2}{x}\right) \;=\; \frac{p}{2}\cdot\log_2\left(\lim_{\delta\to\frac{1}{2}}\frac{(1-2\delta)^2}{x}\right),$$

as claimed.

It is easy to see that the limit is $\frac{1}{p-1}$, and we get

$$g'(0) \;=\; \frac{p\log_2(p-1)}{2}.$$

We proceed to argue that $g'' < 0$ for $x > 0$, and hence g is strongly concave. It is easy to see that g'' is proportional to $\frac{1}{x(1-x)} + \frac{4\delta'}{1-2\delta}$. It will be convenient to state the inequality $g'' < 0$ in terms of δ. Recalling the definition of δ, and rearranging, we need to show that

$$4x(1-x) \;\geq\; -(1-2\delta)x'(\delta),$$

with equality holding only at $\delta = \frac{1}{2}$. Here we write $x(\delta) = \left(\frac{1}{2}-\delta\right)\cdot\frac{(1-\delta)^{p-1}-\delta^{p-1}}{(1-\delta)^p+\delta^p}$. Note that both sides vanish at $\delta = \frac{1}{2}$, so we need only to show strict inequality for $0 \leq \delta < \frac{1}{2}$.

We now introduce some notation to make the following calculations easier to write. Let $L_+(p) = L_+(p,\delta) = (1-\delta)^p + \delta^p$ and let $L_-(p) = L_-(p,\delta) = (1-\delta)^p - \delta^p$. Then the inequality above transforms into (after some simplification):

$$(2-4\delta)L_-(p-1)\cdot(2L_+(p) - (1-2\delta)L_-(p-1)) \;>$$
$$(p-1)(1-2\delta)^2L_+(p-2)L_+(p) + (2-4\delta)L_-(p-1)L_+(p) - p(1-2\delta)^2L_-^2(p-1).$$

Simplifying, the LHS is $(2-4\delta)L_-(2p-2)$.

We also have the following identities: $L_+(p-2)L_+(p) = L_+(2p-2) + \big((1-\delta)^2+\delta^2\big)\cdot\big((\delta(1-\delta)\big)^{p-2}$; $L_-(p-1)L_+(p) = L_-(2p-1) - (1-2\delta)\big((\delta(1-\delta)\big)^{p-1}$; $L_-^2(p-1) = L_+(2p-2) - 2\big((\delta(1-\delta)\big)^{p-1}$.

Substituting, collecting similar terms together, and simplifying, we get to

$$2L_-(2p-2) + (1-2\delta)L_+(2p-2) - 2L_-(2p-1) \;>\; (p-1)(1-2\delta)\big(\delta(1-\delta)\big)^{p-2}.$$

Expaning the "L" notation, this is the same as

$$\frac{(1-\delta)^{2p-2}-\delta^{2p-2}}{1-2\delta} \;>\; (p-1)\big(\delta(1-\delta)\big)^{p-2}. \tag{26}$$

To show this, we start with an auxiliary claim.

Lemma 5.3: *Let $p > 2$. Then the function $f(\delta) = \frac{(1-\delta)^p - \delta^p}{1-2\delta}$ decreases on $\left[0, \frac{1}{2}\right)$.*

Proof: Computing the derivative and simplifying, we have that f' is proportional to $p \cdot \left(\delta(1-\delta)^{p-1} - \delta^{p-1}(1-\delta)\right) - (p-2) \cdot ((1-\delta)^p - \delta^p)$. So, we need to show that the second term is greater than the first one. Both are equal to zero at $\delta = \frac{1}{2}$, so it suffices to show that the derivative of the second term is smaller than that of the first term, which, after simplifying, amounts to $(1-\delta)^{p-1} + \delta^{p-1} > \delta(1-\delta)^{p-2} + \delta^{p-2}(1-\delta)$, which is true for $p > 2$, and for $0 \leq \delta < \frac{1}{2}$. ∎

Now consider (26). Note that for $\delta = \frac{1}{2}$ both sides (LHS at the limit for $\delta \to \frac{1}{2}$) equal $(p-1)\left(\frac{1}{2}\right)^{2p-4}$. In addition, by the preceding lemma, the LHS decreases, while it is easy to see that the RHS increases in δ. ∎

Proof of Lemma 2.8

First,

$$\pi(x,y) \;=\; \tau(x,y) - \frac{1 + H(x) - H(y)}{2} \;=\; \tau(y,x) - \frac{1 - H(x) + H(y)}{2} \;=\; \pi(y,x).$$

The second equality is by Lemma 2.3.

Next, by the proof of Lemma 2.9, we have that for $y < \frac{1}{2} - \sqrt{x(1-x)}$ the derivative $\frac{\partial \pi(x,y)}{\partial y} = -\log_2(1-2\delta)$, where $\delta = \delta(x,y)$ is easily seen to be strictly positive if $y > 0$. Hence π is strongly increasing in y and, by symmetry, also in x.

Finally, recall that π is continuous on $\left[0, \frac{1}{2}\right] \times \left[0, \frac{1}{2}\right]$ and that $\pi(x,y) = 0$ if $y \geq \frac{1}{2} - \sqrt{x(1-x)}$. This implies that $\pi(x,y) < 0$ for $y < \frac{1}{2} - \sqrt{x(1-x)}$. ∎

Proof of Lemma 2.9

Fix σ and κ, and let $F(\delta) = \sigma H\left(\frac{x}{\sigma}\right) + (1-\sigma) H\left(\frac{x}{1-\sigma}\right) + 2x\log_2(\delta) + (1-2x)\log_2(1-\delta) - \kappa \log_2(1-2\delta)$. Then

$$F'(\delta) \;=\; \frac{\partial x}{\partial \delta} \cdot \log_2\left(\frac{\delta^2(\sigma - x)(1-\sigma-x)}{(1-\delta)^2 x^2}\right) + \frac{1}{\ln 2} \cdot \left(\frac{2x}{\delta} - \frac{1-2x}{1-\delta} + \frac{2\kappa}{1-2\delta}\right) \;=$$

$$\frac{1}{\ln 2} \cdot \left(\frac{2x}{\delta} - \frac{1-2x}{1-\delta} + \frac{2\kappa}{1-2\delta}\right) \;=\; \frac{1}{(\ln 2)\delta(1-\delta)(1-2\delta)} \cdot \Big(2x(1-2\delta) + 2\kappa\delta(1-\delta) - \delta(1-2\delta)\Big),$$

where the first equality follows since it is easy to check that the term multiplying $\frac{\partial x}{\partial \delta}$ vanishes by the definition of x.

Recalling the definition of x and simplifying, we have

$$2x(1-2\delta) + 2\kappa\delta(1-\delta) - \delta(1-2\delta) \;=\; \delta \cdot \left(\sqrt{\delta^2 + 4\sigma(1-\sigma)(1-2\delta)} - (1-2\kappa)(1-\delta)\right).$$

Now there are two cases to consider.

1. $\kappa \geq \frac{1}{2} - \sqrt{\sigma(1-\sigma)}$.

   In this case it is easy to see, by squaring both sides and analyzing the obtained quadratic inequality, that $\sqrt{\delta^2 + 4\sigma(1-\sigma)(1-2\delta)} \geq (1-2\kappa)(1-\delta)$ for all $0 \leq \delta \leq \frac{1}{2}$, and hence $F'(\delta) \geq 0$ for all δ. It follows that F is increasing and its minimum is given by $F(0) = 0$. On the other hand, by the definition of π, in this case $\pi(\sigma, \kappa) = 0$, and the claim of the lemma holds.

2. $\kappa < \frac{1}{2} - \sqrt{\sigma(1-\sigma)}$. In this case, again analyzing the appropriate quadratic inequality, it is easy to see that the minimum of F is attained at the only zero of F' on $\left[0, \frac{1}{2}\right]$, that is at

$$\delta = \delta(\sigma, \kappa) \quad = \quad \frac{(1-2\sigma)\sqrt{(1-2\sigma)^2 - 4\kappa(1-\kappa)} - \Big((1-2\sigma)^2 - 4\kappa(1-\kappa)\Big)}{4\kappa(1-\kappa)}.$$

   So, we need to verify that for this value of δ holds

$$\sigma H\left(\frac{x}{\sigma}\right) + (1-\sigma)\, H\left(\frac{x}{1-\sigma}\right) + 2x\log_2(\delta) + (1-2x)\log_2(1-\delta) - \kappa\log_2(1-2\delta) = 2\pi\,(\sigma, \kappa)\,.$$

   Let $L(\sigma, \kappa)$ denote the LHS of the above. We need to verify $L = 2\pi$. First, note that for $\kappa = \frac{1}{2} - \sqrt{\delta(1-\delta)}$, we have $\delta(\sigma, \kappa) = 0$ and both L and π vanish. So we have an equality at an endpoint, and hence it suffices to show that the derivatives $\frac{\partial L}{\partial \kappa}$ and $\frac{\partial (2\pi)}{\partial \kappa}$ coincide. We have

$$\frac{\partial L}{\partial \kappa} = \frac{\partial x}{\partial \kappa} \cdot \log_2\left(\frac{\delta^2(\sigma - x)(1-\sigma-x)}{(1-\delta)^2 x^2}\right) + \frac{\partial \delta}{\partial \kappa} \cdot \frac{1}{\ln 2} \cdot \left(\frac{2x}{\delta} - \frac{1-2x}{1-\delta} + \frac{2\kappa}{1-2\delta}\right) - \log_2(1-2\delta) =$$

$$-\log_2(1-2\delta).$$

   To see the equality, note that the first summand vanishes by the definition of x, and the second summand vanishes by the definition of δ.

   On the other hand, by the definition of π, we have $2\pi(\sigma, \kappa) = H(\sigma) + H(\kappa) - 1 + 2I(\kappa, \sigma) - 2I(0, \sigma)$. Hence, using the notation of Section 2.1.1,

$$\frac{\partial (2\pi)}{\partial \kappa} \quad = \quad \log_2\left(\frac{1-\kappa}{\kappa}\right) + 2\log_2\left(r(\sigma, \kappa)\right) \quad =$$

$$\log_2\left(\frac{1-\kappa}{\kappa}\right) + 2\log_2\left(\frac{(1-2\sigma) + \sqrt{(1-2\sigma)^2 - 4\kappa(1-\kappa)}}{2-2\kappa}\right) \quad =$$

$$\log_2\left(\frac{\Big((1-2\sigma) + \sqrt{(1-2\sigma)^2 - 4\kappa(1-\kappa)}\Big)^2}{4\kappa(1-\kappa)}\right).$$

   Let $C(\sigma, \kappa) = (1-2\sigma)^2 - 4\kappa(1-\kappa)$. Then the last expression can be written as $\log_2\left(\frac{\left((1-2\sigma)+\sqrt{C}\right)^2}{4\kappa(1-\kappa)}\right)$. We can also write $\delta = \frac{(1-2\sigma)\sqrt{C} - C}{4\kappa(1-\kappa)}$. It is easy to see that this implies $1 - 2\delta = \frac{\left((1-2\sigma)-\sqrt{C}\right)^2}{4\kappa(1-\kappa)}$.

It remains to observe that

$$(1-2\delta)\cdot\frac{\left((1-2\sigma)+\sqrt{C}\right)^2}{4\kappa(1-\kappa)} \;=\; \frac{\left((1-2\sigma)-\sqrt{C}\right)^2}{4\kappa(1-\kappa)}\cdot\frac{\left((1-2\sigma)+\sqrt{C}\right)^2}{4\kappa(1-\kappa)} \;=\; 1,$$

which means that

$$\frac{\partial L}{\partial\kappa}=\log_2\left(\frac{1}{1-2\delta}\right)=\log_2\left(\frac{\left((1-2\sigma)+\sqrt{C}\right)^2}{4\kappa(1-\kappa)}\right)=\frac{\partial(2\pi)}{\partial\kappa}.$$

This completes the proof of the lemma.

∎

Proof of Lemma 2.10

For fixed σ and ϵ, let $\alpha(x)=\sigma H\left(\frac{x}{\sigma}\right)+(1-\sigma)H\left(\frac{x}{1-\sigma}\right)+2x\log_2(\epsilon)+(1-2x)\log_2(1-\epsilon)$. As observed in [3], the maximum of α is attained at the only zero of α', that is at $x=x(\sigma,\epsilon)=\frac{-\epsilon^2+\epsilon\sqrt{\epsilon^2+4(1-2\epsilon)\sigma(1-\sigma)}}{2(1-2\epsilon)}$. For this value of x holds $\phi(\sigma,\epsilon)=H(\sigma)-1+\alpha(x)$.

Similarly, for fixed σ and ϵ, let $\beta(y)=y\log_2(1-2\epsilon)+H(y)+2\tau(\sigma,y)$. We are interested in the maximum of β on $0\le y\le\frac{1}{2}$. First, note that, by the definition of τ, we have $\beta(y)=\log_2(1-2\epsilon)\cdot y+1+H(\sigma)$ for $y\ge\frac{1}{2}-\sqrt{\sigma(1-\sigma)}$, and hence β decreases for $\frac{1}{2}-\sqrt{\sigma(1-\sigma)}\le y\le\frac{1}{2}$. For $0\le y<\frac{1}{2}-\sqrt{\sigma(1-\sigma)}$, we have $\beta'=\log_2(1-2\epsilon)+\log_2\left(\frac{1-y}{y}\right)+2\frac{\partial\tau}{\partial y}=(1-2\epsilon)+\log_2\left(\frac{1-y}{y}\right)+2\log_2\left(r(\sigma,y)\right)$. It is easy to see that the maximum of β is attained at the only zero of β', that is at $y=y(\sigma,\epsilon)=\frac{(1-\epsilon)-\sqrt{\epsilon^2+4(1-2\epsilon)\sigma(1-\sigma)}}{2-2\epsilon}$.

Write $\phi_2(\sigma,\epsilon)$ for $\beta(y(\sigma,\epsilon))-2$. We need to verify $\phi=\phi_2$. First, we check the boundary conditions $\phi(\sigma,0)=\phi_2(\sigma,0)$ for all $0\le\sigma\le\frac{1}{2}$. We have $x(\sigma,0)=0$ and hence $\phi(\sigma,0)=H(\sigma)-1+\alpha(0)=H(\sigma)-1$. On the other hand, $y(\sigma,0)=\frac{1}{2}-\sqrt{\sigma(1-\sigma)}$. We have $\tau\left(\sigma,\frac{1}{2}-\sqrt{\sigma(1-\sigma)}\right)=\frac{1+H(\sigma)-H\left(\frac{1}{2}-\sqrt{\sigma(1-\sigma)}\right)}{2}$, and hence $\phi_2(\sigma,0)=\beta\left(\frac{1}{2}-\sqrt{\sigma(1-\sigma)}\right)-2=H\left(\frac{1}{2}-\sqrt{\sigma(1-\sigma)}\right)+2\tau\left(\sigma,\frac{1}{2}-\sqrt{\sigma(1-\sigma)}\right)-2=H(\sigma)-1$ as well.

Next, we verify that $\frac{\partial\phi}{\partial\epsilon}=\frac{\partial\phi_2}{\partial\epsilon}$, which will complete the proof. Writing x for $x(\sigma,\epsilon)$ and y for $y(\sigma,\epsilon)$, it is easy to see that $\frac{\partial\phi}{\partial\epsilon}=\frac{\partial x}{\partial\epsilon}\cdot\left(\frac{\partial\alpha}{\partial x}_{|x(\sigma,\epsilon)}\right)+\frac{2x-\epsilon}{\ln(2)\epsilon(1-\epsilon)}=\frac{2x-\epsilon}{\ln(2)\epsilon(1-\epsilon)}$. Similarly, $\frac{\partial\phi_2}{\partial\epsilon}=\frac{\partial y}{\partial\epsilon}\cdot\left(\frac{\partial\beta}{\partial y}_{|y=y(\sigma,\epsilon)}\right)-\frac{2y}{\ln(2)(1-2\epsilon)}=-\frac{2y}{\ln(2)(1-2\epsilon)}$. So, it remains to verify $\frac{2x-\epsilon}{\epsilon(1-\epsilon)}=-\frac{2y}{1-2\epsilon}$, which is easy to do directly.

∎

Proof of Lemma 2.11

Recall that $\phi(\sigma,\epsilon) = H(\sigma) - 1 + \sigma H\left(\frac{x}{\sigma}\right) + (1-\sigma)H\left(\frac{x}{1-\sigma}\right) + 2x\log_2(\epsilon) + (1-2x)\log_2(1-\epsilon)$, where $x = x(\sigma,\epsilon) = \frac{-\epsilon^2+\epsilon\sqrt{\epsilon^2+4(1-2\epsilon)\sigma(1-\sigma)}}{2(1-2\epsilon)}$. Substituting, we need to show that

$$\min_{0<\epsilon\leq\frac{1}{2}}\left\{\sigma H\left(\frac{x}{\sigma}\right) + (1-\sigma)H\left(\frac{x}{1-\sigma}\right) + (2x-y)\log_2\left(\frac{\epsilon}{1-\epsilon}\right)\right\} = \sigma H\left(\frac{y}{2\sigma}\right) + (1-\sigma)H\left(\frac{y}{2(1-\sigma)}\right).$$

Fix σ and y, and let $F(\epsilon) = \sigma H\left(\frac{x}{\sigma}\right) + (1-\sigma)H\left(\frac{x}{1-\sigma}\right) + (2x-y)\log_2\left(\frac{\epsilon}{1-\epsilon}\right)$. Then

$$F'(\epsilon) \;=\; \frac{\partial x}{\partial\epsilon}\cdot\log_2\left(\frac{\epsilon^2(\sigma-x)(1-\sigma-x)}{(1-\epsilon)^2x^2}\right) + (2x-y)\cdot\frac{1}{\ln 2}\frac{1-2\epsilon}{\epsilon(1-\epsilon)} \;=\; (2x-y)\cdot\frac{1}{\ln 2}\frac{1-2\epsilon}{\epsilon(1-\epsilon)},$$

since the term multiplying $\frac{\partial x}{\partial\epsilon}$ vanishes by the definition of x.

It is not hard to verify that x strictly increases in ϵ from 0 to $\sigma(1-\sigma)$, and hence F has a unique minimum at ϵ for which $x = \frac{y}{2}$. Substituting this value of x in F, gives the claim of the lemma. ∎

5.0.1 Proof of Lemma 2.12

We write α for $\alpha_{\sigma,\epsilon}$. Let $x = x^*(\sigma,\epsilon) = \frac{-\epsilon^2+\epsilon\sqrt{\epsilon^2+4(1-2\epsilon)\sigma(1-\sigma)}}{2(1-2\epsilon)}$ be the point of maximum of α, that is $A = \alpha(x^*)$. We distinguish between four cases.

1. $\sigma = 0$. In this case $x = 0$ as well, and $A = \alpha(0) = B$, and the claim holds. So in the remaining cases we may and will assume $\sigma \geq \frac{1}{n}$.

2. $0 < x \leq \frac{1}{5n}$. In this case we will compare A with $B' = \alpha(0) = \log_2(1-\epsilon)$. Clearly $A \geq B \geq B'$ and hence $A - B \leq A - B'$. Recall that x satisfies: $(1-\epsilon)^2x^2 = \epsilon^2(\sigma-x)(1-\sigma-x)$. So, $\epsilon^2 = \Theta\left(\frac{x^2}{\sigma-x}\right)$, with asymptotic notation hiding absolute constants. Expanding the entropies, and using the fact that $\sigma \geq \frac{1}{n} \geq 5x$, it is easy to verify that $A - B' \;=\; O\left(\frac{1}{n}\right)$ in this case.

3. $x > \frac{1}{5n}$ and $\sigma - x < \frac{2}{n}$.
   In this case we will compare A with $B' = \alpha(\sigma) = \alpha\left(\frac{s}{n}\right)$. It is not hard to verify (expanding the entropies and rearranging) that in this case

   $$A - B' \;\leq\; x\log_2\left(\frac{\sigma}{x}\right) + (\sigma-x)\log_2\left(\frac{\sigma^2}{(1-2\sigma)x^2}\right) + O\left(\frac{1}{n}\right).$$

   Observe that $x\log_2\left(\frac{\sigma}{x}\right) = O\left(\frac{1}{n}\right)$, and that $\frac{\sigma^2}{(1-2\sigma)x^2}$ is bounded from above by an absolute constant. To see the second claim, note that $\frac{\sigma}{x}$ is bounded, and that the maximal value of x, attained for $\sigma = \frac{1}{2}$ is $\frac{\epsilon}{2}$, which is at most $\frac{1}{4}$. So σ cannot be close to $\frac{1}{2}$. Taking all of this into account, we get $A - B' \;=\; O\left(\frac{1}{n}\right)$.

4. $x > \frac{1}{5n}$ and $\sigma - x > \frac{2}{n}$.

   Recall that $\alpha'(y) = \log_2\left(\frac{\epsilon^2(\sigma-y)(1-\sigma-y)}{(1-\epsilon)^2y^2}\right)$. We will choose y to be the nearest fraction of the form $\frac{i}{n}$ approximating x from above, and set $B' = \alpha(y)$. Then $|x-y| \leq \frac{1}{n}$ and for any z between x and y holds

   $$\frac{\epsilon^2(\sigma-z)(1-\sigma-z)}{(1-\epsilon)^2z^2} \;\leq\; O\left(\frac{\epsilon^2(\sigma-x)(1-\sigma-x)}{(1-\epsilon)^2x^2}\right) \;\leq\; O(1).$$

   To see the first inequality note that all the terms in the first expression change by at most a costant factor compared to the second expression. For the second inequality, recall that $\frac{\epsilon^2(\sigma-x)(1-\sigma-x)}{(1-\epsilon)^2x^2} = 1$. Hence $\alpha'(z) \leq O(1)$ for all $y \leq z \leq x$ and hence $A - B \leq A - B' = \alpha(x) - \alpha(y) \leq O\left(\frac{1}{n}\right)$.

∎

5.0.2 Proof of Lemma 2.13

Recall that $\phi(\sigma,\epsilon) = H(\sigma) - 1 + \sigma H\left(\frac{x}{\sigma}\right) + (1-\sigma)H\left(\frac{x}{1-\sigma}\right) + 2x\log_2(\epsilon) + (1-2x)\log_2(1-\epsilon)$, where $x = x(\sigma,\epsilon) = \frac{-\epsilon^2+\epsilon\sqrt{\epsilon^2+4(1-2\epsilon)\sigma(1-\sigma)}}{2(1-2\epsilon)}$.

We have that

$$\frac{\partial\phi}{\partial\sigma} \;=\; \log_2\left(\frac{(\sigma-x)(1-\sigma-x)\epsilon^2}{(1-\epsilon)^2x}\right)\cdot\frac{\partial x^2}{\partial\sigma} + \log_2\left(\frac{1-\sigma}{\sigma}\right) + H\left(\frac{x}{\sigma}\right) - \frac{x}{\sigma}\log_2\left(\frac{\sigma-x}{x}\right) - H\left(\frac{x}{1-\sigma}\right) + \frac{x}{1-\sigma}\log_2\left(\frac{1-\sigma-x}{x}\right).$$

The first summand vanishes, since its first term vanishes by the definition of x , and it is easy to see that the rest can be simplified to $\log_2\left(\frac{1-\sigma-x}{\sigma-x}\right)$. Hence $\frac{\partial\phi}{\partial\sigma} \;=\; \log_2\left(\frac{1-\sigma-x}{\sigma-x}\right)$. Since $\tilde{\phi}(\alpha,\epsilon) = \phi\left(H^{-1}(\alpha),\epsilon\right)$, we have

$$\frac{\partial\tilde{\phi}}{\partial\alpha} \;=\; \left(H^{-1}\right)'(\alpha)\cdot\frac{\partial\phi}{\partial\sigma}_{|\sigma=H^{-1}(\alpha)} \;=\; \frac{\ln\left(\frac{1-\sigma-x}{\sigma-x}\right)}{\ln\left(\frac{1-\sigma}{\sigma}\right)},$$

where $\sigma = H^{-1}(\alpha)$. Computing the second derivative, we have that, similarly,

$$\frac{\partial^2\tilde{\phi}}{\partial\alpha^2} \;=\; \frac{1}{\log_2\left(\frac{1-\sigma}{\sigma}\right)}\cdot\frac{\partial}{\partial\sigma}_{|\sigma=H^{-1}(\alpha)}\frac{\ln(1-\sigma-x)-\ln(\sigma-x)}{\ln(1-\sigma)-\ln(\sigma)},$$

where $\sigma = H^{-1}(\alpha)$.

In order to show that $\tilde{\phi}$ is concave, we need to show that $\frac{\partial}{\partial\sigma}\frac{\ln(1-\sigma-x)-\ln(\sigma-x)}{\ln(1-\sigma)-\ln(\sigma)} \leq 0$. Computing the derivative and rearranging, it is easy to see that this is equivalent to (writing x' for $\frac{\partial x}{\partial\sigma}$):

$$\sigma(1-\sigma)\ln\left(\frac{1-\sigma}{\sigma}\right) \;\geq\; \frac{(\sigma-x)(1-\sigma-x)}{(1-2x)-(1-2\sigma)x'}\cdot\ln\left(\frac{1-\sigma-x}{\sigma-x}\right).$$

Next, note that $x' = \frac{1-2\sigma}{2\frac{(1-\epsilon)^2-\epsilon^2}{\epsilon^2}\cdot x+1}$. Substituting and using the fact that $(\sigma-x)(1-\sigma-x)\epsilon^2 = (1-\epsilon)^2x^2$, the first term on the right can be simplified to $\sigma(1-\sigma)-\frac{1}{2}x$. Observing that, for fixed value of σ, the value of $x(\sigma,\epsilon)$ increases from 0 to $\sigma(1-\sigma)$, as ϵ goes from 0 to $\frac{1}{2}$, it remains to verify that

$$\sigma(1-\sigma)\ln\left(\frac{1-\sigma}{\sigma}\right) \quad \geq \quad \left(\sigma(1-\sigma)-\frac{1}{2}x\right)\cdot\ln\left(\frac{1-\sigma-x}{\sigma-x}\right), \tag{27}$$

for any $0 \leq \sigma \leq \frac{1}{2}$ and $0 \leq x \leq \sigma(1-\sigma)$. We proceed to show this. For a fixed σ, let $f(x)$ denote the RHS of this inequality. It is easy to see that $f(0) = f(\sigma(1-\sigma)) = \sigma(1-\sigma)\ln\left(\frac{1-\sigma}{\sigma}\right)$. We claim that f is convex, which will imply (27). Indeed, direct calculation shows that

$$f'' \quad = \quad \frac{(1-2\sigma)x}{(\sigma-x)^2(1-\sigma-x)^2}\cdot\left(\frac{1}{2}-2\sigma(1-\sigma)\right) \quad \geq \quad 0.$$

This completes the proof of concavity of $\tilde{\phi}$.

Let us also observe, for application in the proof of Lemma 2.14 below, that $f'' > 0$ for all $0 < x < \sigma(1-\sigma)$, which means that the inequality in (27) is strong for all $0 < x < \sigma(1-\sigma)$. This means that $\frac{\partial^2\tilde{\phi}}{\partial\alpha^2} < 0$, for any $0 < \alpha < 1$, assuming $0 < \epsilon < \frac{1}{2}$.

Next, we compute $\tilde{\phi}$ at 1. Note that $x\left(\frac{1}{2},\epsilon\right) = \frac{\epsilon}{2}$, and hence $\tilde{\phi}(1,\epsilon) = \phi\left(\frac{1}{2},\epsilon\right) = H(\epsilon) - H(\epsilon) = 0$.

We proceed to compute the right derivative of $\tilde{\phi}$ at 0. Using the calculations above, and recalling that x is between 0 and $\sigma(1-\sigma)$, we have, by two applications of L'Hospital's rule, writing $\tilde{\phi}'(0,\epsilon)$ for the right derivative at 0:

$$\tilde{\phi}'(0,\epsilon) \quad = \quad \lim_{\sigma\to 0}\frac{\ln\left(\frac{1-\sigma-x}{\sigma-x}\right)}{\ln\left(\frac{1-\sigma}{\sigma}\right)} \quad = \quad \lim_{\sigma\to 0}\left[\sigma(1-\sigma)\cdot\frac{(1-2x)-(1-2\sigma)x'}{(\sigma-x)(1-\sigma-x)}\right] \quad =$$

$$\lim_{\sigma\to 0}\frac{\sigma(1-\sigma)}{\sigma(1-\sigma)-\frac{1}{2}x} \quad = \quad \lim_{\sigma\to 0}\frac{1}{1-\frac{1}{2}x'} \quad = \quad 2,$$

where in the last equality we have used $\lim_{\sigma\to 0} x' = \lim_{\sigma\to 0}\frac{1-2\sigma}{2\frac{(1-\epsilon)^2-\epsilon^2}{\epsilon^2}\cdot x+1} = 1$.

To compute the left derivative of $\tilde{\phi}$ at $\frac{1}{2}$, recall that $x\left(\frac{1}{2},\epsilon\right) = \frac{\epsilon}{2}$, and hence, proceeding as in the preceding calculation,

$$\tilde{\phi}'(1,\epsilon) \quad = \quad \lim_{\sigma\to\frac{1}{2}}\frac{\ln\left(\frac{1-\sigma-x}{\sigma-x}\right)}{\ln\left(\frac{1-\sigma}{\sigma}\right)} \quad = \quad \lim_{\sigma\to\frac{1}{2}}\frac{\sigma(1-\sigma)}{\sigma(1-\sigma)-\frac{1}{2}x} \quad = \quad \frac{1}{1-\epsilon}.$$

Finally, we observe that since $\tilde{\phi}$ is concave, the value of its derivative on $(0,1)$ is bounded from below by $\tilde{\phi}'(1,\epsilon) = \frac{1}{1-\epsilon}$ and hence it is strongly increasing. This completes the proof of the lemma. ∎

5.0.3 Proof of Lemma 2.14

We rely on the results in Lemma 2.13. The concavity of $\eta_p(x,\epsilon)$ in x follows immediately from that of $\tilde{\phi}$. Next, we compute the derivative of η_p w.r.t. x. Since $\tilde{\phi}(y,\epsilon)$ is concave in y, its derivative w.r.t. to y is lower-bounded by the derivative at $y=1$, that is by $\frac{1}{1-\epsilon}$. Hence

$$\frac{\partial \eta_p}{\partial x} \;=\; -\frac{p}{2p-2}\frac{\partial \phi(\tilde{y},\epsilon)}{\partial y}\left(1-\frac{p}{p-1}\cdot x,\; 2\epsilon(1-\epsilon)\right)+\frac{1}{p-1} \;\leq\; -\frac{p}{2p-2}\cdot\frac{1}{1-2\epsilon(1-\epsilon)}+\frac{1}{p-1} \;=$$

$$\frac{1}{p-1}\cdot\left(-\frac{p}{1+(1-2\epsilon)^2}+1\right) \;\leq\; 0.$$

Therefore, η_p is decreasing in x.

Let now $0<\epsilon<\frac{1}{2}$. the second derivative of $\tilde{\phi}(x,\epsilon)$ w.r.t. x is negative for all $0<x<1$, and hence the inequality in the above computation is sharp for all $x<\frac{p-1}{p}$. Hence η_p is strongly decreasing in x. Observing that $\eta_p(0,\epsilon)=\frac{1}{2}\tilde{\phi}\,(1,\; 2\epsilon(1-\epsilon))=0$, this means that η_p is strictly negative for $0<x\leq\frac{p-1}{p}$.

∎

Acknowledgement

We are grateful to Yuzhou Gu, Ilia Krasikov, Elchanan Mossel, Or Ordentlich, and Yury Polyanskiy for valuable remarks.

References

[1] J. Aaronson, *Functions with large additive energy supported on a Hamming sphere*, arXiv:1805.05295, 2018.

[2] A. Ashikhmin, A. Barg, and S. Litsyn, *Estimates of the distance distribution of codes and designs*, IEEE Trans. Inf. Theory, vol. 47, no. 2, pp. 10501061, Mar. 2001.

[3] A. Ashikhmin, G. Cohen, M. Krivelevich, and S. Litsyn *Bounds on distance distributions in codes of known size*, IEEE Trans. Inform. Theory, vol. 51, 2005, pp. 250-258.

[4] W. Beckner, *Inequalities in Fourier Analysis*, Annals of Math., 102(1975), pp. 159-182

[5] I. Benjamini, G. Kalai, and O. Schramm, *Noise sensitivity of boolean functions and applications to percolation*, Publications mathmatiques de lI.H..S., tome 90 (1999), p. 5-43.

[6] S. Bezrukov, *Isoperimetric problems in discrete spaces*, In Extremal Problems for Finite Sets, Vol. 3 of Bolyai Soc. Math. Stud. (P. Frankl, Z. Furedi, G. Katona and D. Miklos, eds), 1994, pp. 5991.

[7] A. Bogdanov and E. Mossel, *On Extracting Common Random Bits From Correlated Sources*, Trans. Inform. Theory, vol. 57, 2011, pp. 6351 - 6355.

[8] A. Bonami, *Etude des coefficients Fourier des fonctions de Lp(G)*, Annales de lInstitut Fourier, 20(2) (1970), 335-402.

[9] T. Cover and J. Thomas, **Elements of Information Theory**, Wiley 2006.

[10] P. Delsarte, *An algebraic approach to the association schemes of coding theory*, Philips Res. Rep., Suppl., vol. 10, 1973.

[11] L. Gross, *Logarithmic Sobolev inequalities*, Amer. J. of Math., 97, 1975, pp. 1061-1083.

[12] L. H. Harper, *Optimal assignments of numbers to vertices*, SIAM J. Appi Math. vol. 12, 1964, pp. 131-135.

[13] L. H. Harper, *On a problem of Kleitman and West*, Discrete Math. 93 (1991), pp. 169182.

[14] G. H. Hardy, J. E. Littlewood, and G. Polya, **Inequalities**, Cambridge University Press, 1988.

[15] S. Hart, *A note on the edges of the n-cube*, Discrete Math. vol. 14, 1976, pp. 157-163.

[16] P. Ivanisvili and T. Tkocz, *Comparison of moments of Rademacher Chaoses*, arXiv preprint arXiv:1807.04358, 2018

[17] J. Kahn, G. Kalai, and N. Linial, *The influence of variables on Boolean functions*, FOCS 1988, pp. 68-80.

[18] G. Kalai and N. Linial, *On the distance distribution of codes*, IEEE Trans. Inform. Theory 41 (1995), pp. 1467-1472.

[19] N. Kirshner and A. Samorodnitsky, *On $\ell_4 : \ell_2$ ratio of functions with restricted Fourier support*, ECCC report TR18-016.

[20] A. Khintchine: *Uber dyadische Brüche*, Math. Z. 18, 109-116 (1923)

[21] T. Kløve and V. Korzhik, **Error Detecting Codes, General Theory and Applications**, in Feedback Communication Systems. Norwell, MA: Kluwer, 1995.

[22] I. Krasikov, *Nonnegative quadratic forms and bounds on orthogonal polynomials*, Journal of Approximation Theory 111, 31-49 (2001).

[23] I. Krasikov and S. Litsyn, *Estimates for the Range of Binomiality in Codes Spectra*, IEEE Trans. Inform. Theory 43 (1997), pp. 987-990.

[24] I. Krasikov and A. Zarkh, *On zeroes of discrete orthogonal polynomials*, Journal of Approximation Theory, Volume 156, Issue 2, 2009, pp. 121-141.

[25] S. Kudekar, S. Kumar, M. Mondelli, H. D. Pfister, E. Sasoglu and R. L. Urbanke, *Reed-Muller Codes Achieve Capacity on Erasure Channels*, IEEE Trans. Inform. Theory, vol. 63, 2017, pp. 4298-4316.

[26] L. Larsson-Cohn, *Lp-norms of Hermite polynomials and an extremal problem on Wiener chaos*, Ark. Mat. 40 (2002), no. 1, 133-144.

[27] T. Y. Lee and H. T. Yau, *Logarithmic Sobolev inequality for some models of random walks*, Ann. Prob. 26, 4, 1998, pp. 1855-1873.

[28] V. I. Levenshtein, *Krawtchouk polynomials and universal bounds for codes and designs in Hamming spaces*, IEEE Trans. Inform. Theory 41 (1995), pp. 1303-1321.

[29] V. I. Levenshtein, *Universal bounds for codes and designs*, in "Handbook of Coding Theory" (V. S. Pless and W. C. Huffman, Eds.), Elsevier, Amsterdam, 1998.

[30] E. H. Lieb abd M. Loss, **Analysis**, AMS, 2001.

[31] J.H. van Lint, **Introduction to Coding Theory**, third edition, Graduate Texts in Mathematics, vol. 86, Springer-Verlag, Berlin, 1999.

[32] R. J. McEliece, E. R. Rodemich, H. Rumsey, Jr., and L. R. Welch, *New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities*, IEEE Trans. Inform. Theory 23 (1977), 157-166.

[33] E. Mossel, R. ODonnell, and K. Oleszkiewicz. *Noise stability of functions with low influences: Invariance and optimality*, Annals of Math., vol. 171 (2010), pp. 295-341.

[34] R. O'Donnel, **Analysis of Boolean functions**, Cambridge University Press, 2014.

[35] O. Ordentlich, Y. Polyanskiy, and O. Shayevitz, *A note on the probability of rectangles for correlated binary strings*, arXiv:1909.01221, 2019.

[36] Y. Polyanskiy, *Hypercontractivity of spherical averages in Hamming space*, SIAM J. Discrete Math., vol. 33, no. 2, pp. 731–754, 2019.

[37] Y. Polyanskiy, personal communication, 2016.

[38] Y. Polyanskiy, personal communication, 2019.

[39] Y. Polyanskiy, *Hypercontractivity for sparse functions on the discrete hypercube*, manuscript 2019.

[40] Y. Polyanskiy and A. Samorodnitsky, *Improved log-Sobolev inequalities, hypercontractivity and uncertainty principle on the hypercube*, Journal of Functional Analysis, to appear.

[41] A. Samorodnitsky, *On the optimum of Delsarte's linear program*, Journal of Combinatorial Theory, Series A 96, 261-287 (2001).

[42] G. Szego, **Orthogonal Polynomials**, Amer. Math. Soc., Providence, 1939.

[43] M. Talagrand, *How much are increasing sets positively correlated?*, Combinatorica 16 (1996), 243-258.