

## MIT Open Access Articles

### *Hypercontractivity of Spherical Averages in Hamming Space*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**As Published:** 10.1137/15M1046575

**Publisher:** Society for Industrial & Applied Mathematics (SIAM)

**Persistent URL:** <https://hdl.handle.net/1721.1/134821>

**Version:** Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

**Terms of Use:** Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



## HYPERCONTRACTIVITY OF SPHERICAL AVERAGES IN HAMMING SPACE\*

YURY POLYANSKIY†

**Abstract.** Consider the linear space of functions on the binary hypercube and the linear operator  $S_\delta$  acting by averaging a function over a Hamming sphere of radius  $\delta n$  around every point. It is shown that this operator has a dimension-independent bound on the norm  $L_p \rightarrow L_2$  with  $p = 1 + (1 - 2\delta)^2$ . This result evidently parallels a classical estimate of Bonami and Gross for  $L_p \rightarrow L_q$  norms for the operator of convolution with a Bernoulli noise. The estimate for  $S_\delta$  is harder to obtain since the latter is neither a part of a semigroup nor a tensor power. The result is shown by a detailed study of the eigenvalues of  $S_\delta$  and  $L_p \rightarrow L_2$  norms of the Fourier multiplier operators  $\Pi_a$  with symbol equal to a characteristic function of the Hamming sphere of radius  $a$  (in the notation common in boolean analysis  $\Pi_a f = f^{\wedge a}$ , where  $f^{\wedge a}$  is a degree- $a$  component of function  $f$ ). A sample application of the result is given: Any set  $A \subset \mathbb{F}_2^n$  with the property that  $A + A$  contains a large portion of some Hamming sphere (counted with multiplicity) must have cardinality a constant multiple of  $2^n$ .

**Key words.** Hamming space, hypercontractivity, Krawtchouk polynomials, Fourier analysis on hypercube, additive combinatorics

**AMS subject classifications.** 60C05, 60D05

**DOI.** 10.1137/15M1046575

**1. Main result and discussion.** Consider a linear space  $\mathcal{L}$  of functions on  $n$ -dimensional Hamming cube  $f : \mathbb{F}_2^n \rightarrow \mathbb{C}$ . We endow  $\mathcal{L}$  with the following norms and an inner product:

$$(1.1) \quad \|f\|_p \triangleq \mathbb{E}^{\frac{1}{p}} [|f(X)|^p], \quad 1 \leq p \leq \infty,$$

$$(1.2) \quad (f, g) \triangleq \mathbb{E} [f(X)\bar{g}(X)],$$

where  $X$  is uniform on  $\mathbb{F}_2^n$ . For any linear operator  $T : \mathcal{L} \rightarrow \mathcal{L}$  we define

$$\|T\|_{p \rightarrow q} \triangleq \sup_{f \in \mathcal{L}} \frac{\|Tf\|_q}{\|f\|_p}.$$

Let  $Z = (Z_1, \dots, Z_n)$  be a random element of  $\mathbb{F}_2^n$  with components independent and identically distributed (i.i.d.) according to Bern( $\delta$ ) distribution:  $\mathbb{P}[Z_i = 1] = 1 - \mathbb{P}[Z_i = 0] = \delta$ . For the operator

$$(1.3) \quad N_\delta f(x) \triangleq \mathbb{E} [f(x + Z)], \quad x \in \mathbb{F}_2^n, 0 \leq \delta \leq 1,$$

the so-called hypercontractive inequality was established by Bonami [3], Gross [12], and others (see [22, Chapter 9, notes] for the history):

$$(1.4) \quad \|N_\delta f\|_q \leq \|f\|_p \quad \forall q \geq p \geq 1, p - 1 \geq (q - 1)(1 - 2\delta)^2, p, q \geq 1.$$

\*Received by the editors November 2, 2015; accepted for publication (in revised form) February 22, 2019; published electronically May 2, 2019.

<http://www.siam.org/journals/sidma/33-2/M104657.html>

**Funding:** The research was supported by NSF grant CCF-12-53205 and by the NSF Center for Science of Information (CSoI) under grant agreement CCF-09-39370.

†Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA 02139 (yp@mit.edu).

There are a number of applications of hypercontractive inequalities. For example, we mention an early result in information theory [1], which has recently become known as the “small-set expansion.” A number of applications in theoretical computer science are presented in [22, Chapters 9–10]. One of the pillars of the analysis of boolean functions, the KKL lemma [14], is an ingenious application of (1.4). Hypercontractivity is also an indispensable tool in probability for analyzing mixing of Markov chains [6] and isoperimetry [21, Theorem 3.4].

In this paper we analyze the  $L_p \rightarrow L_2$  norm for an operator  $S_\delta$  of averaging over a Hamming sphere  $\mathbb{S}_{\delta n}$ . Specifically, for  $x = (x_1, \dots, x_n) \in \mathbb{F}_2^n$  denote the Hamming weight of  $x$  and the Hamming sphere centered at zero as

$$(1.5) \quad |x| \triangleq |\{j : x_j = 1\}|,$$

$$(1.6) \quad \mathbb{S}_j \triangleq \{x : |x| = j\}.$$

The operator  $S_\delta$  is defined as follows:

$$S_\delta f(x) \triangleq \binom{n}{j}^{-1} \sum_{y \in \mathbb{F}_2^n, |y|=j} f(x+y),$$

where  $j = \lceil \delta n \rceil$  if  $\delta < 1/2$  and  $j = \lfloor \delta n \rfloor$  if  $\delta \geq 1/2$ . In other words, we may write

$$S_\delta f \triangleq \frac{f * 1_{\mathbb{S}_j}}{|\mathbb{S}_j|},$$

where  $*$  denotes the convolution

$$f * g(x) \triangleq \sum_{y \in \mathbb{F}_2^n} f(x-y)g(y).$$

This definition ensures  $S_\delta f(x) = S_{1-\delta} f(\bar{x})$  for  $\delta \neq \frac{1}{2}$ , where  $\bar{x} = (1-x_1, \dots, 1-x_n)$ .

Our main result is that  $S_\delta$  satisfies an inequality entirely similar to  $N_\delta$ , namely,

$$(1.7) \quad \|S_\delta f\|_2 \leq C_\delta \|f\|_p \quad \forall p \geq 1 + (1-2\delta)^2, \delta \neq \frac{1}{2},$$

where the crucial part is that  $C_\delta > 1$  does not depend on dimension  $n$ . Note also that the constant cannot be tightened to 1. Indeed, taking  $f = 1_{\text{even}}$  to be the characteristic function of the set of all even-weight vectors we get

$$\|S_\delta\|_{p \rightarrow 2} \geq 2^{\frac{1}{2} - \frac{1}{p}}, \quad 1 \leq p \leq 2, 0 < \delta < 1,$$

regardless of dimension  $n$ . More precisely, we show the following.

**THEOREM 1.1.** *Consider the set  $F \subset [0, 1] \times [1, 2]$*

$$F = \{(\delta, p) : p \geq 1 + (1-2\delta)^2, 0 \leq \delta \leq 1, 1 < p \leq 2\}.$$

*For every compact subset  $K$  of  $F$  there exists a constant  $C = C(K)$  such that for all  $(\delta, p) \in K$ ,  $n \geq 1$ , and  $f : \mathbb{F}_2^n \rightarrow \mathbb{C}$  we have*

$$(1.8) \quad \|S_\delta f\|_2 \leq C \|f\|_p.$$

Conversely, for any  $(\delta, p) \notin F$  there is  $E > 0$  such that

$$(1.9) \quad \sup_f \frac{\|S_\delta f\|_2}{\|f\|_p} \geq e^{nE+o(n)}, \quad n \rightarrow \infty,$$

with the exception of  $\delta = 1/2, p = 1$  for which we have

$$(1.10) \quad \sup_f \frac{\|S_{1/2} f\|_2}{\|f\|_1} = 2^{n/2} \binom{n}{\lfloor n/2 \rfloor}^{-\frac{1}{2}} \sim \left(\frac{\pi n}{2}\right)^{\frac{1}{4}}.$$

*Remark.* The constants that can be extracted from our proof method (after numerical evaluations) are as follows: for  $\delta \leq 0.16$  we have  $C = \sqrt{2}$ , while for larger  $\delta$  we can take  $C$  to be arbitrarily close to  $\sqrt{2}$  for sufficiently large  $n$ .

The full proof is given in section 3, while here we give a high-level sketch. We note first that the standard methods for showing hypercontractivity do not apply since they require the operator to be a tensor power or be part of a semigroup. The semigroup could be continuous-time, as in [6], or discrete-time as in [20], but  $S_\delta$  is a member of neither. Instead, our proof proceeds by noticing that  $S_\delta$  and  $N_\delta$  commute and are self-adjoint, hence have common orthogonal eigenspaces (given by the Fourier transform, also known as degree- $d$  components). Consequently, decomposing a function  $f = \sum_j f_j$  into the sum of its projections on eigenspaces we have from (1.4)

$$(1.11) \quad \|N_\delta f\|_2^2 = \sum_j \lambda_j(N_\delta)^2 \|f_j\|_2^2 \leq \|f\|_p^2.$$

Writing a similar expansion for  $S_\delta$  we have

$$(1.12) \quad \|S_\delta f\|_2^2 = \sum_j \lambda_j(S_\delta)^2 \|f_j\|_2^2.$$

If we had that  $\lambda_j(S_\delta) \leq \lambda_j(N_\delta)$ , then we could just upper bound (1.12) with (1.11) and conclude the proof. It turns out that such estimate does hold but only for a range of  $j$ , and thus the bulk of the proof consists of showing that contribution to (1.12) of the eigenspaces outside of this range is small. This part crucially depends on a curious relation between norms of certain Fourier multiplier operators on  $\mathbb{F}_2^n$  and eigenvalues of  $S_\delta$ . The corresponding estimates that bound energies in the degree- $a$  components of functions on the hypercube are, perhaps, of independent interest.

**1.1. Discussion.** Why would one conjecture that  $S_\delta$  is hypercontractive? Note that [6, Theorem 3.7] shows that a discrete-time Markov chain on state space  $\mathcal{X}$  and whose kernel satisfies hypercontractive inequality mixes in time of order  $O(\log \log |\mathcal{X}|)$ . For  $S_\delta$ , this Markov chain is a nonstandard random walk on a hypercube  $\mathbb{F}_2^n$  which jumps by a distance exactly  $\delta n$  at each step. A simple coupling argument shows that indeed such a random walk must mix in time  $O(\log n)$ , therefore giving some probabilistic intuition as to why Theorem 1.1 might hold.

We note that our main goal was to show an  $O(1)$  estimate for  $\|S_\delta\|_{p \rightarrow q}$ . Indeed, an  $O(\sqrt{n})$  estimate is much easier.

**THEOREM 1.2.** *For any  $\delta$  and  $p \geq 1 + (q - 1)(1 - 2\delta)^2$  we have*

$$\|S_\delta\|_{p \rightarrow q} = O(\sqrt{n}).$$

*Proof.* Assuming without loss of generality that  $f \geq 0$  it is easy to see from Stirling's formula that

$$\frac{1}{\binom{n}{\delta n}} \sum_{|y|=\delta n} f(x+y) \leq O(\sqrt{n}) \sum_{|y|=\delta n} f(x+y) \delta^{|y|} (1-\delta)^{n-|y|}.$$

Then extending summation to all of  $y$  we get

$$S_\delta f(x) \leq O(\sqrt{n}) N_\delta f(x) \quad \forall x \in \mathbb{F}_2^n.$$

The result then follows from (1.4).  $\square$

The importance of having an  $O(1)$  estimate for the  $p \rightarrow q$  norm is due to the following general result of Semenov and Shneiberg [26], which generalized earlier results of Fefferman and Shapiro [8] and Segal [25]. Semenov and Shneiberg showed that if  $T$  is any operator with  $\|T\|_{p \rightarrow q} < \infty$ , then for all  $\epsilon < \epsilon_0 = \epsilon_0(p, q, \|T\|_{p \rightarrow q})$  we have

$$\|(1-\epsilon)\mathbb{E} + \epsilon T\|_{p \rightarrow q} = 1,$$

provided that  $\mathbb{E} \circ T = T \circ \mathbb{E}$ ,  $T1 = 1$  and  $(\mathbb{E}f)(x) \triangleq \mathbb{E}[f(X)]$ . The key point is that  $\epsilon_0$  only depends on  $T$  through the norm  $\|T\|_{p \rightarrow q}$ . Paired with our Theorem 1.1 this allows us to establish that certain permutation-invariant (or  $S_n$ -equivariant) operators in Hamming space have  $L_p \rightarrow L_q$  norm equal to 1.

**1.2. Application: Sumsets in Hamming space.** Our original interest in hypercontractivity was motivated by a remarkably simple solution it yields to a problem that the author attempted to solve using more conventional semidefinite programming (SDP); compare sections IV in [23] and [24]. Here is an application of the new result (Theorem 1.1), similar in spirit.

**COROLLARY 1.3.** *For every  $\epsilon \in (0, 1)$  there are constants  $C_1, C_2 > 0$  such that for any dimension  $n$  and any set  $A \subset \mathbb{F}_2^n$  we have*

$$\sup_{j \in [\epsilon n, (1-\epsilon)n]} \frac{2^n (1_A * 1_A, 1_{\mathbb{S}_j})}{|\mathbb{S}_j| |A|} \geq \lambda \quad \implies \quad |A| \geq C_1 \lambda^{C_2} 2^n.$$

*In other words,  $\mathbb{P}[X + Y \in A] \geq \lambda$  implies  $|A| \geq C_1 \lambda^{C_2} 2^n$ , where  $(X, Y)$  is uniform on  $A \times \mathbb{S}_j$ .*

*Remark.* It is known that any linear subspace  $V \subset \mathbb{F}_2^n$  which contains an  $\Omega(1)$ -fraction of any  $\mathbb{S}_{\delta n}$  must have co-dimension  $O(1)$  (in  $n \rightarrow \infty$ ). This corollary is a generalization: if a sumset  $A + A$  contains a  $\lambda$ -fraction of any Hamming sphere  $\mathbb{S}_j$  (counted with multiplicity normalized by  $|A|$ ), then the set must be of cardinality  $\Omega(2^n)$ .

*Proof.* We prove a stronger statement:

$$(1.13) \quad \left( \phi * \phi, \frac{1_{\mathbb{S}_j}}{|\mathbb{S}_j|} \right) \geq \lambda \|\phi\|_2^2 \quad \implies \quad \frac{\|\phi\|_2^2}{\|\phi\|_1^2} \leq \frac{1}{C_1} \lambda^{-C_2},$$

from which the result follows by taking  $\phi = 1_A$ . To show (1.13) denote  $\delta = \frac{j}{n}$  and consider the chain

$$(1.14) \quad \lambda \|\phi\|_2^2 \leq \left( \phi * \phi, \frac{\mathbf{1}_{\mathbb{S}_j}}{|\mathbb{S}_j|} \right)$$

$$(1.15) \quad = (\phi, S_\delta \phi)$$

$$(1.16) \quad \leq \|\phi\|_2 \|S_\delta \phi\|_2$$

$$(1.17) \quad \leq C \|\phi\|_2 \|\phi\|_p, \quad p = 1 + (1 - 2\epsilon)^2 < 2$$

$$(1.18) \quad \leq C \|\phi\|_2 \|\phi\|_1^{\frac{2}{p}-1} \|\phi\|_2^{2-\frac{2}{p}},$$

where (1.16) is Cauchy–Schwarz, (1.17) is from Theorem 1.1, and (1.18) is from log-convexity of  $\frac{1}{p} \mapsto \|\phi\|_p$ . Rearranging terms yields (1.13).  $\square$

In fact, this corollary can be interpreted in terms of the Frankl–Rödl graphs  $\text{FR}_\gamma^n$ , which are defined on the vertex set  $\mathbb{F}_2^n$  with  $v \sim v'$  if  $|v - v'| = (1 - \gamma)n$ . Denoting by  $E(A, A)$  the number of internal edges of a set  $A$ , our corollary says

$$|A| \leq \mu 2^n \implies E(A, A) \leq C'_1 \mu^{C'_2} |\mathbb{S}_{\gamma n}| |A|.$$

In the regime of constant  $\mu$  this is essentially tight. Indeed, an estimate in the opposite direction has been obtained by Benabbas, Hatami, and Magen [2] (see [16, section 5] for a public account of these results):

$$(1.19) \quad |A| \geq \mu 2^n \implies E(A, A) \geq \left( (\mu/2)^{\frac{1}{\gamma}} - o_n(1) \right) 2^n |\mathbb{S}_{\gamma n}|,$$

provided  $\gamma < 1/2$ . In particular, this implies that if  $A$  is an independent set of  $\text{FR}_\gamma^n$  (so that  $E(A, A) = 0$ ) we must have  $|A| \leq o(1)2^n$ . This is a weak form of the famous Frankl–Rödl theorem [9] showing that  $\alpha(\text{FR}_\gamma^n) \leq (2 - \epsilon(\gamma))^n$ , where  $\alpha(\cdot)$  denotes the maximal independent set of the graph. Similar to our result, (1.19) was obtained by employing a reverse hypercontractivity result of Borell [4], which states

$$(1.20) \quad \|N_\delta f\|_q \geq \|f\|_p \quad \forall -\infty < q < p < 1, p - 1 \leq (q - 1)(1 - 2\delta)^2$$

for any  $f > 0$ . Note that (1.20) cannot be extended to  $S_\delta$ , but in [2] the authors show that the eigenvalues of  $N_\delta$  and  $\frac{1}{2}(S_\delta + S_{\delta+1/n})$  are similar enough that the latter operator is almost reverse-hypercontractive. We will further discuss results of [2] below.

**1.3. Hypercontractivity and SDP.** Part of our motivation to study hypercontractivity is that it may be employed as an improvement to the method of SDP relaxation in various constraint satisfaction problems. For example, the best known bound [19] on the size of error correcting codes in Hamming space are obtained by the SDP relaxation of Delsarte [5], and there has long been interest in using hypercontractivity to improve the SDP relaxation; see [15].

The relation between hypercontractivity and SDP has also been known in the computer science literature.<sup>1</sup> For example, [11] shows that any (fixed) number of rounds of Lovász–Schrijver SDPs is unable to prove a bound better than  $\alpha(\text{FR}(m, \gamma)) < (\frac{1}{2} - \epsilon)2^m$ , whereas we know from [9] that  $\alpha(\text{FR}(m, \gamma)) < (2 - \epsilon)^m$ . At the same time, [2] shows that reverse hypercontractivity proves  $\alpha(\text{FR}(m, \gamma)) < o(2^m)$ . Following up on the latter, [16] shows that reverse hypercontractivity itself is provable in

<sup>1</sup>This paper was originally written before some of the discussed results were published. We thank the reviewers for pointing out these references.

a sum-of-squares (SOS) proof system, thereby showing that  $\alpha(\text{FR}(m, \gamma)) < o(2^m)$  is provable via Lasserre's SOS algorithm of a fixed (but dependent on  $\gamma$ ) degree.

This section gives another example where (direct, as opposed to reverse) hypercontractivity supersedes SDP methods. We mention that while the previously mentioned examples deal with integer-programming problems, our example below is inherently "continuous."

Define  $B_\delta(x) = \delta^{|x|}(1-\delta)^{n-|x|}$  to be a distribution function of an i.i.d. Bernoulli noise. For  $\lambda \in (0, 1)$  we define

$$(1.21) \quad V_n(\lambda) = \max \left\{ \frac{(\phi, \phi)}{(\phi, 1)^2} : \phi \geq 0, (\phi * \phi, B_\delta) \geq \lambda \|\phi\|_2^2 \right\}.$$

An argument entirely similar to (1.16)–(1.18) invoking Bonami–Gross (1.4) instead of Theorem 1.1 demonstrates<sup>2</sup>

$$(1.22) \quad V_n(\lambda) \leq \lambda^{-s}$$

for some  $s > 0$  and all dimensions  $n$ .

Note that the problem in (1.21) is completely " $L_2$ " and thus escaping to  $L_p$  space in order to solve it looks somewhat unusual. Indeed, a more natural approach (at least to us) would be to apply Fourier analysis or an SDP relaxation. Here is the "spectral gap" type of argument: Since the second-largest eigenvalue of  $N_\delta$  equals  $(1 - 2\delta)$  we get

$$(\phi_0, N_\delta \phi_0) \leq (1 - 2\delta) \|\phi_0\|^2,$$

where  $\phi_0 = \phi - (\phi, 1)$ . Simple manipulations then imply

$$V_n(\lambda) \leq \frac{2\delta}{\lambda - (1 - 2\delta)} \quad \text{if } \lambda > (1 - 2\delta).$$

This proves a correct estimate of  $O(1)$  but only for large values of  $\lambda$ .

An improvement of this method comes with the use of an SDP relaxation. The latter is obtained by considering  $\psi = \phi * \phi$  and retaining only the nonnegative definiteness property of  $\psi$ , i.e. we have the following upper bound:

$$V_n(\lambda) \leq \text{SDP}(n, \lambda) \triangleq \max \left\{ 2^n \frac{(\psi, B_0)}{(\psi, 1)} : \psi \succeq 0, (\psi, B_\delta) \geq \lambda(\psi, B_0) \right\},$$

where  $B_0(x) = 1\{x = 0\}$  and  $\psi \succeq 0$  denotes that  $f \mapsto f * \psi$  is a nonnegative definite operator. It can be shown that<sup>3</sup>

$$\text{SDP}(n, \lambda) = O(1), \quad \lambda > (1 - 2\delta)^2,$$

while for smaller values of  $\lambda$   $\text{SDP}(n, \lambda)$  grows polynomially in  $n$ . Thus, while SDP improves on the "spectral gap" argument, it is still unable to yield the correct estimate of  $V_n(\lambda)$  for the entire range of  $\lambda$ .

<sup>2</sup>The original question was to check whether there exists a small set  $A \subset \mathbb{F}_2^n$  such that  $\mathbb{P}[X + X' = Z] \geq \lambda \mathbb{P}[X + X' = 0]$ , where  $X \perp\!\!\!\perp X' \sim \text{uniform on } A$  and  $Z \sim \text{Bern}(\delta)$ . Bound (1.22) shows any such set occupies a nonvanishing fraction of  $\mathbb{F}_2^n$ .

<sup>3</sup>These observations were made in collaboration with Prof. A. Megretski.

**2. Auxiliary results.**

**2.1. Notation.** For  $x = (x_1, \dots, x_n) \in \mathbb{F}_2^n$  define  $\bar{x} \triangleq (1 - x_1, \dots, 1 - x_n)$ . For each  $j = 1, \dots, n$  let

$$\chi_j(x_1, \dots, x_n) \triangleq 1_{\{x_j=0\}} - 1_{\{x_j=1\}}.$$

Define the characters, indexed by  $v \in \mathbb{F}_2^n$ ,

$$\chi_v(x) \triangleq \prod_{j:v_j=1} \chi_j(x) = (-1)^{\langle v, x \rangle},$$

where  $\langle v, x \rangle = \sum_{j=1}^n v_j x_j$  is a nondegenerate bilinear form on  $\mathbb{F}_2^n$ . The Fourier transform of  $f : \mathbb{F}_2^n \rightarrow \mathbb{C}$  is

$$\hat{f}(\omega) \triangleq \sum_{x \in \mathbb{F}_2^n} \chi_\omega(x) f(x) = 2^n (f, \chi_\omega), \quad \omega \in \mathbb{F}_2^n.$$

$L_p$ -norms are monotonic,

$$(2.1) \quad \|f\|_p \leq \|f\|_{p_1}, \quad p \leq p_1,$$

and satisfy the Young inequality:

$$(2.2) \quad \|f * g\|_p \leq 2^n \|f\|_q \|g\|_r, \quad \frac{1}{p} + 1 = \frac{1}{q} + \frac{1}{r}, \quad 1 \leq p, q, r \leq \infty.$$

For the size of Hamming spheres we have

$$(2.3) \quad |\mathbb{S}_{\delta n}| = \binom{n}{\lfloor \delta n \rfloor} = e^{nh(\delta) - \frac{1}{2} \ln n + O(1)}, \quad n \rightarrow \infty,$$

where the estimate is a consequence of Stirling's formula,  $O(1)$  is uniform in  $\delta$  on compact subsets of  $(0, 1)$ , and

$$(2.4) \quad h(\delta) = -\delta \ln \delta - (1 - \delta) \ln(1 - \delta).$$

Furthermore, for all  $0 \leq j \leq n$

$$(2.5) \quad e^{nh(\frac{j}{n})} \sqrt{\frac{1}{2n}} \leq |\mathbb{S}_j| < e^{nh(\frac{j}{n})},$$

and for  $1 \leq j \leq n - 1$  (cf. [10, exercise 5.8]),

$$(2.6) \quad e^{nh(\frac{j}{n})} \sqrt{\frac{n}{8j(n-j)}} \leq |\mathbb{S}_j| \leq e^{nh(\frac{j}{n})} \sqrt{\frac{n}{2\pi j(n-j)}}.$$

**2.2. Asymptotics of Krawtchouk polynomials.** Krawtchouk polynomials are defined as Fourier transforms of Hamming spheres:

$$(2.7) \quad K_j(x) \triangleq \widehat{1_{\mathbb{S}_j}}(x) = \sum_{k=0}^n (-1)^k \binom{|x|}{k} \binom{n - |x|}{j - k}.$$

Since  $K_j(x)$  only depends on  $x$  through its Hamming weight  $|x|$ , we will abuse notation and write  $K_j(2)$  to mean value of  $K_j$  at a point with weight 2, etc.



Some useful properties of  $K_j$  are as follows (cf. [18]):

$$(2.8) \quad K_j(x) = (-1)^j K_j(n-x),$$

$$(2.9) \quad K_j(x) = (-1)^x K_{n-j}(x),$$

$$(2.10) \quad \frac{K_j(x)}{K_j(0)} = \frac{K_x(j)}{K_x(0)},$$

$$(2.11) \quad K_j(0) = \|K_j\|_2^2 = |\mathbb{S}_j| = \binom{n}{j},$$

$$(2.12) \quad K_j(x) = \sum_{|v|=j} \chi_v(x).$$

It is also well-known that  $K_j(x)$  has  $j$  simple real roots. For  $j \leq n/2$  all of them are in the following interval (see [18, equation (71)]):

$$\frac{n}{2} - \sqrt{j(n-j)} \leq x \leq \frac{n}{2} + \sqrt{j(n-j)}.$$

For large  $n$  the above bounds become tight, so that for  $j = \delta n$  the location of the first root is at roughly

$$\xi_{crit}(\delta) \triangleq \frac{1}{2} - \sqrt{\delta(1-\delta)}.$$

The following gives a convenient nonasymptotic estimate of the magnitude of  $K_j(x)$ .

LEMMA 2.1. *For all  $x, j = 0, \dots, n$  we have*

$$(2.13) \quad |K_j(x)| \leq e^{nE_j/n(x/n)},$$

where the function  $E_\delta(\xi) = E_{1-\delta}(\xi)$ , and for  $\delta \in [0, 1/2]$

$$(2.14) \quad E_\delta(\xi) = \begin{cases} \frac{1}{2}(h(\delta) + \ln 2 - h(\xi)), & \xi_{crit}(\delta) \leq \xi \leq 1 - \xi_{crit}(\delta), \\ \phi(\xi, \omega), & \xi = \frac{1}{2}(1 - (1-\delta)\omega - \delta\omega^{-1}), \end{cases}$$

where in the second case  $\omega$  ranges in

$$\omega \in \left[ -\sqrt{\frac{\delta}{1-\delta}}, -\frac{\delta}{1-\delta} \right] \cup \left[ \frac{\delta}{1-\delta}, \sqrt{\frac{\delta}{1-\delta}} \right]$$

and

$$(2.15) \quad \phi(\xi, \omega) \triangleq \xi \ln |1 - \omega| + (1 - \xi) \ln |1 + \omega| - \delta \ln |\omega|.$$

*Remark.* Exponent  $E_\xi(\delta)$  was derived in [15] for  $\xi \leq \xi_{crit}(\delta)$ . Subsequently, a refined asymptotic expansion for all  $\xi \in [0, 1]$  was found in [13]:

$$(2.16) \quad K_{\delta n}(\xi n) = \frac{O(1)}{\sqrt{n}} e^{nE_\delta(\xi)},$$

where the  $O(1)$  term is  $\theta(1)$  for  $\xi \leq \xi_{crit}$ , while for  $\xi \in [\xi_{crit}, 1/2]$  the factor  $O(1)$  is oscillating and may reduce the exponent for a few integer points  $x \in [\xi_{crit}n, (1 - \xi_{crit})n]$ , which are close to one of the roots of  $K_j(\cdot)$ .

*Proof.* Following [13]<sup>4</sup> we have

$$(2.17) \quad K_j(x) = \frac{1}{2\pi i} \oint_{\mathcal{C}} (1-z)^x (1+z)^{n-x} z^{-j} \frac{dz}{z},$$

where integration is over an arbitrary circle  $\mathcal{C}$  with center at  $z = 0$ . The derivative of the function  $(1-z)^x (1+z)^{n-x} z^{-j}$  is zero when

$$(2.18) \quad n - 2x = (n-j)z + jz^{-1}.$$

Due to (2.9) it is sufficient to consider  $j \leq n/2$ . Among the two solutions of (2.18) denote by  $\omega$  the unique one with smallest  $|z|$  and  $\Im(z) \geq 0$ . Set, for convenience,

$$\xi = x/n, \quad \delta = j/n \in [0, 1/2],$$

and note that we have the following relation between  $\omega$  and  $\xi$ :

$$(2.19) \quad \omega = \frac{1}{2(1-\delta)} \left( 1 - 2\xi - \operatorname{sgn}(1-2\xi) \cdot \sqrt{(1-2\xi)^2 - 1 + (1-2\delta)^2} \right),$$

$$(2.20) \quad 1 - 2\xi = (1-\delta)\omega + \frac{\delta}{\omega}.$$

As  $\xi$  ranges from 0 to 1 the saddle point  $\omega$  traverses the path

$$\omega : \frac{\delta}{1-\delta} \rightarrow \sqrt{\frac{\delta}{1-\delta}} \rightarrow -\sqrt{\frac{\delta}{1-\delta}} \rightarrow -\frac{\delta}{1-\delta},$$

where the middle segment is along the arc  $e^{i\phi} \sqrt{\frac{\delta}{1-\delta}}$ ,  $\phi \in [0, \pi]$ ; corresponding to these corner points  $\xi$  ranges as follows

$$\xi : 0 \rightarrow \xi_{crit} \rightarrow 1 - \xi_{crit} \rightarrow 1.$$

It is more convenient to reparameterize the answer in terms of location of the saddle point  $\omega$ . If we take  $\mathcal{C}$  to be the circle passing through  $\omega$ , then as shown in [13, (3.4) and paragraph after (3.19)] the maximum

$$\max_{z \in \mathcal{C}} |(1-z)^x (1+z)^{n-x} z^{-j}|$$

is attained at  $z = \omega$  and is equal to  $e^{nE_\delta(\xi)}$ , where

$$(2.21) \quad E_\delta(\xi) = \phi(\xi, \omega),$$

and  $\xi$  is a function of  $\omega$  defined via (2.20). Thus, upper-bounding the integrand  $\{\cdot\}$  in (2.17) by the maximal value and noting that for any circle

$$\oint_{\mathcal{C}} \left| \frac{dz}{z} \right| \leq 2\pi$$

we conclude that (2.13) holds.

It remains to show the simplified expression in (2.14) for  $\xi \in [\xi_{crit}, 1 - \xi_{crit}]$ . To that end, notice that such  $\xi$  corresponds to

$$\omega = e^{i\phi} \sqrt{\frac{\delta}{1-\delta}}, \quad \phi \in [0, \pi].$$

<sup>4</sup>Note that  $K_j(\cdot)$  in [13] corresponds to  $(-1)^j K_j(\cdot)$  in this paper.

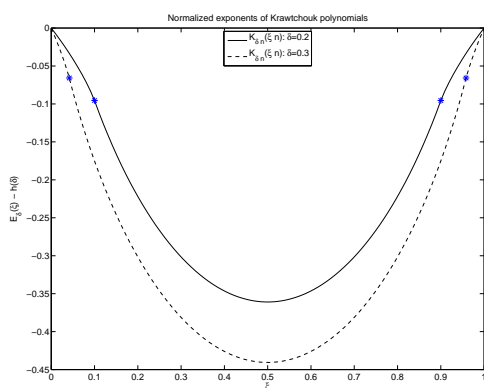


FIG. 2.1. The exponent of  $\frac{K_{\delta n}(\xi n)}{K_{\delta n}(0)}$  is equal to  $E_{\delta}(\xi) - h(\delta)$ . The figure compares these exponents for two values of  $\delta$ . Asterisks mark the interval  $[\xi_{crit}, 1 - \xi_{crit}]$  containing all the roots of  $K_{\delta n}(\cdot)$ . In this interval  $K_{\delta n}(\cdot)$  is oscillatory.

Substituting this  $\omega$  into (2.21) we see that (2.14) is equivalent to

$$(2.22) \quad \xi \ln \frac{|1 - \omega|}{\sqrt{\xi}} + (1 - \xi) \ln \frac{|1 + \omega|}{\sqrt{1 - \xi}} = \frac{1}{2} \ln \frac{2}{1 - \delta}.$$

But for  $\omega$  on the arc we have

$$\frac{|1 - \omega|}{\sqrt{\xi}} = \frac{|1 + \omega|}{\sqrt{1 - \xi}} = \sqrt{\frac{2}{1 - \delta}},$$

thus verifying (2.22) and (2.14).  $\square$

Some of the properties of  $E_{\delta}(\xi)$  are summarized below (see Figure 2.1 for an illustration):

1.  $(\delta, \xi) \mapsto E_{\delta}(\xi)$  is continuous on  $[0, 1] \times [0, 1]$  and has two symmetries:  $E_{\delta}(\xi) = E_{1-\delta}(\xi)$ ,  $E_{\delta}(\xi) = E_{\delta}(1 - \xi)$ .
2.  $E_{\delta}(0) = E_{\delta}(1) = h(\delta)$ ,  $E_{\delta}(1/2) = h(\delta)/2$ .
3.  $E_{1/2}(\xi) = \ln 2 - h(\xi)/2$ .
4.  $E_{\delta}(\xi) = h(\delta) - h(\xi) + E_{\xi}(\delta)$ .
5.  $\xi \mapsto E_{\delta}(\xi)$  is monotonically decreasing on  $[0, 1/2]$  and has continuous derivative on  $[0, 1]$ .
6.  $\delta \mapsto E_{\delta}(\xi)$  is monotonically increasing on  $[0, 1/2]$ .
7.  $\delta \mapsto E_{\delta}(\xi) - h(\delta)$  is monotonically decreasing on  $[0, 1/2]$ .
8. For fixed  $\delta$  and all  $\xi \leq \xi_{crit}(\delta)$  we have

$$(2.23) \quad E_{\delta}(\xi) \leq \xi \ln(1 - 2\delta) + h(\delta).$$

We will also need a more refined estimate for  $K_j(x)$  when  $x$  is small.

LEMMA 2.2. For  $j \leq n/2$  and  $0 \leq x \leq n\xi_{crit}(j/n) = n/2 - \sqrt{j(n-j)}$  we have

$$(2.24) \quad \frac{K_j(x)}{K_j(0)} \leq \left(1 - \frac{2j}{n}\right)^x.$$

*Remark.* With the additional factor  $O(\sqrt{n})$  the estimate (2.24) follows from (2.13). Lemma 2.2 establishes the crucial relation between spectra of operators  $N_{\delta}$  and  $S_{\delta}$  powering Theorem 1.1.

*Proof.* In the mentioned range of  $x$  the polynomial  $K_j(x)$  is monotonically decreasing since  $K_j(0) > 0$  and all roots are to the right of  $x$ . Hence, for any  $x + 1 \leq n\xi_{crit}(j/n)$  we have

$$(2.25) \quad 0 \leq \frac{K_j(x+1)}{K_j(x)} < 1.$$

On the other hand (e.g., [18, (15)]),  $K_j(\cdot)$  satisfies a three-term recurrence

$$(2.26) \quad (n-x)K_j(x+1) - (n-2j)K_j(x) + xK_j(x-1) = 0.$$

Dividing by  $nK_j(x)$  we get

$$(2.27) \quad \frac{K_j(x+1)}{K_j(x)} = \left(1 - \frac{2j}{n}\right) - \frac{x}{n} \left(\frac{K_j(x-1)}{K_j(x)} - \frac{K_j(x+1)}{K_j(x)}\right)$$

$$(2.28) \quad \leq \left(1 - \frac{2j}{n}\right),$$

where (2.28) is from (2.25). The (2.24) then follows by iterating (2.28). □

Note that for  $j \approx \frac{n}{2}$  conditions of Lemma 2.2 are not satisfied for any  $x$ . For such  $j$  we prove another (somewhat loose) estimate below.

LEMMA 2.3. *Fix arbitrary  $\theta_1 \in (0, 1/2)$ . Then for all  $x, j$  such that*

$$(2.29) \quad n - 2j \leq n\theta_1,$$

$$(2.30) \quad 0 \leq x \leq 1 + \frac{\theta_1}{1 + \theta_1^2}(n\theta_1 - (n - 2j))$$

we have

$$(2.31) \quad \left| \frac{K_j(x)}{K_j(0)} \right| \leq \theta_1^x.$$

*Proof.* Denote  $\theta = 1 - 2\frac{j}{n} \leq \theta_1$ . Clearly (2.31) holds for  $x = 0$ . From (2.27) and (2.29) it also holds for  $x = 1$ . Let the induction hypothesis be that (2.31) holds for  $x \leq x_0$ . Then

$$(2.32) \quad \left| \frac{K_j(x_0+1)}{K_j(0)} \right| = \left| \frac{n\theta}{n-x_0} \frac{K_j(x_0)}{K_j(0)} - \frac{x_0}{n-x_0} \frac{K_j(x_0-1)}{K_j(0)} \right|$$

$$(2.33) \quad \leq \frac{n\theta}{n-x_0} \theta_1^{x_0} + \frac{x_0}{n-x_0} \theta_1^{x_0-1},$$

where (2.32) is from (2.26) and (2.33) is by induction hypothesis. Finally, it is easy to see that whenever  $n - x_0 > 0$  it holds that

$$x_0 \leq n \frac{\theta_1}{1 + \theta_1^2} (\theta_1 - \theta) \iff \frac{n\theta}{n-x_0} \theta_1^{x_0} + \frac{x_0}{n-x_0} \theta_1^{x_0-1} \leq \theta_1^{x_0+1},$$

which concludes the proof of (2.31) for  $x = x_0 + 1$ . □

On the other extreme, for small values of  $j$  we can extend Lemma 2.2 to the whole range  $0 \leq x \leq \frac{n}{2}$ .

LEMMA 2.4. *There exist  $C_1 \geq 1$  and  $\delta_0 \in (0, 1)$  such that for all  $0 \leq j \leq \delta_0 n$  we have*

$$\left| \frac{K_j(x)}{K_j(0)} \right| \leq C_1 \cdot \left( 1 - \frac{2j}{n} \right)^x, \quad 0 \leq x \leq \frac{n}{2}.$$

*Remark.* In fact, one can show the statement with  $C_1 = 1$  and  $\delta_0 = 0.16$ . This is achieved by carefully following constants in the analysis and showing that  $\max_{\delta \in [0, \delta_0]}$  over the right-hand side of (2.34) is  $\leq 1$  for  $n \geq 300$ . For smaller  $n$  the statement is checkable numerically, e.g., by running the recurrence (2.26) for normalized functions  $\frac{K_j(x)}{K_j(0)(1-2\delta)^x}$  (to avoid large numbers).

*Proof.* For  $j = 0$  the inequality is trivial. For  $x \leq \xi_{crit}(j/n)$  it follows from Lemma 2.2. Thus, it is sufficient to consider  $x \geq \xi_{crit}(j/n)$ ,  $j \geq 1$ . Denote  $\delta = j/n$ . Then from Lemma 2.1 and (2.6) we have for all  $n \geq 1$

$$(2.34) \quad \left| \frac{K_j(x)}{K_j(0)(1-2\delta)^x} \right| \leq \sqrt{8(1-\delta)} \cdot e^{n(f(\delta) - \frac{1}{2}h(\delta))} \sqrt{n\delta},$$

where

$$f(\delta) = \max_{\xi \in [\xi_{crit}(\delta), 1/2]} \frac{1}{2} (\ln 2 - h(\xi)) - \xi \ln(1-2\delta).$$

From convexity of the function under maximization, we conclude

$$f(\delta) = \frac{\ln 2}{2} - \frac{1}{2} \min(h(\xi_{crit}(\delta)) + 2\xi_{crit}(\delta) \ln(1-2\delta), \ln 2(1-2\delta)).$$

Taking derivative at  $\delta = 0$  we conclude that for some  $\delta'_0 > 0$  we have

$$h(\xi_{crit}(\delta)) + 2\xi_{crit}(\delta) \ln(1-2\delta) \leq \ln 2(1-2\delta) \quad \forall \delta \in [0, \delta'_0].$$

Consequently, for such  $\delta$

$$f(\delta) = \frac{1}{2} (\ln 2 - h(\xi_{crit}(\delta))) - \xi_{crit}(\delta) \ln(1-2\delta).$$

Evidently,  $f$  is continuously differentiable and

$$f(\delta) = 2\delta + o(\delta), \quad \delta \rightarrow 0.$$

Therefore for some  $\delta_0 \in (0, \delta'_0]$  we must have

$$f(\delta) - \frac{1}{2}h(\delta) < 0 \quad \forall \delta \in (0, \delta_0].$$

The statement of the lemma then follows with  $C_1 = \max(1, \sqrt{8C'_1})$ , where  $C'_1$  is the finite supremum found in the following lemma.  $\square$

LEMMA 2.5. *Let  $\alpha, \delta_0, C > 0$  and  $f$  be a continuous function on  $[0, \delta_0]$  with  $f(0) = 0$ , derivative (one-sided at 0) bounded by  $C$  and satisfying*

$$(2.35) \quad f(\delta) - \alpha h(\delta) < 0 \quad \forall \delta \in (0, \delta_0].$$

*Then*

$$(2.36) \quad \sup_{n \geq 1} \max_{\delta \in [0, \delta_0]} e^{n(f(\delta) - \alpha h(\delta))} \sqrt{n\delta} < \infty.$$

*Proof.* Under conditions of the theorem there exists  $0 < \delta_1 < \delta_0$  such that

$$f(\delta) \leq \frac{\alpha}{2}h(\delta) \quad \forall \delta \in [0, \delta_1].$$

Thus we have

$$(2.37) \quad \max_{\delta \in [0, \delta_1]} n(f(\delta) - \alpha h(\delta)) + \frac{1}{2} \ln(\delta n) \leq \frac{1}{2} \max_{\delta \in [0, \delta_1]} -\alpha n h(\delta) + \ln \delta n$$

$$(2.38) \quad \leq \frac{1}{2} \max_{\delta \in [0, \delta_1]} \alpha n \delta \ln \delta + \ln(\delta n).$$

Without loss of generality we may assume  $\delta_1 < \frac{1}{e}$  and  $n > \frac{e^2}{\alpha}$ . In this case, maximization in (2.38) is attained at  $\delta^* \in (0, \frac{1}{n\alpha})$ . Consequently, upper-bounding the first term by zero and the second by  $\ln(\frac{1}{n\alpha} \cdot n)$  we get

$$\frac{1}{2} \max_{\delta \in [0, \delta_1]} \alpha n \delta \ln \delta + \ln(\delta n) \leq \frac{-\ln \alpha}{2}.$$

On the other hand, from (2.35) and continuity we get

$$\max_{\delta \in [\delta_1, \delta_0]} f(\delta) - \alpha h(\delta) = -C_2 < 0.$$

Therefore, putting both bounds together,

$$\max_{n \geq 1, \delta \in [0, \delta_0]} e^{n(f(\delta) - \alpha h(\delta))} \leq \max \left( \frac{1}{\sqrt{\alpha}}, \sup_n \sqrt{\delta_0 n} e^{-C_2 n} \right) < \infty. \quad \square$$

*Remark.* Reference [2] establishes the following estimate:

$$\left| \frac{1}{2} \left( \frac{K_c(n)}{K_c(0)} + \frac{K_{c-1}(n)}{K_{c-1}(0)} \right) - \left( 1 - \frac{2c}{n} \right)^n \right| \leq O \left( \max \left( n^{-\frac{1}{5}}, \frac{n}{c^2} \log^2 \frac{c^2}{n} \right) \right),$$

for all  $e^2 \sqrt{nc} \leq \frac{n}{2}$ . This result is incomparable to ours: it bounds deviation from  $(1 - \frac{2c}{n})^n$  on both sides, albeit much less precisely.

Finally, for illustrating tightness of the bounds in the next section we will need the following lemma, proved in the appendix. It is not used in the proof of Theorem 1.1.

**LEMMA 2.6.**  *$L_p$  norms of Krawtchouk polynomials are given asymptotically by the following parametric formula: Let  $\omega \in [0, 1]$ ; then for  $p \geq 2$*

$$(2.39) \quad \|K_{[\delta n]}\|_p = \exp \left\{ n \left( \frac{h(\xi) - \ln 2}{p} + \phi(\xi, \omega) \right) + O(\log n) \right\}, \quad n \rightarrow \infty,$$

$$(2.40) \quad c = \frac{(1 + \omega)^p - (1 - \omega)^p}{(1 + \omega)^p + (1 - \omega)^p},$$

$$(2.41) \quad \xi = \frac{1 - c}{2} = \frac{1}{2}(1 - (1 - \delta)\omega - \delta\omega^{-1}),$$

$$(2.42) \quad \delta = \frac{c\omega - \omega^2}{1 - \omega^2}$$

and  $\phi(\xi, \omega)$  is given by (2.15). For  $p \leq 2$  we have

$$(2.43) \quad \|K_{[\delta n]}\|_p = \exp \left\{ \frac{n}{2} h(\delta) + O(\log n) \right\}$$

as  $n \rightarrow \infty$  along a subsequence such that both  $[\delta n]$  and  $n$  are even.

**2.3. Norms of Fourier projection operators.** The Fourier projection operators  $\Pi_a$  are defined as

$$(2.44) \quad \widehat{\Pi_a f} \triangleq \hat{f} \cdot 1_{S_a} \quad a = 0, 1, \dots, n,$$

or, equivalently,

$$\Pi_a f \triangleq 2^{-n} f * K_a.$$

On the other hand from Young's inequality (2.2) we have for any convolution operator

$$\|\phi * (\cdot)\|_{1 \rightarrow 2} = 2^n \|\phi\|_2.$$

Thus we have

$$(2.45) \quad \|\Pi_a\|_{1 \rightarrow 2} = \sqrt{\binom{n}{a}}.$$

Also, we note that

$$\|\Pi_a\|_{p \rightarrow q} = \|\Pi_{n-a}\|_{p \rightarrow q},$$

and thus we only consider  $a \leq \frac{n}{2}$  below.

Estimates for other  $L_p \rightarrow L_2$  follow from the Bonami–Gross inequality (1.4) and complex interpolation.

LEMMA 2.7. *For any  $1 \leq p \leq 2$  and  $0 \leq a = n\delta \leq \frac{n}{2}$  we have*

$$(2.46) \quad \|\Pi_a\|_{p \rightarrow 2} \leq \begin{cases} (p-1)^{-\frac{a}{2}}, & p > p^*, \\ (p^*-1)^{-\frac{(1-s)a}{2}} \binom{n}{a}^{\frac{s}{p}-\frac{s}{2}}, & \frac{1}{p} = \frac{1-s}{p^*} + s, 0 \leq s \leq 1, \end{cases}$$

where  $p^* = p^*(a) = 2$  if  $\frac{h(\delta)}{\delta} \leq 2$ , and otherwise  $p^* \in (1, 2)$  is a solution of

$$p^* - \ln(p^* - 1) = \frac{h(\delta)}{\delta}.$$

We also have two weaker bounds,

$$(2.47) \quad \|\Pi_a\|_{p \rightarrow 2} \leq (p-1)^{-\frac{a}{2}},$$

$$(2.48) \quad \|\Pi_a\|_{p \rightarrow 2} \leq \binom{n}{a}^{\frac{1}{p}-\frac{1}{2}}.$$

*Remark.* The estimate (2.47) has been the basis of Kahn–Kalai–Linial results [14], so we refer to (2.47) as the KKL bound. Note that  $p^*(a) = 2$  corresponds to  $a > 0.3093n$ , and then bound (2.46) coincides with (2.48).

*Proof.* From the Riesz–Thorin interpolation [7, section VI.10.8], we know that the map  $\frac{1}{p} \mapsto \|\Pi_a\|_{p \rightarrow 2}$  is log-convex. Thus (2.46) follows from (2.47) and (2.48) by convexification (the value of  $p^*$  is chosen to minimize the resulting exponent when  $a = \delta n$ ). Thus, it is sufficient to prove (2.47) and (2.48). The second one again follows from interpolating between (2.45) and  $\|\Pi_a\|_{2 \rightarrow 2} = 1$ . For the first one notice that for any  $\tau$  we have

$$N_\tau \Pi_a = \Pi_a N_\tau = (1 - 2\tau)^a \Pi_a.$$

And thus from (1.4) with  $(1 - 2\tau)^2 = p - 1$  we get

$$\|\Pi_a f\|_2 = |1 - 2\tau|^{-a} \|\Pi_a N_\tau f\|_2 \leq |1 - 2\tau|^{-a} \|N_\tau f\|_2 \leq |1 - 2\tau|^{-a} \|f\|_p.$$

To verify the tightness of our bounds we derive a simple lower bound by considering permutation invariant functions.

LEMMA 2.8. *For any  $a \in \{0, \dots, n\}$  and any  $q, p \geq 1$  we have*

$$\|\Pi_a\|_{p \rightarrow q} \geq \frac{\|K_a\|_q \|K_a\|_{p'}}{\|K_a\|_2^2},$$

where  $p' = \frac{p}{p-1}$  is the Hölder conjugate.

*Proof.* The lower bound is shown by optimizing over a class of permutation invariant functions,

$$f(x) = K_a(x) + \sum_{j \neq a}^n c_j K_j(x) \triangleq K_a(x) + \Phi(x),$$

where  $\Phi \perp K_a$ . Note that

$$(2.49) \quad \inf_{\Phi \perp K_a} \|f\|_p = \inf_{\Phi \perp K_a} \sup_{g: \|g\|_{p'} \leq 1} (K_a + \Phi, g)$$

$$(2.50) \quad = \inf_{\Phi \perp K_a} \sup_{g\text{-sym.}: \|g\|_{p'} \leq 1} (K_a + \Phi, g)$$

$$(2.51) \quad = \sup_{g\text{-sym.}: \|g\|_{p'} \leq 1} \inf_{\Phi \perp K_a} (K_a + \Phi, g)$$

$$(2.52) \quad = \left( K_a, \frac{K_a}{\|K_a\|_{p'}} \right) = \frac{\|K_a\|_2^2}{\|K_a\|_{p'}},$$

where (2.49) is by duality  $(L_p)^* = L_{p'}$ , (2.50) states the obvious fact that supremization can be restricted to permutation-symmetric  $g$ , (2.51) is by Kneser’s minimax theorem [17] (for bi-affine function over  $X \times Y$  with  $X$  convex-compact,  $Y$  convex, and  $f$  upper semicontinuous on  $X$ ), and (2.52) is because the inner inf can only be finite if  $g$  belongs to the one-dimensional subspace spanned by  $K_a$ , i.e.,  $g = cK_a$  for a suitable  $c$ .

Since  $\Pi_a(K_a + \Phi) = K_a$  we conclude that

$$\|\Pi_a\|_{p \rightarrow q} \geq \frac{\|K_a\|_q}{\inf_{\Phi \perp K_a} \|K_a + \Phi\|_p} = \frac{\|K_a\|_q \|K_a\|_{p'}}{\|K_a\|_2^2}$$

as claimed. □

In Figure 2.2 we compare the upper and lower bounds on  $\|\Pi_a\|_{p \rightarrow 2}$  as  $a$  ranges from 0 to  $n/2$  for two values of  $p$ . We note that the KKL bound (2.47) is significantly suboptimal for small  $p$  and large  $a$ . For example, for  $a > 0.3093n$  the bound (2.48) is strictly better than KKL.

Before proceeding to the proof of the main result, we need one last estimate relating the magnitude of Krawtchouk polynomials (in the oscillating strip) to the norms of projectors  $\Pi_a$ .

LEMMA 2.9. *Fix arbitrary  $0 < \delta_0 < \Delta < 1/2$ . Then there exist constants  $C'_1, C'_2 > 0$  such that for all  $n \geq 1$ , all  $j \in [\delta_0 n, \Delta n]$ , and all*

$$\frac{n}{2} - \sqrt{j(n-j)} \leq x \leq \frac{n}{2} + \sqrt{j(n-j)}$$



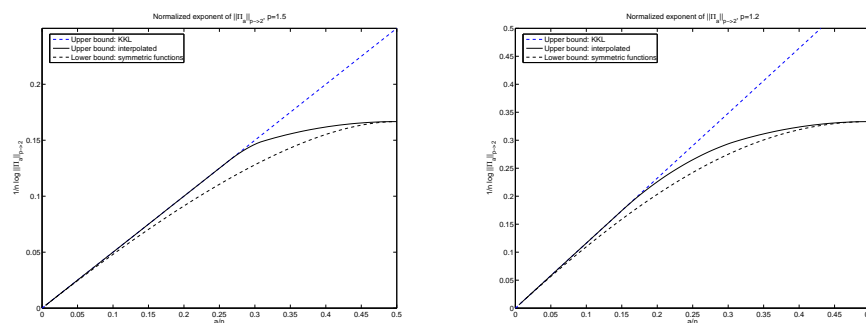


FIG. 2.2. Exponent of  $\|\Pi_a\|_{p \rightarrow 2}$  as a function of  $a$  for two values of  $p$ . Two upper bounds correspond to Kahn–Kalai–Linial (2.47) and the interpolated one (2.46). The lower bound is given by considering only permutation invariant functions (cf. Lemmas 2.6 and 2.8).

we have

$$(2.53) \quad \left| \frac{K_j(x)}{K_j(0)} \right| \cdot \|\Pi_j\|_{p(\frac{j}{n}) \rightarrow 2} \leq C'_1 \sqrt{ne}^{-C_2 n},$$

where  $p(\delta) = 1 + (1 - 2\delta)^2$ .

*Proof.* Let  $\xi = \frac{a}{n}$  and  $\delta = \frac{j}{n}$ . From symmetry, we can and will assume  $\xi \leq \frac{1}{2}$ . Since  $\xi$  is restricted to a critical strip of Krawtchouk polynomial  $K_{\delta n}$  from Lemma 2.1, bound (2.6), and Lemma 2.7 it is sufficient to show

$$(2.54) \quad \max_{\delta_0 \leq \delta \leq \Delta} \max_{\xi: (1-2\xi)^2 \leq (1-2\delta)^2 \leq 1} \frac{1}{2} (\ln 2 - h(\xi) - h(\delta)) + \pi(p(\delta), \xi) \leq -C_2 < 0,$$

where  $p(\delta) = 1 + (1 - 2\delta)^2$  and

$$\frac{1}{p} \mapsto \pi(p, \xi)$$

is the convexification of the function (cf. Lemma 2.7)

$$(2.55) \quad \frac{1}{p} \mapsto \min \left\{ -\frac{\xi}{2} \ln(p-1), \left( \frac{1}{p} - \frac{1}{2} \right) h(\xi) \right\}.$$

To show (2.54) we first change variable  $\delta$  to  $p = p(\delta) = 1 + (1 - 2\delta)^2$ . Set

$$(2.56) \quad p_0 = 1 + (1 - 2\Delta)^2,$$

$$(2.57) \quad p_1 = 1 + (1 - 2\delta_0)^2.$$

Then (2.54) is equivalent to (we also interchange the maxima in  $\xi$  and  $\delta$ )

$$(2.58) \quad \max_{\xi: (1-2\xi)^2 \leq 2-p_0} \max_{p: p_0 \leq p \leq \min(p_1, 2-(1-2\xi)^2)} \eta(\xi, p) + \frac{\ln 2 - h(\xi)}{2} \leq -C_2 < 0,$$

where

$$\eta(\xi, p) \triangleq \pi(p, \xi) - \frac{1}{2} h \left( \frac{1 - \sqrt{p-1}}{2} \right).$$

By construction,  $\frac{1}{p} \mapsto \pi(p, \xi)$  is convex. Taking derivatives one can show that  $h\left(\frac{1-\sqrt{p-1}}{2}\right)$  is concave in  $\frac{1}{p}$ . Thus, the maximization over  $p$  in (2.58) is applied to

a convex function and therefore must be achieved at one of the boundaries. Consequently, to verify (2.58) it is sufficient to show the following three strict inequalities:

$$(2.59) \quad \max_{\xi:(1-2\xi)^2 \leq 2-p_0} \eta(\xi, p_0) + \frac{\ln 2 - h(\xi)}{2} < 0,$$

$$(2.60) \quad \max_{\xi:(1-2\xi)^2 \leq 2-p_1} \eta(\xi, p_1) + \frac{\ln 2 - h(\xi)}{2} < 0,$$

$$(2.61) \quad \max_{\xi:2-p_1 \leq (1-2\xi)^2 \leq 2-p_0} \eta(\xi, 2 - (1 - 2\xi)^2) + \frac{\ln 2 - h(\xi)}{2} < 0$$

(the maximum value of the three left-hand sides is then taken to be  $-C_2$ ). The first two are verified as follows: From (2.55) we have

$$\pi(p, \xi) \leq -\frac{\xi}{2} \ln(p - 1).$$

Plugging this upper bound in (2.59) we arrive at the optimization

$$\max_{\xi:(1-2\xi)^2 \leq 2-p} -\frac{\xi}{2} \ln(p - 1) - \frac{1}{2}h(\xi).$$

Equating derivative in  $\xi$  to zero, we find solution  $\xi^*(p) = 1 - \frac{1}{p}$ . Since for  $p > 1$  we have  $(1 - 2\xi^*(p))^2 < 2 - p$ , this is also the maximizer. Consequently, substituting  $\xi = \xi^*(p)$  we get

$$\begin{aligned} \max_{(1-2\xi)^2 \leq 2-p} \eta(\xi, p) + \frac{\ln 2 - h(\xi)}{2} &\leq -\frac{\xi^*(p)}{2} \ln(p - 1) \\ &+ \frac{1}{2} \left[ \ln 2 - h(\xi^*(p)) - h\left(\frac{1 - \sqrt{p-1}}{2}\right) \right]. \end{aligned}$$

The function of a single variable  $p$  on the right is continuous and nonpositive and attains Zero only at the endpoints of  $p \in [1, 2]$ . Since both  $p_0$  and  $p_1$  belong to the interior of  $[1, 2]$ , this completes the proof of (2.59) and (2.60).

To show (2.61) we apply the bound in (2.55) (without convexification),

$$(2.62) \quad \max_{\xi} \eta(\xi, 2 - (1 - 2\xi)^2) + \frac{\ln 2 - h(\xi)}{2} \leq \max_{\xi} \frac{1}{2}f(\xi),$$

where maximization is over

$$(2.63) \quad 2 - p_1 \leq (1 - 2\xi)^2 \leq 2 - p_0$$

and  $f(\xi)$  is defined as

$$(2.64) \quad \begin{aligned} f(\xi) \triangleq \min \left\{ \left( \frac{(1 - 2\xi)^2}{2 - (1 - 2\xi)^2} \right) h(\xi), -\xi \ln(4\xi(1 - \xi)) \right\} \\ + \ln 2 - h(\xi) - h\left(\frac{1}{2} - \sqrt{\xi(1 - \xi)}\right). \end{aligned}$$

The minimum in this expression selects the first term for  $\xi \in [\xi^*, 1/2]$  and the second term otherwise, where  $\xi^* \approx 0.3082$  is the solution of

$$8\xi^2(1 - \xi) \ln \xi + (2\xi - (1 - 2\xi)^2) \ln(1 - \xi) + 2\xi(2 - (1 - 2\xi)^2) \ln 2 = 0$$

in the interior of  $(0, 1/2)$ . Furthermore, the function in (2.64) is nonpositive and continuous and attains zero only at  $\xi = 0, \frac{1}{2}$ , both of which are excluded by the constraints (2.63). Thus (2.61) holds.  $\square$

**3. Proof of Theorem 1.1.** Denote the boundary of  $F$  as

$$p(\delta) \triangleq 1 + (1 - 2\delta)^2.$$

Note that every compact subset  $K'$  of  $F$  is contained in  $F \cap \{p \geq p_0\}$  for sufficiently small  $p_0$  and in turn in some

$$(3.1) \quad K = (F \cap \{\delta : |1 - 2\delta| \geq \theta\}) \cup \{(\delta, p) : |1 - 2\delta| \leq \theta, p \geq p_0\}$$

for sufficiently small  $\theta$ . In particular, we may choose  $\theta$  so small that  $p_0 > 1 + \theta^2$ . Next note that

$$(f * 1_{S_{n-a}})(x) = (f * 1_{S_a})(\bar{x})$$

and thus estimates for  $S_\delta$  and  $S_{1-\delta}$  coincide asymptotically. Due to this symmetry and thanks to the monotonicity (2.1) of norms, to prove the theorem it is sufficient to prove the following pair of statements, corresponding to the boundary of  $K$ :

S1. (critical estimate for  $\delta < 1/2$ ) For each  $\delta$  there is  $C_\delta$  such that for all  $n \geq 1$  and all functions  $f$  we have

$$(3.2) \quad \|S_\delta f\|_2 \leq C_\delta \|f\|_{p(\delta)},$$

and function  $\delta \mapsto C_\delta$  is bounded on each  $[0, \Delta]$ ,  $\Delta < 1/2$ .

S2. (subcritical estimate around  $\delta = 1/2$ ) For any  $p > 1$  and sufficiently small  $\theta$  (in particular,  $p > 1 + \theta^2$ ) there is  $C$  such that for all  $\delta \in [(1 - \theta)/2, 1/2]$ ,  $n \geq 1$ , and functions  $f$  we have

$$(3.3) \quad \|S_\delta f\|_2 \leq C \|f\|_p.$$

First we show S1. In accordance with (2.7)

$$(3.4) \quad \|S_\delta f\|_2^2 = \sum_{a=0}^n \left| \frac{K_{\delta n}(a)}{K_{\delta n}(0)} \right|^2 \|f_a\|_2^2,$$

where we denoted

$$f_a \triangleq \Pi_a f.$$

The scheme of our proof is illustrated by Figure 3.1:

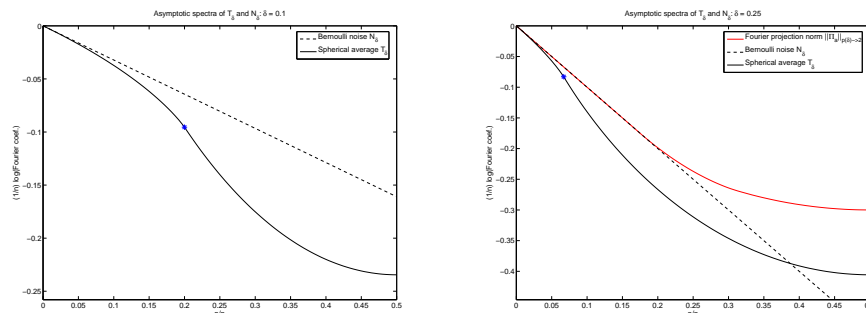


FIG. 3.1. Comparison of exponents of  $a$ th eigenvalue of  $S_\delta$  and  $N_\delta$ . For larger  $\delta$  we also show the negative of the exponent of  $\|\Pi_a\|_{p(\delta) \rightarrow 2}$ ,  $p(\delta) = 1 + (1 - 2\delta)^2$ . As before asterisks denote the critical value  $\xi_{crit}(\delta)$ , i.e., the smallest root of Krawtchouk polynomial  $K_{\delta n}(\cdot)$ .

1. First, we show that summation in (3.4) can be truncated to  $a \leq \frac{n}{2}$ .
2. Second, we show that for small values of  $\delta$  eigenvalues of  $S_\delta$  are upper-bounded by a constant multiple of eigenvalues of  $N_\delta$  defined in (1.3). This is the content of Lemma 2.4.
3. Third, for larger values of  $\delta$  we show that although eigenvalues of  $S_\delta$  can be exponentially larger than those of  $N_\delta$ , such eigenvalues correspond to large  $a$  for which  $\frac{\|f_a\|_2}{\|f\|_p}$  is exponentially smaller.

For the first step note that any  $f$  can be written as

$$f = f_{even} + f_{odd},$$

where each of the summands is supported on vectors  $x \in \mathbb{F}_2^n$  of even/odd weight. Note that  $S_\delta f_{even}$  and  $S_\delta f_{odd}$  are also of opposite parity. Thus,

$$\|S_\delta f\|_2^2 = \|S_\delta f_{even}\|_2^2 + \|S_\delta f_{odd}\|_2^2.$$

On the other hand, we have

$$(3.5) \quad (\|f_{even}\|_p^2 + \|f_{odd}\|_p^2)^{\frac{1}{2}} \leq \left\| \sqrt{f_{even}^2 + f_{odd}^2} \right\|_p$$

$$(3.6) \quad = \|f\|_p,$$

where (3.5) is from Minkowski's inequality and (3.6) is because the supports of  $f_{even}$  and  $f_{odd}$  are disjoint. Thus, if (3.2) is established for both odd and even functions, then (3.2) follows for all functions with the same constant  $C$ .

Note that for both odd and even functions we have

$$|\hat{f}(\omega)| = |\pm \hat{f}(\bar{\omega})| = |\hat{f}(\bar{\omega})|,$$

and for any such  $f$  from (3.4) and (2.8) we get

$$(3.7) \quad \|S_\delta f\|_2^2 \leq 2 \sum_{0 \leq a \leq n/2} \left| \frac{K_{\delta n}(a)}{K_{\delta n}(0)} \right|^2 \|f_a\|_2^2.$$

In the remaining we show that (3.7) is upper-bounded by  $C\|f\|_{p(\delta)}$  uniformly in  $f$  and  $\delta \leq \Delta < 1/2$ . For all  $\delta \in [0, \delta_0]$  from Lemma 2.4 we have

$$(3.8) \quad \|S_\delta f\|_2^2 \leq 2C_1^2 \sum_{0 \leq a \leq n/2} (1 - 2\delta)^{2a} \|f_a\|_2^2$$

$$(3.9) \quad = 2C_1^2 \|N_\delta f\|_2^2$$

$$(3.10) \quad \leq 2C_1^2 \|f\|_{p(\delta)}^2,$$

where the last step follows from Bonami–Gross (1.4). For  $\delta \in [\delta_0, \Delta]$  we have from Lemma 2.2

$$(3.11) \quad \left| \frac{K_{\delta n}(a)}{K_{\delta n}(0)} \right| \leq (1 - 2\delta)^a, \quad 0 \leq a \leq n\xi_{crit}(\delta).$$

On the other hand, for  $a \in [n\xi_{crit}(\delta), n/2]$  we have the estimate given by Lemma 2.9. Putting together (3.11) and (2.53) we get, similar to (3.10),

$$(3.12) \quad \|S_\delta f\|_2^2 \leq 2C_1^2 \|N_\delta f\|_2^2 + 2\|f\|_{p(\delta)}^2 \sum_{a \in [n\xi_{crit}(\delta), n/2]} (C'_1)^2 n e^{-2C_2 n}$$

$$(3.13) \quad \leq 2C_1^2 \|N_\delta f\|_2^2 + 2(C'_1)^2 \|f\|_{p(\delta)}^2 \cdot n^2 e^{-2C_2 n}$$

$$(3.14) \quad \leq 2(C_1^2 + (C'_1)^2 n^2 e^{-2C_2 n}) \|f\|_{p(\delta)}^2,$$

where in the last step we applied (1.4). Since constants  $C'_1$  and  $C_2$  only depend on  $\delta_0$  and  $\Delta$  we finish the proof of (3.2) and of statement S1.

We proceed to statement S2. Showing (3.3) is significantly simpler since  $p > p(\delta)$  this time. Take  $\theta_1 = \sqrt{p-1} > \theta$  and  $\delta_1 = \frac{1-\theta_1}{2}$ . Let

$$\xi_1 \triangleq \frac{\theta_1}{1+\theta_1^2}(\theta_1 - \theta)$$

and assume that  $\theta$  is so small that  $\xi_{crit}(\delta) < \xi_1$  for all  $\delta \in [\frac{1-\theta}{2}, \frac{1}{2}]$ . Then, on one hand, for all  $0 \leq a \leq n\xi_1$  and all  $\delta \in [\frac{1-\theta}{2}, \frac{1}{2}]$  we have from Lemma 2.3

$$\left| \frac{K_j(a)}{K_j(0)} \right| \leq (1 - 2\delta_1)^a.$$

Thus, from (1.4) we get

$$(3.15) \quad \sum_{a \in [0, n\xi_1]} \left| \frac{K_j(a)}{K_j(0)} \right|^2 \|f_a\|_2^2 \leq \|N_{\delta_1} f\|_2^2 \leq \|f\|_p^2.$$

On the other hand, for  $a > n\xi_1$  we have for some  $C_1, E > 0$

$$(3.16) \quad \left| \frac{K_j(a)}{K_j(0)} \right| \cdot \frac{\|f_a\|_2}{\|f\|_p} \leq C_1 \sqrt{n} e^{-nE} \quad \forall a \in \left[ n\xi_1, \frac{n}{2} \right].$$

Indeed, from Lemma 2.1 and (2.48) the exponent of the left-hand side of (3.16) is upper-bounded by

$$\frac{1}{2}(\ln 2 - h(\delta)) + \left( \frac{1}{p} - 1 \right) h(\xi), \quad \xi \triangleq \frac{a}{n}, \delta \triangleq \frac{j}{n},$$

since  $\xi \in (\xi_{crit}(\delta), 1/2]$ . The largest value is attained when  $\delta = \frac{1-\theta}{2}$  and  $\xi = \xi_1$ , yielding

$$\frac{1}{2}(\ln 2 - h(\delta)) + \left( \frac{1}{p} - 1 \right) h(\xi) \leq \frac{1}{2} \left( \ln 2 - h\left(\frac{1-\theta}{2}\right) \right) + \left( \frac{1}{p} - 1 \right) h\left(\frac{\theta_1(\theta_1 - \theta)}{1 + \theta_1^2}\right).$$

Since  $p > 1$  as  $\theta \rightarrow 0$  the function on the right-hand side becomes negative. Thus the exponent of left-hand side in (3.16) is negative for sufficiently small  $\theta$ .

Estimating the sum in (3.7) via (3.15) and (3.16) we get similar to (3.14) that

$$\|S_\delta f\|_2^2 \leq 2(1 + (C_1)^2 n^2 e^{-2En}) \|f\|_p^2 \quad \forall \delta \in \left[ \frac{1-\theta}{2}, \frac{1}{2} \right].$$

This completes the proof of (3.3) and statement S2.

We proceed to lower bounds on  $\|S_\delta\|_{p \rightarrow 2}$ . To show (1.9) consider the function

$$f(x) = \prod_{j=1}^n (1 + \epsilon \chi_j) = \sum_{t=0}^n (1 + \epsilon)^{n-t} (1 - \epsilon)^t 1_{\mathbb{S}_t} = \sum_{k=0}^n \epsilon^k K_k(x).$$

On one hand,

$$(3.17) \quad \|f\|_p = \left( \frac{(1 + \epsilon)^p}{2} + \frac{(1 - \epsilon)^p}{2} \right)^{\frac{n}{p}}$$

$$(3.18) \quad = e^{n \frac{p-1}{2} \epsilon^2 + o(\epsilon^2)}, \quad \epsilon \rightarrow 0.$$

On the other hand, from Lemma 2.1 and (2.16) we have

$$(3.19) \quad \|S_\delta f\|_2^2 = \sum_{a=0}^n e^{2n(E_\delta(\frac{a}{n}) - h(\delta) + \frac{a}{n} \ln \epsilon + \frac{1}{2}h(\frac{a}{n})) + o(n)},$$

where we also used

$$\|f_a\|_2 = \epsilon^a \binom{n}{a}^{\frac{1}{2}} = e^{a \ln \epsilon + nh(\frac{a}{n}) + o(n)}.$$

For convenience, set  $\xi = \frac{a}{n}$ . Then it is not hard to show from (2.14) that

$$E_\delta(\xi) - h(\delta) = \xi \ln(1 - 2\delta) + o(\xi).$$

Then setting  $\xi = \epsilon^2(1 - 2\delta)^2$  we find that

$$E_\delta(\xi) - h(\delta) + \xi \ln \epsilon + \frac{1}{2}h(\xi) = \frac{(1 - 2\delta)^2}{2} \epsilon^2 + o(\epsilon^2), \quad \epsilon \rightarrow 0.$$

Thus from (3.19) and (3.18) we get

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \ln \frac{\|S_\delta f\|_2}{\|f\|_p} \geq \frac{(1 - 2\delta)^2 - (p - 1)}{2} \epsilon^2 + o(\epsilon^2).$$

Evidently, for  $p < 1 + (1 - 2\delta)^2$  the norm  $\|S_\delta\|_{p \rightarrow 2}$  grows exponentially in dimension.

Finally, estimate (1.10) follows from Young's inequality (2.2):

$$(3.20) \quad \|S_{1/2} f\|_2 \leq 2^n \|f\|_1 \frac{\|1_{\mathbb{S}_{n/2}}\|_2}{|\mathbb{S}_{n/2}|}$$

$$(3.21) \quad = 2^n \cdot \left( 2^{-n/2} \binom{n}{\lfloor n/2 \rfloor} \right)^{-1/2} \|f\|_1$$

$$(3.22) \quad = (1 + o(1)) \left( \frac{\pi n}{2} \right)^{\frac{1}{4}} \|f\|_1.$$

This upper bound is tight as  $f(x) = 1_{\{x=0\}}$  shows.

**Appendix A. Proof of Lemma 2.6.**

*Proof.* Let  $j = \lfloor \delta n \rfloor$  and note that from Plancherel we have

$$(A.1) \quad \|K_j\|_2 = \sqrt{|\mathbb{S}_j|} = \exp \left\{ \frac{n}{2} h(\delta) + O(\log n) \right\}.$$

Consequently, we only consider  $p \neq 2$  from now on.

The lemma is shown by analyzing with exponential precision the expression

$$(A.2) \quad \|K_j\|_p^p = \sum_{a=0}^n 2^{-n} \binom{n}{a} |K_j(a)|^p,$$

so that

$$nE(p, \delta) \leq \ln \|K_j\|_p^p \leq \ln(n + 1) + nE(p, \delta),$$

where

$$(A.3) \quad E(p, \delta) \triangleq \frac{1}{n} \max_{a \leq n/2} \ln \binom{n}{a} - n \ln 2 + p \ln |K_j(a)|,$$

and we used the symmetry to restrict analysis to  $a \leq n/2$ . We will show below that for  $p > 2$  the term exponentially dominating this sum occurs at  $a \leq n\xi_{crit}(j/n)$ , while for  $p < 2$  the dominating term is at  $a = n/2$ .

First, consider  $p > 2$ . From Lemma 2.1, we have

$$(A.4) \quad E(p, \delta) \leq \max_{0 \leq \xi \leq 1/2} h(\xi) - \ln 2 + pE_\delta(\xi) + O\left(\frac{\log n}{n}\right).$$

In the regime  $\xi_{crit}(\delta) \leq \xi \leq 1/2$  we have

$$h(\xi) - \ln 2 + pE_\delta(\xi) = \frac{1}{2}(h(\delta) - \ln 2) + (1 - p/2)h(\xi),$$

which is decreasing in  $\xi$ , and hence we may restrict maximization in (A.4) to  $\xi \leq \xi_{crit}(\delta)$ . We introduce parametrization  $\xi = \xi(\omega)$  as in (2.14), with

$$\frac{\delta}{1-\delta} \leq \omega \leq \sqrt{\frac{\delta}{1-\delta}}.$$

Then using identity

$$(A.5) \quad \frac{d}{d\omega} \phi(\xi(\omega), \omega) = \xi'(\omega) \ln \frac{1-\omega}{1+\omega}$$

we get that the derivative of the expression under the max in (A.4) is

$$(A.6) \quad \frac{d}{d\omega}(\dots) = \xi'(\omega) \left( \ln \frac{1-\xi}{\xi} + p \ln \frac{1-\omega}{1+\omega} \right). \quad \square$$

It is clear that this function is strictly increasing as  $\omega$  ranges in (A.5). For the right endpoint in (A.5) we have  $\xi = 0$  and thus the derivative tends to  $-\infty$ ; for the left endpoint, notice that when  $p = 2$  and  $\omega = \sqrt{\frac{\delta}{1-\delta}}$  the expression (A.6) is exactly zero and thus  $> 0$  for  $p > 2$ . So there does exist a unique  $\omega^*(p, \delta)$  such that (A.6) equals zero. Instead of finding the function  $\omega^*(p, \delta)$  and  $\xi^* = \xi(\omega^*)$  we fix an arbitrary value  $\omega \in [0, 1]$  and find the  $\delta$  for which  $\omega^*(p, \delta) = \omega$ . This gives expression for  $\delta = \delta(\omega)$  given in (2.42). Plugging the values  $\delta = \delta(\omega)$  and  $\xi^* = \xi^*(\delta(\omega), \omega)$  into (A.4) we conclude that

$$E(p, \delta) \leq h(\xi^*) - \ln 2 + E_\delta(\xi^*) + O\left(\frac{\log n}{n}\right),$$

where furthermore  $E_\delta(\xi^*) = \phi(\xi, \omega)$ . This completes proof of the upper bound in (2.39).

To prove a matching lower bound, notice that for any fixed  $\delta$  we have argued that  $\omega = \sqrt{\frac{\delta}{1-\delta}}$  yields a positive value of (A.6). Consequently, the optimal value of  $\xi^*$  in (A.4) is always  $< \xi_{crit}(\delta) - \epsilon$  for some  $\epsilon = \epsilon(p, \delta) > 0$ . Thus, taking  $a = \lfloor \xi^* n \rfloor$ , we can apply the result of [15, section IV] establishing

$$K_j(a) = \exp\{nE_\delta(\xi^*) + O(\log n)\},$$

which shows that  $E(p, \delta) \geq h(\xi^*) - \ln 2 + pE_\delta(\xi^*) + O(\frac{\log n}{n})$  matching the previous upper bound.

We now prove (2.43). The upper bound follows from  $\|K_j\|_p \leq \|K_j\|_2$  and (A.1). For the lower bound, assume  $j$  and  $n$  are even. From (2.8) we have  $K_k(n/2) = 0$  for

any odd  $k$ , and thus from (2.10), we have that roots of  $K_{n/2}(\cdot)$  are precisely all odd integers in  $[n]$ , so that

$$K_{n/2}(x) = c \prod_{m=1}^{n/2} (x - 2m - 1),$$

where constant  $c$  is found from  $K_{n/2}(0) = \binom{n}{n/2}$ . Applying (2.10) again, we find

$$K_j(n/2) = \frac{\binom{n}{j}}{\binom{n}{n/2}} K_{n/2}(j).$$

When  $j$  is even,  $K_j(n/2)$  is nonzero, so analyzing this similarly to the proof of the Stirling formula we get

$$K_j(n/2) = \exp\{nh(\delta)/2 + O(\log n)\}.$$

The lower bound in (2.43) then follows from (cf. (A.2))

$$\|K_j\|_p^p \geq 2^{-n} \binom{n}{n/2} |K_j(n/2)|^p.$$

**Acknowledgment.** We are grateful to Prof. Y. Peres for a stimulating discussion.

#### REFERENCES

- [1] R. AHLWEDE AND P. GACS, *Spreading of sets in product spaces and hypercontraction of the Markov operator*, Ann. Probab., 4 (1976), pp. 925–939.
- [2] S. BENABBAS, H. HATAMI, AND A. MAGEN, *An Isoperimetric Inequality for the Hamming Cube with Applications for Integrality Gaps in Degree-Bounded Graphs*, unpublished manuscript, 2012.
- [3] A. BONAMI, *Étude des coefficients de Fourier des fonctions de  $l_p(g)$* , Ann. Inst. Fourier (Grenoble), 20 (1970), pp. 335–402.
- [4] C. BORELL, *Positivity improving operators and hypercontractivity*, Math. Z., 180 (1982), pp. 225–234.
- [5] P. DELSARTE, *An algebraic approach to the association schemes of coding theory*, Philips Research Rep. Supp., 10 (1973).
- [6] P. DIACONIS AND L. SALOFF-COSTE, *Logarithmic Sobolev inequalities for finite Markov chains*, Ann. Appl. Probab., 6 (1996), pp. 695–750.
- [7] N. DUNFORD AND J. SCHWARTZ, *Linear Operators: General Theory*, Vol. 1, Interscience, New York, 1958.
- [8] C. FEFFERMAN AND H. S. SHAPIRO, *A planar face on the unit sphere of the multiplier space  $m_p$ ,  $1 < p < \infty$* , Proc. Amer. Math. Soc., 36 (1972).
- [9] P. FRANKL AND V. RÖDL, *Forbidden intersections*, Trans. Amer. Math. Soc., 300 (1987), pp. 259–286.
- [10] R. G. GALLAGER, *Information Theory and Reliable Communication*, Wiley, New York, 1968.
- [11] K. GEORGIU, A. MAGEN, T. PITASSI, AND I. TOURLAKIS, *Integrality gaps of  $2 - o(1)$  for vertex cover SDPs in the Lovász–Schrijver hierarchy*, SIAM J. Comput., 39 (2010), pp. 3553–3570.
- [12] L. GROSS, *Logarithmic Sobolev inequalities*, Amer. J. Math., 97 (1975), pp. 1061–1083.
- [13] M. E. H. ISMAIL AND P. SIMEONOV, *Strong asymptotics for Krawtchouk polynomials*, J. Comput. Appl. Math., 100 (1998), pp. 121–144.
- [14] J. KAHN, G. KALAI, AND N. LINIAL, *The influence of variables on Boolean functions*, in Proceedings of the 29th Annual Symposium on Foundations of Computer Science, Los Alamitos, CA, 1988, pp. 68–80.
- [15] G. KALAI AND N. LINIAL, *On the distance distribution of codes*, IEEE Trans. Inform. Theory, 41 (1995), pp. 1467–1472.



- [16] M. KAUIERS, R. O'DONNELL, L.-Y. TAN, AND Y. ZHOU, *Hypercontractive inequalities via SOS, and the Frankl–Rödl graph*, in Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, 2014, pp. 1644–1658.
- [17] H. KNESER, *Sur un théorème fondamental de la théorie des jeux*, C. R. Acad. Sci. Paris, 234 (1952), pp. 2418–2420.
- [18] I. KRASIKOV AND S. LITSYN, *Survey of binary Krawtchouk polynomials*, in Codes and Association Schemes, DIMACS Ser. Discrete Math. Theoret. Comput. Sci. 56, AMS, Providence, RI, 2001, pp. 199–212.
- [19] R. MCELIECE, E. RODEMICH, H. RUMSEY, AND L. WELCH, *New upper bounds on the rate of a code via the Delsarte–MacWilliams inequalities*, IEEE Trans. Inform. Theory, 23 (1977), pp. 157–166.
- [20] L. MICLO, *Remarques sur l'hypercontractivité et l'évolution de l'entropie pour des chaînes de Markov finies*, in Séminaire de Probabilités XXXI, Springer, New York, 1997, pp. 136–167.
- [21] E. MOSSEL, R. O'DONNELL, O. REGEV, J. E. STEIF, AND B. SUDAKOV, *Non-interactive correlation distillation, inhomogeneous Markov chains, and the reverse Bonami–Beckner inequality*, Israel J. Math., 154 (2006), pp. 299–336.
- [22] R. O'DONNELL, *Analysis of Boolean Functions*, Cambridge University Press, Cambridge, UK, 2014.
- [23] Y. POLYANSKIY, *Hypothesis testing via a comparator*, in Proceedings of ISIT, Cambridge, MA, 2012.
- [24] Y. POLYANSKIY, *Hypothesis Testing via a Comparator and Hypercontractivity*, preprint, 2013.
- [25] I. SEGAL, *Construction of non-linear local quantum processes: I*, Ann. of Math., 92 (1970), pp. 462–481.
- [26] E. M. SEMENOV AND I. Y. SHNEIBERG, *Hypercontractive operators and Khinchin's inequality*, Func. Anal. Appl., 22 (1988), pp. 244–246.