Some applications of hypercontractive inequalities in quantum information theory

Ashley Montanaro[*]

November 9, 2012

Abstract

Hypercontractive inequalities have become important tools in theoretical computer science and have recently found applications in quantum computation. In this note we discuss how hypercontractive inequalities, in various settings, can be used to obtain (fairly) concise proofs of several results in quantum information theory: a recent lower bound of Lancien and Winter on the bias achievable by local measurements which are 4-designs; spectral concentration bounds for k-local Hamiltonians; and a recent result of Pellegrino and Seoane-Sepúlveda giving general lower bounds on the classical bias obtainable in multiplayer XOR games.

1 Introduction

Hypercontractive inequalities, which have been used in theoretical physics for many years, have more recently become increasingly important tools in theoretical computer science. The prototypical example of such an inequality is hypercontractivity of a certain "noise" operator on the boolean cube. For a function $f : \{0,1\}^n \to \mathbb{R}$, and $-1 \le \epsilon \le 1$, define the noise operator T_ϵ as follows:

$$(T_\epsilon f)(x) = \mathbb{E}_{y \sim_\epsilon x}[f(y)], \tag{1}$$

where the expectation is over bit-strings y obtained from x by flipping each bit of x with probability $(1 - \epsilon)/2$. Thus, if $\epsilon = 1$, $T_\epsilon f = f$, whereas if $\epsilon = 0$, $T_\epsilon f$ is constant. If f has Fourier expansion $f = \sum_{S \subseteq [n]} \hat{f}(S) \chi_S$, where $\chi_S(x) = (-1)^{\sum_{i \in S} x_i}$, then

$$T_\epsilon f = \sum_{S \subseteq [n]} \epsilon^{|S|} \hat{f}(S) \chi_S.$$

We thus see that T_ϵ suppresses the higher order Fourier coefficients of f. Let $\|f\|_p$ be the normalised ℓ_p norm, $\|f\|_p = \left(\frac{1}{2^n} \sum_{x \in \{0,1\}^n} |f(x)|^p \right)^{1/p}$. It is fairly straightforward to show that, for any $p \ge 1$, $\|T_\epsilon f\|_p \le \|f\|_p$, i.e. T_ϵ is a contraction. However, a stronger result also holds.

Theorem 1 (Bonami [14], Gross [31]). *For any $f : \{0,1\}^n \to \mathbb{R}$, and any p and q such that $1 \le p \le q \le \infty$ and $\epsilon \le \sqrt{\frac{p-1}{q-1}}$,*

$$\|T_\epsilon f\|_q \le \|f\|_p.$$

---

[*]Centre for Quantum Information and Quantum Foundations, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, UK; `am994@cam.ac.uk`.

Although in general $\|f\|_q \geq \|f\|_p$ for $q \geq p$, we see that for small enough ϵ, $\|T_\epsilon f\|_q \leq \|f\|_p$. T_ϵ is thus said to be *hypercontractive*. This precisely accords with the intuition that applying noise to f should make it more smooth; as p increases, $\|f\|_p$ gives more weight to extreme values of f, but applying noise smoothes out these extreme values.

Theorem 1 was proven independently by Bonami [14] and Gross [31]. In the computer science literature, Theorem 1 is often known as the Bonami-Beckner inequality (Beckner proved generalisations of this result to complex-valued functions and other settings [6]). The concept of hypercontractivity had its genesis in important work by Nelson [48] on quantum field theory (the term "hypercontractive" was coined shortly afterwards [57]), and has since found many applications in this and other areas of physics. For detailed reviews from a physics perspective of the history and more recent developments, see [23, 32].

Theorem 1 has subsequently also found a number of applications in computer science, the first of which was the celebrated result of Kahn, Kalai and Linial that every boolean function has an influential variable [38]. Following this, many important results in the theory of boolean functions have also made crucial use of hypercontractivity of T_ϵ (e.g. [15, 40, 47]; see [59, 51] for reviews). These results often use Theorem 1 through the following corollary, which relates different norms of low-degree polynomials. For the (simple) proof of this corollary from Theorem 1, see e.g. [50].

Corollary 2. *Let $f : \{\pm 1\}^n \to \mathbb{R}$ be a degree d polynomial. Then, for any $1 \leq p \leq 2$, $\|f\|_p \geq (p-1)^{d/2}\|f\|_2$, and for any $q \geq 2$, $\|f\|_q \leq (q-1)^{d/2}\|f\|_2$.*

Theorem 1 is just one example of a hypercontractive inequality; a more general setting in which to study hypercontractivity is as follows. Let (S, μ) be a measure space such that $\mu(S) = 1$. For any function $f : S \to \mathbb{R}$, define the L^p norm of f as $\|f\|_p = \left(\int |f(x)|^p d\mu(x)\right)^{1/p}$, for any $p \geq 1$. Note that these norms are nondecreasing with p. Let $L^p(S, \mu)$ be the space of all functions $f : S \to \mathbb{R}$ such that $\|f\|_p$ is finite. For linear operators $M : L^p(S, \mu) \to L^q(T, \nu)$, define the operator norm of M as

$$\|M\| := \sup_{f \neq 0} \frac{\|Mf\|_q}{\|f\|_p}.$$

M is said to be a contraction from L^p to L^q when $\|M\| \leq 1$. If $q > p$ and $\|M\| \leq 1$, M is said to be hypercontractive. For more on general hypercontractive inequalities, see for example [23, 6, 7, 37].

1.1 Hypercontractivity in quantum information

Applications of hypercontractivity are now being found in quantum information theory. In particular, Gavinsky et al. used Theorem 1 (via an inequality of Kahn, Kalai and Linial [38]) to give the first exponential separation between one-way quantum and classical communication complexity of partial boolean functions [30]. Later, Klartag and Regev used hypercontractivity on the n-sphere to resolve the long-standing conjecture that one-way quantum communication is exponentially stronger than even two-way classical communication [42]. Ben-Aroya, Regev and de Wolf generalised Theorem 1 to matrix-valued functions [8], and used their generalised inequality to prove limitations on quantum random access codes. Finally, the original version of Theorem 1 has been used by Buhrman et al. to find an upper bound on the classical success probability of a certain non-local game [18], and by Ambainis and de Wolf to give a general lower bound on quantum query complexity [4].

The purpose of the present work is to give several further examples of how hypercontractive inequalities can be used, from a computer science perspective, as tools in quantum information

theory. The results themselves are largely not new, and the proofs given here sometimes lead to worse constants than the original proofs. However, the use of hypercontractivity makes the proofs (arguably) more concise and intuitive, and in some cases allows immediate generalisations. Notably, each result we discuss uses hypercontractivity in a different setting (the real n-sphere, the space of Hermitian operators on n qubits, and the boolean cube).

1.1.1 Bias achievable by local measurements

Given a quantum state which is promised to be either ρ (with probability p) or σ (with probability $1 - p$), it is well known that the optimal measurement for distinguishing ρ and σ achieves success probability $\frac{1}{2}(1 + \|p\rho - (1 - p)\sigma\|_1)$, where $\|M\|_1 = \operatorname{tr}|M|$ is the usual trace norm [35, 36]. Setting $\Delta = p\rho - (1 - p)\sigma$, we thus obtain that the optimal bias (i.e. the difference between the probability of success and failure) over all measurements is just $\|\Delta\|_1$.

However, what if we are not allowed to perform an arbitrary measurement, but are restricted to performing a single fixed quantum measurement, followed by arbitrary classical postprocessing? Given a measurement $M = (M_i)$ (a partition of the identity into positive operators, i.e. a POVM), let ρ^M, σ^M be the probability distributions on measurement outcomes induced by performing M on ρ, σ. The optimal bias one can achieve by performing M is then equal to

$$\|\Delta\|_M := \|p\rho^M - (1 - p)\sigma^M\|_1 = \sum_i |p \operatorname{tr} M_i\rho - (1 - p) \operatorname{tr} M_i\sigma| = \sum_i |\operatorname{tr} M_i\Delta|.$$

Generalising work of Ambainis and Emerson [5], Matthews, Wehner and Winter studied this quantity for measurements which are 4-designs (defined below), and bipartite measurements M comprised of a product of local 4-designs, and gave essentially tight bounds on $\|\Delta\|_M$ [45]. Very recently, Lancien and Winter have extended these results to general multipartite 4-design measurements [43], achieving the following result.

Theorem 3 (Lancien and Winter [43])**.** *Let M be a k-fold tensor product 4-design and set $\Delta = p\rho - (1-p)\sigma$, for quantum states ρ, $\sigma \in B((\mathbb{C}^n)^{\otimes k})$. Then there is a universal constant $C > 1$ such that*

$$\|\Delta\|_M \geq C^{-k}\left(\sum_{S \subseteq [k]} \|\operatorname{tr}_S \Delta\|_2^2\right)^{1/2},$$

where $\|\Delta\|_2 = \left(\operatorname{tr}|\Delta|^2\right)^{1/2}$ is the Schatten 2-norm.

Lancien and Winter give two proofs of Theorem 3. While both are arguably elementary (being ultimately based on the use of Cauchy-Schwarz and some combinatorial bounds), the details of each appear somewhat intricate. We give an alternative proof of this result (with a worse constant C) based on hypercontractivity on the real n-sphere.

1.1.2 Spectral concentration for k-local Hamiltonians

A Hamiltonian (Hermitian operator) H on the space of n qubits is said to be k-local if it can be written as a sum $H = \sum_i H_i$, where each H_i acts non-trivially on at most k qubits. It is well-known that Theorem 1 can be used to prove concentration inequalities for low-degree multivariate polynomials (otherwise known as "higher order Chernoff bounds", see e.g. [27, 51]). We observe that a noncommutative generalisation of Theorem 1 allows an easy proof of the following bound on the spectra of k-local operators.

3

Theorem 4. *Let M be a k-local Hermitian operator on n qubits with eigenvalues (λ_i) such that* $\operatorname{tr} M^2 = 2^n$. *Then, for any $t \geq (2e)^{k/2}$,*

$$\frac{|\{i : |\lambda_i| \geq t\}|}{2^n} \leq \exp(-kt^{2/k}/(2e)).$$

1.1.3 General bounds on multiplayer XOR games

A simple and natural way of exploring the power of quantum correlations is via the framework of XOR games [21, 22], which have recently been intensively studied (see e.g. [53, 16, 17]). A k-player XOR game G is specified as follows. Let $A \in (\{\pm 1\}^n)^k$ be a multidimensional array (tensor). The j'th player is given an input $i_j \in \{1, \ldots, n\}$ and must reply with an output $x^j_{i_j} \in \{\pm 1\}$. The inputs are picked by a referee according to a known joint probability distribution π on $\{1, \ldots, n\}^k$. The players win if and only if the product of all their outputs is equal to the corresponding entry of A. The maximal bias achievable by deterministic strategies is therefore

$$\beta(G) := \max_{x^1, \ldots, x^k \in \{\pm 1\}^n} \left| \sum_{i_1, \ldots, i_k = 1}^n \pi_{i_1, \ldots, i_k} A_{i_1, \ldots, i_k} x^1_{i_1} \ldots x^k_{i_k} \right|.$$

One can easily see that shared randomness does not help classical players to increase the bias; however, in some cases sharing entanglement can be advantageous [21]. In the case where π is uniform, the problem of calculating or bounding $\beta(G)$ has been studied in several other settings classically, under the title of "unbalancing lights" [3] or "Gale-Berlekamp switching games" (e.g. [10]), and has applications in communication complexity [20, 28]. Even when $k = 2$, $\beta(G)$ is known to be NP-hard to compute [2, 54].

It is an interesting question to determine what the "hardest" XOR game is for k classical players, i.e. the game for which the bias of the best classical strategy is minimised. It is known that there exist games for which classical players can achieve a bias of at most $n^{-(k-1)/2}$ [28]. On the other hand, an inequality of Bohnenblust and Hille from 1931 [13] implies that any XOR game has maximal bias of at least $2^{-O(k)}n^{-(k-1)/2}$. Based on recent work by Defant, Popa and Schwarting [24], Pellegrino and Seoane-Sepúlveda have significantly improved this result by showing that the exponential dependence on k can be made polynomial.

Theorem 5 (Defant, Popa and Schwarting [24]; Pellegrino and Seoane-Sepúlveda [52])**.** *Let G be a k-player XOR game with n possible inputs per player. Then there is a universal constant $c > 0$ such that $\beta(G) = \Omega(k^{-c}n^{-(k-1)/2})$.*

One can replace a component of their proof with Corollary 2, which leads to a simple and explicit bound on $\beta(G)$ and illustrates the power of using hypercontractivity as a "black box". As we discuss below, Theorem 5 implies a very special case of a conjecture of Aaronson and Ambainis [1] that every bounded low-degree polynomial on the boolean cube has an influential variable.

In the remainder of this paper we elaborate on and prove each of these results in turn.

2 The bias of local 4-design measurements

A rank-one POVM $M = (M_i)$ in n dimensions is called a t-design [5] if

$$\sum_i p_i P_i^{\otimes t} = \int d\psi |\psi\rangle\langle\psi|^{\otimes t}, \tag{2}$$

where $p_i = \frac{1}{n} \operatorname{tr} M_i$ and $P_i = \frac{1}{\operatorname{tr} M_i} M_i$, and the integral is taken according to Haar measure on the complex unit n-sphere, normalised such that $\int d\psi |\psi\rangle\langle\psi| = I/n$. Observe that the right-hand side of (2) is equal to the projector onto the symmetric subspace of t n-dimensional systems, normalised by a factor of $\binom{n+t-1}{t}^{-1}$. A t-design is automatically an s-design for any $1 \leq s < t$ [5]. t-designs can be viewed as discrete approximations to the continuous POVM which puts uniform weight on each measurement vector $|\psi\rangle$. In particular, approximate 4-designs can be used to give an efficient derandomisation of the operation of measurement in a random basis [5], a primitive which has been used in quantum algorithms [55].

We will be interested in proving bounds on the bias of k-partite measurement operators which are products of local 4-designs. In other words, each measurement operator $M_{i_1,\ldots,i_k}$ is a tensor product $M_{i_1}^1 \otimes M_{i_2}^2 \otimes \cdots \otimes M_{i_k}^k$, where each individual measurement $M^i = (M_j^i)$ is a 4-design. We assume for notational simplicity the unnecessary restriction that the local dimensions are the same.

To gain intuition, we begin by considering the unipartite setting $k = 1$. In this case we will prove the following theorem, which reproduces a result of Ambainis and Emerson [5] (see also the proof of Matthews, Wehner and Winter [45]), with a worse constant.

Theorem 6. *Let M be a 4-design and set $\Delta = p\rho - (1-p)\sigma$, for quantum states ρ, $\sigma \in B(\mathbb{C}^n)$ and $0 \leq p \leq 1$. Then*

$$\|\Delta\|_M \geq \frac{1}{9(1+1/n)^{1/2}} \left((1-2p)^2 + \operatorname{tr}\Delta^2\right)^{1/2}.$$

In order to prove Theorem 6 we follow the strategy of [5, 45]. To start with, we use the definition of $\|\Delta\|_M$ and Hölder's inequality (in a form popularised by Berger as the "fourth moment method[1]" [10]) to obtain

$$
\begin{aligned}
\|\Delta\|_M &= \sum_i |\operatorname{tr} M_i \Delta| = n \sum_i p_i |\operatorname{tr} P_i \Delta| \geq n \frac{\left(\sum_i p_i (\operatorname{tr} P_i \Delta)^2\right)^{3/2}}{\left(\sum_i p_i (\operatorname{tr} P_i \Delta)^4\right)^{1/2}} \\
&= n \frac{\left(\operatorname{tr}\left(\sum_i p_i P_i^{\otimes 2}\right) \Delta^{\otimes 2}\right)^{3/2}}{\left(\operatorname{tr}\left(\sum_i p_i P_i^{\otimes 4}\right) \Delta^{\otimes 4}\right)^{1/2}}.
\end{aligned}
$$

As M is a 4-design (and hence automatically a 2-design), we can replace both the numerator and denominator with the corresponding quantities on the right-hand side of eqn. (2) to obtain

$$\|\Delta\|_M \geq n \frac{\left(\operatorname{tr}\left(\int d\psi |\psi\rangle\langle\psi|^{\otimes 2}\right) \Delta^{\otimes 2}\right)^{3/2}}{\left(\operatorname{tr}\left(\int d\psi |\psi\rangle\langle\psi|^{\otimes 4}\right) \Delta^{\otimes 4}\right)^{1/2}} = n \frac{\left(\int (\operatorname{tr}\Delta |\psi\rangle\langle\psi|)^2 d\psi\right)^{3/2}}{\left(\int (\operatorname{tr}\Delta |\psi\rangle\langle\psi|)^4 d\psi\right)^{1/2}}.$$

We now observe that the quantity $\int (\operatorname{tr}\Delta |\psi\rangle\langle\psi|)^2 d\psi$ can easily be calculated explicitly in terms of the Schatten 2-norm. We have

$$
\begin{aligned}
\operatorname{tr}\left(\int d\psi |\psi\rangle\langle\psi|^{\otimes 2}\right) \Delta^{\otimes 2} &= \operatorname{tr}\left(\frac{I+F}{n(n+1)}\right) \Delta^{\otimes 2} = \frac{1}{n(n+1)}\left((\operatorname{tr}\Delta)^2 + \operatorname{tr}\Delta^2\right) \\
&= \frac{1}{n(n+1)}\left((1-2p)^2 + \operatorname{tr}\Delta^2\right),
\end{aligned}
$$

where F is the swap operator that exchanges two n-dimensional systems. Thus, if we can upper bound $\int (\operatorname{tr}\Delta |\psi\rangle\langle\psi|)^4 d\psi$ in terms of $\int (\operatorname{tr}\Delta |\psi\rangle\langle\psi|)^2 d\psi$, this will give a lower bound on $\|\Delta\|_M$. Hypercontractivity on the sphere will allow us to do precisely this.

---

[1]This method seems to have been first used by Littlewood [44] in 1930.

2.1 Hypercontractivity and spherical harmonics

For any n, let S^n be the real n-sphere (i.e. $\{x \in \mathbb{R}^{n+1} : \sum_i x_i^2 = 1\}$), and for $f : S^n \to \mathbb{R}$ define the $L^p(S^n)$ norms as

$$\|f\|_{L^p(S^n)} := \left(\int |f(\xi)|^p d\xi \right)^{1/p},$$

where we integrate with respect to the uniform measure on S^n, normalised so that $\int d\xi = 1$. Identify each n-dimensional quantum state $|\psi\rangle$ (element of the unit sphere in $\mathbb{C}^n$) with a real vector $\xi \in S^{2n-1}$ by taking real and imaginary parts, and consider the function $f(\xi) = \operatorname{tr} \Delta |\psi\rangle\langle\psi|$. It is easy to convince oneself that $f : S^{2n-1} \to \mathbb{R}$ is a homogeneous degree 2 polynomial in the components of ξ, which will allow us to apply hypercontractivity to relate $\int (\operatorname{tr} \Delta |\psi\rangle\langle\psi|)^4 d\psi$ to $\int (\operatorname{tr} \Delta |\psi\rangle\langle\psi|)^2 d\psi$.

In order to state and use a hypercontractive inequality on the n-sphere, we will need some basic ideas from the theory of spherical harmonics (see e.g. the book [58] or the online notes [29] for background). The restriction of every degree d polynomial $f : \mathbb{R}^{n+1} \to \mathbb{R}$ to the sphere S^n can be written as

$$f(x) = \sum_{k=0}^{d} Y_k(x),$$

where $Y_k : S^n \to \mathbb{R}$ is called a *spherical harmonic*, and is the restriction of a degree k polynomial to the sphere, satisfying $\int Y_j(\xi)Y_k(\xi)d\xi = 0$ for $j \neq k$. Hence

$$\|f\|_{L^2(S^n)}^2 = \int f(\xi)^2 d\xi = \sum_{k=0}^{d} \|Y_k\|_{L^2(S^n)}^2,$$

a generalisation of Parseval's equality. The *Poisson semigroup* (which can be thought of as a "noise operator" for the sphere) is defined by

$$(P_\epsilon f)(x) = \sum_k \epsilon^k Y_k(x).$$

Crucially, it is known that the Poisson semigroup is indeed hypercontractive.

Theorem 7 (Beckner [7]). *If $1 \leq p \leq q \leq \infty$ and $\epsilon \leq \sqrt{\frac{p-1}{q-1}}$, then*

$$\|P_\epsilon f\|_{L^q(S^n)} \leq \|f\|_{L^p(S^n)}.$$

Theorem 7 will allow us to relate different L^p norms of low-degree polynomials on the sphere, as follows.

Corollary 8. *Let $f : \mathbb{R}^{n+1} \to \mathbb{R}$ be a degree d polynomial. Then, for any $q \geq 2$,*

$$\|f\|_{L^q(S^n)} \leq (q-1)^{d/2} \|f\|_{L^2(S^n)}.$$

Proof. The proof is exactly the same as that of Corollary 2 (see e.g. [50]). Explicitly, we write

$$
\begin{aligned}
\|f\|_{L^q(S^n)}^2 &= \left\| \sum_{k=0}^{d} Y_k \right\|_{L^q(S^n)}^2 = \left\| P_{1/\sqrt{q-1}} \left(\sum_{k=0}^{d} (q-1)^{k/2} Y_k \right) \right\|_{L^q(S^n)}^2 \\
&\leq \left\| \sum_{k=0}^{d} (q-1)^{k/2} Y_k \right\|_{L^2(S^n)}^2 = \sum_{k=0}^{d} (q-1)^k \|Y_k\|_{L^2(S^n)}^2 \leq (q-1)^d \|f\|_{L^2(S^n)}^2.
\end{aligned}
$$

6

The first equality is just expanding f in terms of spherical harmonics and the second follows from the definition of P_ϵ. The first inequality is Theorem 7, the third equality follows from Parseval's theorem, and the last inequality is obvious. $\square$

By Corollary 8, for any $p \geq 2$,

$$\left(\int (\operatorname{tr} \Delta |\psi\rangle\langle\psi|)^p d\psi \right)^{1/p} \leq (p-1) \left(\int (\operatorname{tr} \Delta |\psi\rangle\langle\psi|)^2 d\psi \right)^{1/2}.$$

Taking $p = 4$ this implies Theorem 6, via

$$
\begin{aligned}
\|\Delta\|_M &\geq n \frac{\left(\int (\operatorname{tr} \Delta |\psi\rangle\langle\psi|)^2 d\psi \right)^{3/2}}{\left(\int (\operatorname{tr} \Delta |\psi\rangle\langle\psi|)^4 d\psi \right)^{1/2}} \geq \frac{n}{9} \left(\int (\operatorname{tr} \Delta |\psi\rangle\langle\psi|)^2 d\psi \right)^{1/2} \\
&= \frac{1}{9(1 + 1/n)^{1/2}} \left((1 - 2p)^2 + \operatorname{tr} \Delta^2 \right)^{1/2}.
\end{aligned}
$$

The works [5, 45] considered the special case where Δ is traceless and gave a better constant of $1/3$ rather than $1/9$; the more recent result of Lancien and Winter [43] extends the inequality to the case where Δ is not traceless and achieves a constant $1/\sqrt{18} \approx 1/4.243$.

2.2 Hypercontractivity on k copies of the sphere

We will now apply the approach of the previous section to the multipartite setting, where the power of the hypercontractive approach will become apparent. We will prove the following, reproducing the main result of [43] with a worse constant.

Theorem 9. *Let M be a k-fold tensor product 4-design and set $\Delta = p\rho - (1-p)\sigma$, for quantum states $\rho, \sigma \in B((\mathbb{C}^n)^{\otimes k})$. Then*

$$\|\Delta\|_M \geq \frac{1}{(81(1 + 1/n))^{k/2}} \left(\sum_{S \subseteq [k]} \| \operatorname{tr}_S \Delta \|_2^2 \right)^{1/2},$$

where $\|\Delta\|_2$ is the Schatten 2-norm.

To prove Theorem 9, we begin by mimicking the start of the proof of Theorem 6 and using the fact that M is a tensor product of local 4-designs to obtain

$$\|\Delta\|_M \geq n^k \frac{\left(\int \ldots \int d\psi_1 \ldots d\psi_k (\operatorname{tr} \Delta (|\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_k\rangle\langle\psi_k|))^2 \right)^{3/2}}{\left(\int \ldots \int d\psi_1 \ldots d\psi_k (\operatorname{tr} \Delta (|\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_k\rangle\langle\psi_k|))^4 \right)^{1/2}} = n^k \frac{\|f\|_{L^2((S^{2n-1})^k)}^3}{\|f\|_{L^4((S^{2n-1})^k)}^2}, \quad (3)$$

where we define the function $f : (S^{2n-1})^k \to \mathbb{R}$ by

$$f(\xi_1, \ldots, \xi_k) = \operatorname{tr} \Delta (|\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_k\rangle\langle\psi_k|),$$

where $|\psi_i\rangle$ is the n-dimensional complex unit vector whose real and imaginary parts are given by $\xi_i \in S^{2n-1}$ in the obvious way. In order to relate the denominator in (3) to the numerator (which can easily be calculated, as we will see below), we require an extension of Theorem 7 and Corollary 2 to $(S^{2n-1})^k$. A suitable extension of Theorem 7 immediately follows from multiplicativity of the operator norm of the Poisson operator, which can be proven using standard and generic arguments (see e.g. the proof in [6] or the version in [26]).

Lemma 10. *Fix $1 \leq p \leq q$ and consider linear operators $M_1 : L^p(S_1, \mu_1) \to L^q(T_1, \nu_1)$ and $M_2 : L^p(S_2, \mu_2) \to L^q(T_2, \nu_2)$. Then $\|M_1 \otimes M_2\| \leq \|M_1\|\|M_2\|$.*

We thus immediately have the following corollary of Theorem 7.

Corollary 11. *Let $f : (S^n)^k \to \mathbb{R}$. If $1 \leq p \leq q \leq \infty$ and $\epsilon \leq \sqrt{\frac{p-1}{q-1}}$, then*

$$\|P_\epsilon^{\otimes k} f\|_{L^q((S^n)^k)} \leq \|f\|_{L^p((S^n)^k)}.$$

This in turn allows us to prove the required generalisation of Corollary 2, which is also almost immediate, but for which we include a proof for completeness.

Corollary 12. *Let $f : (\mathbb{R}^{n+1})^k \to \mathbb{R}$ be a degree d polynomial in the components of each $x^1, \ldots, x^k \in \mathbb{R}^{n+1}$. Then, for any $q \geq 2$,*

$$\|f\|_{L^q((S^n)^k)} \leq (q-1)^{dk/2}\|f\|_{L^2((S^n)^k)}.$$

Proof. The restriction of f to $(S^n)^k$ can be written as a sum of products of spherical harmonics, i.e.

$$f(\xi_1, \ldots, \xi_k) = \sum_{i_1, \ldots, i_k=0}^{d} Y_{i_1, \ldots, i_k}(\xi_1, \ldots, \xi_k),$$

where each $Y_{i_1, \ldots, i_k}$ is a sum of products of degree $i_1, \ldots, i_k$ spherical harmonics on S^n, $Y_{i_1, \ldots, i_k} = \sum_m Y_{i_1}^m \ldots Y_{i_k}^m$, and hence $\int Y_{i_1, \ldots, i_k}(\xi_1, \ldots, \xi_k)Y_{j_1, \ldots, j_k}(\xi_1, \ldots, \xi_k)d\xi_1 \ldots d\xi_k = 0$ if $i_\ell \neq j_\ell$ for some ℓ. Then

$$
\begin{aligned}
\|f\|_{L^q((S^n)^k)}^2 &= \left\|\sum_{i_1, \ldots, i_k=0}^{d} Y_{i_1, \ldots, i_k}\right\|_{L^q((S^n)^k)}^2 = \left\|P_{1/\sqrt{q-1}}^{\otimes k}\left(\sum_{i_1, \ldots, i_k=0}^{d} (q-1)^{\sum_j i_j/2} Y_{i_1, \ldots, i_k}\right)\right\|_{L^q((S^n)^k)}^2 \\
&\leq \left\|\sum_{i_1, \ldots, i_k=0}^{d} (q-1)^{\sum_j i_j/2} Y_{i_1, \ldots, i_k}\right\|_{L^2((S^n)^k)}^2 = \sum_{i_1, \ldots, i_k=0}^{d} (q-1)^{\sum_j i_j}\|Y_{i_1, \ldots, i_k}\|_{L^2((S^n)^k)}^2 \\
&\leq (q-1)^{dk}\|f\|_{L^2((S^n)^k)}^2.
\end{aligned}
$$

$\square$

We can now finally complete the proof of Theorem 9. From (3) and Corollary 12, we have

$$\|\Delta\|_M \geq n^k \frac{\|f\|_{L^2((S^{2n-1})^k)}^3}{\|f\|_{L^4((S^{2n-1})^k)}^2} \geq \left(\frac{n}{9}\right)^k \|f\|_{L^2((S^{2n-1})^k)}.$$

All that remains is to explicitly calculate

$$
\begin{aligned}
\|f\|_{L^2((S^{2n-1})^k)}^2 &= \mathrm{tr}\left(\int \ldots \int d\psi_1 \ldots d\psi_k |\psi_1\rangle\langle\psi_1|^{\otimes 2} \otimes \cdots \otimes |\psi_k\rangle\langle\psi_k|^{\otimes 2}\right)\Delta^{\otimes 2} \\
&= \mathrm{tr}\left(\frac{I+F}{n(n+1)}\right)^{\otimes k}\Delta^{\otimes 2} = \frac{1}{n^k(n+1)^k}\sum_{S \subseteq [k]} \|\mathrm{tr}_S \Delta\|_2^2,
\end{aligned}
$$

which completes the proof of Theorem 9.

We have therefore reproduced the main result of [43] (with slightly worse constants). This result is almost optimal, as Lancien and Winter give a simple example ($\Delta = \left(\frac{1}{2}(|0\rangle\langle0| - |1\rangle\langle1|)\right)^{\otimes k}$) where the bias goes down exponentially with k.

It is far from clear that the proof here is "simpler" than that of [43], which is arguably more intuitive and explicit. One potential advantage of the present approach is that it naturally explains the presence of an exponential prefactor in Theorem 9 in terms of the multiplicativity of the operator norm of the Poisson semigroup. The proof technique here is also somehow more general, as Corollary 12 gives bounds on all L^q norms of f, rather than just the L^4 norm considered in [43], which may be useful in the study of q-designs for $q > 4$. By contrast, although extending the combinatorial approach of [5, 45, 43] to arbitrary q would be possible, it seems likely to require more work. In the special case $k = 1$, the approach of [5, 45, 43] can indeed be pushed to arbitrary q without too great an effort, and yields better constants than the result given here (Andreas Winter, personal communication).

It is interesting to note that the quantity $\|M\|_{2(k)} := \left(\sum_{S \subseteq [k]} \|\operatorname{tr}_S M\|_2^2\right)^{1/2}$ which appears in Theorem 9 has occurred before in a different setting. If we let M be a quantum state (rather than the difference of two quantum states as considered above) then $\|M\|_{2(k)}^2$ is equal to the probability of M passing a certain natural test for being a product state [33], up to a constant depending on k and the local dimensions. The main result of [33] is an upper bound on $\|M\|_{2(k)}^2$ in terms of the injective tensor norm of M, or in other words the maximal overlap of M with a product state. It is an intriguing open question whether the perspective taken here of relating different norms of M could lead to a simpler or stronger proof of the results of [33].

3 Spectral concentration for k-local Hamiltonians

There is a natural extension of the idea of hypercontractivity to a noncommutative setting. Here we will only consider one particular such extension, from functions on the boolean cube $\{0,1\}^n$ to linear operators on the space of n qubits [46]. Noncommutative hypercontractive inequalities in quantum information have been studied very recently in much greater generality by Kastoryano and Temme, from the perspective of log-Sobolev inequalities [39].

Let $M \in B((\mathbb{C}^2)^{\otimes n})$ be a Hermitian operator on the space of n qubits. The natural noncommutative analogue of the noise operator defined in eqn. (1) turns out to be the tensor product of n copies of the qubit depolarising channel $D_\epsilon M = (1 - \epsilon)(\operatorname{tr} M)\frac{I}{2} + \epsilon M$. This operator also has a pleasant "Fourier-side" description, as follows. Any operator $M \in B((\mathbb{C}^2)^{\otimes n})$ can be expanded in terms of tensor products of Pauli matrices. Identify each string $s \in \{0,1,2,3\}^n$ with the product $\sigma_s := \sigma_{s_1} \otimes \cdots \otimes \sigma_{s_n}$ (where $\sigma_0 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, $\sigma_1 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $\sigma_2 = \left(\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right)$, $\sigma_3 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$), and write $M = \sum_{s \in \{0,1,2,3\}^n} \widehat{M}(s)\sigma_s$. Then

$$D_\epsilon^{\otimes n} M = \sum_{s \in \{0,1,2,3\}^n} \epsilon^{|s|} \widehat{M}(s)\sigma_s,$$

where $|s|$ is the number of non-zero components of s. Observe that the natural analogue of degree k polynomials on the boolean cube (which have no Fourier coefficients of weight greater than k) is operators which have no Pauli coefficients of weight greater than k, i.e. the class of k-local operators.

As one might hope by analogy with the commutative case, $D_\epsilon^{\otimes n}$ does indeed satisfy a hypercontractive inequality [46].

Theorem 13. *Let M be a Hermitian operator on n qubits and assume that $1 \leq p \leq 2 \leq q \leq \infty$. Then, if $\epsilon \leq \sqrt{\frac{p-1}{q-1}}$, $\|D_\epsilon^{\otimes n} M\|_q \leq \|M\|_p$.*

In this theorem, and the rest of this section, $\|M\|_p$ is the *normalised* Schatten p-norm, $\|M\|_p = \left(\frac{1}{2^n} \operatorname{tr} |M|^p\right)^{1/p}$. Following the completion of this work, King has extended Theorem 13 from the depolarising channel to arbitrary qubit channels that belong to self-adjoint semigroups, and has removed the restriction that $p \leq 2 \leq q$ [41]. (Noncommutative hypercontractive inequalities for certain subalgebras of $B((\mathbb{C}^2)^{\otimes n})$ had previously been given by Carlen and Lieb [19] and Biane [11]; see [41] for a discussion.) The following corollary of Theorem 13 was stated in [46], by exact analogy with the proof of Corollary 2.

Corollary 14. *Let M be a k-local Hermitian operator on n qubits. Then, for any $q \geq 2$, $\|M\|_q \leq (q-1)^{k/2}\|M\|_2$. Also, for any $p \leq 2$, $\|M\|_p \geq (p-1)^{k/2}\|M\|_2$.*

We now observe that Corollary 14 allows one to easily prove general tail bounds for eigenvalues of k-local Hamiltonians, using exactly the same proof as for classical tail bounds on degree k polynomials [27, 51].

Theorem 15. *Let M be a k-local Hermitian operator on n qubits with eigenvalues (λ_i) such that $\|M\|_2 = 1$. Then, for any $t \geq (2e)^{k/2}$,*

$$\frac{|\{i : |\lambda_i| \geq t\}|}{2^n} \leq \exp(-kt^{2/k}/(2e)).$$

Proof. For any $r \geq 2$, we have

$$\frac{|\{i : |\lambda_i| \geq t\}|}{2^n} = \frac{|\{i : |\lambda_i|^r \geq t^r\}|}{2^n} \leq \frac{\sum_i |\lambda_i|^r}{2^n t^r} = \frac{\|M\|_r^r}{t^r} \leq \frac{((r-1)^{k/2}\|M\|_2)^r}{t^r} \leq (r^{k/2}/t)^r,$$

where the first inequality is Markov's inequality, the second is Corollary 14, and the third follows from the conditions of the theorem. Now, minimising this expression by taking $r = t^{2/k}/e$, we have

$$\frac{|\{i : |\lambda_i| \geq t\}|}{2^n} \leq \exp(-kt^{2/k}/(2e))$$

as claimed. □

One can, of course, generalise this bound to systems of d-dimensional qudits by considering each qudit as $\lceil \log_2 d \rceil$ qubits, at the expense of increasing k to $k\lceil \log_2 d \rceil$. Similar (and/or stronger) concentration bounds to Theorem 15 have been proven before by other authors in more restricted settings. For example, Hartmann, Mahler and Hess [34] have shown that in a setting where subsystems interact with nearest neighbours in a linear chain, the spectrum in fact converges to a normal distribution. Theorem 15 gives a less precise result, but makes no assumptions about the interaction geometry and has a much more concise proof.

Physically, Theorem 15 says that the probability that a random state has high energy is exponentially small. One consequence is that, for any k-local Hamiltonian H on n qubits, there must exist very large subspaces of states on which the time evolution e^{-iHt} is slow. Indeed, one can use Theorem 15 to prove bounds on the probability that a state does not significantly change after evolving according to H, as in the following example. Let H be a 2-local Hamiltonian on n qubits

10

with eigendecomposition $H = \sum_k \lambda_k |v_k\rangle\langle v_k|$, such that $\|H\|_2 = 1$. Then, for any initial state $|\psi\rangle$ and any $\mu \geq 0$,

$$
\begin{aligned}
|\langle\psi|e^{-iHt}|\psi\rangle| &= \left| \sum_{k,|\lambda_k|\leq\mu} e^{-i\lambda_k t}|\langle v_k|\psi\rangle|^2 + \sum_{k,|\lambda_k|>\mu} e^{-i\lambda_k t}|\langle v_k|\psi\rangle|^2 \right| \\
&\geq \cos(\mu t)\left(1 - \sum_{k,|\lambda_k|>\mu} |\langle v_k|\psi\rangle|^2\right) - \sum_{k,|\lambda_k|>\mu} |\langle v_k|\psi\rangle|^2 \\
&\geq \cos(\mu t) - 2\sum_{k,|\lambda_k|>\mu} |\langle v_k|\psi\rangle|^2.
\end{aligned}
$$

By Theorem 15, for any $\mu \geq 2e$ the dimension of the subspace spanned by $|v_k\rangle$ such that $|\lambda_k| > \mu$ is at most $e^{-\mu/e}2^n$. Now imagine we pick $|\psi\rangle$ at random (i.e. according to Haar measure on the complex unit sphere). This implies, via a standard tail bound for projector overlaps [9], that for any $\delta > 0$

$$
\Pr_{|\psi\rangle}\left[\sum_{k,|\lambda_k|>\mu} |\langle v_k|\psi\rangle|^2 \geq (1+\delta)e^{-\mu/e} \right] \leq \exp(-e^{-\mu/e}2^n(\delta - \ln(1+\delta))/(\ln 2)).
$$

Fixing (for example) $\delta = 1$, we obtain for any $\mu \geq 2e$ that

$$
\Pr_{|\psi\rangle}\left[|\langle\psi|e^{-iHt}|\psi\rangle| \leq \cos(\mu t) - 4e^{-\mu/e} \right] \leq \exp(-\Omega(e^{-\mu/e}2^n)).
$$

Other spectral inequalities follow from Corollary 14. For example, the following quantum generalisation of the classical Schwartz-Zippel lemma was observed in [46].

Corollary 16. *Let H be a non-zero k-local Hermitian operator on n qubits with rank r. Then $r \geq 2^{n-(2\log_2 e)k} \approx 2^{n-2.89k}$.*

4   General bounds on XOR games

A homogeneous polynomial $f : (\mathbb{R}^n)^k \to \mathbb{R}$ is said to be a *multilinear form* if it is linear in each input, i.e. $f(x^1 + y, x^2, \ldots, x^k) = f(x^1, \ldots, x^k) + f(y, x^2, \ldots, x^k)$ for all $x^i \in \mathbb{R}^n$ and all $y \in \mathbb{R}^n$, and similarly for the other positions. Any multilinear form can be written as

$$
f(x^1, \ldots, x^k) = \sum_{i_1,\ldots,i_k} \hat{f}_{i_1,\ldots,i_k} x^1_{i_1} x^2_{i_2} \ldots x^k_{i_k}
$$

for some multidimensional array (tensor) $\hat{f} \in \mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n$. Observe that, if we consider $x^j \in \{\pm 1\}^n$, this expression is precisely the Fourier expansion of f as a function on the boolean cube $\{\pm 1\}^{nk}$, justifying the use of the notation $\hat{f}$. The (normalised) ℓ_p norms of f as a function on the boolean cube are thus given by

$$
\|f\|_p := \left(\frac{1}{2^{nk}} \sum_{x^1,\ldots,x^k\in\{\pm 1\}^n} |f(x^1,\ldots,x^k)|^p \right)^{1/p},
$$

and in particular

$$\|f\|_\infty := \max_{x^1,\ldots,x^k \in \{\pm 1\}^n} |f(x^1,\ldots,x^k)|.$$

Any XOR game $G = (\pi, A)$ corresponds to a multilinear form f by taking

$$f(x^1,\ldots,x^k) = \sum_{i_1,\ldots,i_k} \pi_{i_1,\ldots,i_k} A_{i_1,\ldots,i_k} x^1_{i_1} x^2_{i_2} \ldots x^k_{i_k},$$

and the bias $\beta(G)$ is precisely $\|f\|_\infty$. The following result, which is known as the Bohnenblust-Hille inequality, will allow us to find a general lower bound on $\beta(G)$.

Theorem 17 (Bohnenblust-Hille inequality [13, 24, 52]). *For any multilinear form $f : (\mathbb{R}^n)^k \to \mathbb{R}$, and any $p \geq 2k/(k+1)$,*

$$\|\hat{f}\|_p := \left(\sum_{i_1,\ldots,i_k} |\hat{f}_{i_1,\ldots,i_k}|^p \right)^{1/p} \leq C_k \|f\|_\infty,$$

where C_k may be taken to be $O(k^{\log_2 e}) \approx O(k^{1.45})$.

The special case $k = 2$ was previously proven by Littlewood [44] and is known as Littlewood's $4/3$ inequality. The original proof of Theorem 17 given by Bohnenblust and Hille had C_k growing exponentially with k [13]. Pellegrino and Seoane-Sepúlveda have recently shown that in fact one can take $C_k = \text{poly}(k)$ in this inequality [52] (see also [25, 49, 56] for further and more precise bounds on C_k). The proof in [52] is formally a consequence of prior, and very general, work by Defant, Popa and Schwarting [24], but makes various careful combinatorial choices to achieve an explicit bound on C_k which is polynomial in k; we therefore give credit for Theorem 17 to both sets of authors. Here we will give a modified proof, based on hypercontractivity, of this result.

First, we observe that Theorem 17 has the following corollary for XOR games.

Corollary 18. *Let G be a k-player XOR game with n possible inputs per player. Then $\beta(G) = \Omega(k^{-3/2} n^{-(k-1)/2})$.*

Proof. Apply Theorem 17 to the multilinear form f corresponding to G, taking $p = 2k/(k+1)$, and use the inequality $\|\hat{f}\|_p \geq n^{k(1/p-1)} \|\hat{f}\|_1 = n^{k(1/p-1)}$. $\square$

In order to prove Theorem 17 we will need the following inequality of Defant, Popa and Schwarting [24] (extending an inequality of Blei [12]), which can be proven by the careful application of Hölder's inequality.

Lemma 19. *Let $A = (A_{ij})$ be a matrix whose columns are (α_i) and whose rows are (β_j). Then, for any $m \geq 1$,*

$$\left(\sum_{i,j} |A_{ij}|^{2m/(m+1)} \right)^{(m+1)/(2m)} \leq \left(\sum_i \|\alpha_i\|_2^{2m/(m+2)} \right)^{(m+2)/(4m)} \left(\sum_j \|\beta_j\|_2^{2m/(m+2)} \right)^{(m+2)/(4m)}.$$

As with the previous two applications of hypercontractivity discussed, the proof of Theorem 17 that we give here will also use hypercontractivity via a corollary: in this case the perhaps more well-known Corollary 2 for functions on the boolean cube. We stress that the proof given here

follows the same lines as that of Pellegrino and Seoane-Sepúlveda; the use of Corollary 2 simply replaces an equivalent step in their proof (and indeed gives worse constants). However, using this corollary seems (to the author) to make the proof somewhat simpler and more transparent, and illustrates how hypercontractivity can be used as a "black box".

Proof of Theorem 17 (Defant, Popa and Schwarting [24]; Pellegrino and Seoane-Sepúlveda [52]).
As $\|\hat{f}\|_p$ is nonincreasing with p, it suffices to prove the theorem for $p = 2k/(k+1)$. The proof will be by induction on k, first assuming that k is a power of 2. The base case $k = 1$ is trivial, and in this case we have $C_1 = 1$. So, assuming the theorem holds for $k/2$, we prove it holds for k. By Lemma 19,

$$
\left(\sum_{i_1,\ldots,i_k} |\hat{f}_{i_1,\ldots,i_k}|^{2k/(k+1)} \right)^{(k+1)/(2k)} \leq \left(\sum_{i_1,\ldots,i_{k/2}} \|(\hat{f}_{i_1,\ldots,i_k})_{i_{k/2+1},\ldots,i_k=1}^n\|_2^{2k/(k+2)} \right)^{(k+2)/4k}
$$
$$
\times \left(\sum_{i_{k/2+1},\ldots,i_k} \|(\hat{f}_{i_1,\ldots,i_k})_{i_1,\ldots,i_{k/2}=1}^n\|_2^{2k/(k+2)} \right)^{(k+2)/4k} \quad (4)
$$

We estimate the second term (the first follows exactly the same procedure). For each $i_{k/2+1},\ldots,i_k \in \{1,\ldots,n\}$, define the function $f_{i_{k/2+1},\ldots,i_k} : (\mathbb{R}^n)^{k/2} \to \mathbb{R}$ by

$$
f_{i_{k/2+1},\ldots,i_k}(x^1,\ldots,x^{k/2}) = \sum_{i_1,\ldots,i_{k/2}} \hat{f}_{i_1,\ldots,i_k} x_{i_1}^1 x_{i_2}^2 \ldots x_{i_{k/2}}^{k/2}.
$$

Also define a "dual" function $f'_{x^1,\ldots,x^{k/2}} : (\mathbb{R}^n)^{k/2} \to \mathbb{R}$ by

$$
f'_{x^1,\ldots,x^{k/2}}(x^{k/2+1},\ldots,x^k) = f(x^1,\ldots,x^k),
$$

i.e. with respect to ± 1-valued inputs, $f'_{x^1,\ldots,x^{k/2}}$ is just the restriction of f to the subcube produced by fixing $x^1,\ldots,x^{k/2}$. It is easy to verify that $f'_{x^1,\ldots,x^{k/2}}$ can be written as

$$
f'_{x^1,\ldots,x^{k/2}}(x^{k/2+1},\ldots,x^k) = \sum_{i_{k/2+1},\ldots,i_k=1}^n f_{i_{k/2+1},\ldots,i_k}(x^1,\ldots,x^{k/2}) x_{i_{k/2+1}}^{k/2+1} \ldots x_{i_k}^k;
$$

of course $\|f'_{x^1,\ldots,x^{k/2}}\|_\infty \leq \|f\|_\infty$. For each tuple $i_{k/2+1},\ldots,i_k$ we have by Parseval's equality

$$
\|(\hat{f}_{i_1,\ldots,i_k})_{i_1,\ldots,i_{k/2}=1}^n\|_2 = \left(\sum_{i_1,\ldots,i_{k/2}=1}^n \hat{f}_{i_1,\ldots,i_k}^2 \right)^{1/2} = \|f_{i_{k/2+1},\ldots,i_k}\|_2.
$$

By the hypercontractive estimate of Corollary 2,

$$
\|f_{i_{k/2+1},\ldots,i_k}\|_2^{2k/(k+2)} \leq \left(\frac{k+2}{k-2} \right)^{\frac{k^2}{2(k+2)}} \|f_{i_{k/2+1},\ldots,i_k}\|_{2k/(k+2)}^{2k/(k+2)}.
$$

We now observe that, for any $p \geq 1$,

$$
\sum_{i_{k/2+1},\ldots,i_k} \|f_{i_{k/2+1},\ldots,i_k}\|_p^p = \mathbb{E}_{x^1,\ldots,x^{k/2}} \left[\sum_{i_{k/2+1},\ldots,i_k} |f_{i_{k/2+1},\ldots,i_k}(x^1,\ldots,x^{k/2})|^p \right]
$$
$$
= \mathbb{E}_{x^1,\ldots,x^{k/2}} \left[\|\hat{f}'_{x^1,\ldots,x^{k/2}}\|_p^p \right].
$$

Hence, taking $p = 2k/(k+2) = 2(k/2)/(k/2+1)$, we have by the inductive hypothesis

$$\sum_{i_{k/2+1},\ldots,i_k} \|(\hat{f}_{i_1,\ldots,i_k})^n_{i_1,\ldots,i_{k/2}=1}\|_2^{2k/(k+2)} \leq \mathbb{E}_{x^1,\ldots,x^{k/2}} \left[\|\hat{f}'_{x^1,\ldots,x^{k/2}}\|_{2k/(k+2)}^{2k/(k+2)}\right]$$

$$\leq \left(\frac{k+2}{k-2}\right)^{\frac{k^2}{2(k+2)}} C_{k/2}^{2k/(k+2)} \|f\|_\infty^{2k/(k+2)},$$

so, combining both terms in the inequality (4),

$$\left(\sum_{i_1,\ldots,i_k} |\hat{f}_{i_1,\ldots,i_k}|^{2k/(k+1)}\right)^{(k+1)/(2k)} \leq \left(\frac{k+2}{k-2}\right)^{k/4} C_{k/2}\|f\|_\infty.$$

Thus

$$C_k \leq \left(1 + \frac{4}{k-2}\right)^{k/4} C_{k/2}.$$

Observing that $(1 + 4/(k-2))^{k/4} \leq (1 + O(1/k))e$, we have $C_k = O(k^{\log_2 e})$ as claimed.

Finally, if k is not a power of 2, we simply increase it to the next smallest power of 2 (redefining f appropriately), which corresponds to at most a constant increase in C_k. □

We remark that Theorem 17 also proves a very special case of a conjecture of Aaronson and Ambainis [1] that every bounded low-degree polynomial on the boolean cube has an influential variable. Define the *influence* of the j'th variable on a function $f : \{\pm 1\}^n \to \mathbb{R}$ as $I_j(f) = \frac{1}{2^{n+2}} \sum_{x\in\{\pm 1\}^n}(f(x) - f(x^j))^2$, where x^j is x with the j'th variable negated. The influence also has a concise Fourier-side description: $I_j(f) = \sum_{S\ni j} \hat{f}(S)^2$. The conjecture of [1] is that for all degree d polynomials $f : \{\pm 1\}^n \to [-1, 1]$, there exists a j such that $I_j(f) \geq \text{poly}(\text{Var}(f)/d)$, where $\text{Var}(f) = \sum_{S\neq\emptyset} \hat{f}(S)^2$ is the ℓ_2 variance of f. If this conjecture were true, it would imply (informally) that all quantum query algorithms could be efficiently simulated by classical query algorithms on most inputs [1].

Using Theorem 17, it is easy to prove the Aaronson-Ambainis conjecture in the very particular case where f is a multilinear form such that $\hat{f}_{i_1,\ldots,i_k} = \pm\alpha$, for some α, as we now show. If f is a multilinear form as above, it depends on nk variables x^j_ℓ, where $1 \leq j \leq k$ and $1 \leq \ell \leq n$. Observe that the influence of variable (j,ℓ) on f is

$$\text{Inf}_{(j,\ell)}(f) = \sum_{i_1,\ldots,i_{j-1},i_{j+1},\ldots,i_k} \hat{f}^2_{i_1,\ldots,i_{j-1},\ell,i_{j+1},\ldots,i_k}.$$

We can now state the following corollary of Theorem 17.

Corollary 20. *If f is a multilinear form such that $\|f\|_\infty \leq 1$ and $\hat{f}_{i_1,\ldots,i_k} = \pm\alpha$ for some α, then $I_{(j,\ell)}(f) = \Omega(\text{Var}(f)^2/k^3)$ for all (j,ℓ).*

Proof. We simply calculate $\text{Var}(f) = n^k\alpha^2$, $I_{(j,\ell)}(f) = n^{k-1}\alpha^2$, and $\|\hat{f}\|_{2k/(k+1)} = n^{(k+1)/2}\alpha = O(k^{3/2})$ by Theorem 17. Thus $\text{Var}(f)^2/I_{(j,\ell)}(f) = n^{k+1}\alpha^2 = O(k^3)$ as required. □

Andris Ambainis has subsequently generalised this result to multilinear forms f such that $\hat{f}_{i_1,\ldots,i_k} \in \{-\alpha, 0, \alpha\}$, and $\hat{f}_{i_1,\ldots,i_k} \neq 0$ at $n^{\Omega(k)}$ positions $(i_1,\ldots,i_k)$ (personal communication). Generalising further to arbitrary multilinear forms might be an interesting way of making progress on the Aaronson-Ambainis conjecture.

5 Outlook

This work has presented several examples of results in quantum information theory which can be proven using hypercontractive inequalities. It is, of course, debatable as to whether the proofs given here are really simpler or more intuitive than previously known proofs, especially if one is not initially familiar with hypercontractivity. The author's feeling is that the proofs given here seem less technical than the original proofs, at the expense of being less explicit; it is hoped that hypercontractivity will continue to develop as a tool in quantum information theory.

Acknowledgements

References

[1] S. Aaronson and A. Ambainis. The need for structure in quantum speedups. In *Proceedings of ICS 2011*, pages 338–352, 2011. `arXiv:0911.0996`.

[2] N. Alon and A. Naor. Approximating the cut-norm via Grothendieck's inequality. *SIAM J. Comput.*, 35(4):787–803, 2006.

[3] N. Alon and J. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley, New York, 2000.

[4] A. Ambainis and R. de Wolf. How low can approximate degree and quantum query complexity be for total boolean functions?, 2012. `arXiv:1206.0717`.

[5] A. Ambainis and J. Emerson. Quantum t-designs: t-wise independence in the quantum world. In *Proc. 22nd Annual IEEE Conf. Computational Complexity*, pages 129–140, 2007. `quant-ph/0701126`.

[6] W. Beckner. Inequalities in Fourier analysis. *Ann. of Math.*, 102:159–182, 1975.

[7] W. Beckner. Sobolev inequalities, the Poisson semigroup, and analysis on the sphere S^n. *Proc. Natl. Acad. Sci. USA*, 89:4816–4819, 1992.

[8] A. Ben-Aroya, O. Regev, and R. de Wolf. A hypercontractive inequality for matrix-valued functions with applications to quantum computing and LDCs. In *Proc. 49th Annual Symp. Foundations of Computer Science*, pages 477–486, 2008. `arXiv:0705.3806`.

[9] C. H. Bennett, P. Hayden, D. Leung, P. Shor, and A. Winter. Remote preparation of quantum states. *IEEE Trans. Inform. Theory*, 51(1):56–74, 2005. `quant-ph/0307100`.

[10] B. Berger. The fourth moment method. *SIAM J. Comput.*, 24(6):1188–1207, 1997.

[11] P. Biane. Free hypercontractivity. *Comm. Math. Phys.*, 184(2):457–474, 1997.

[12] R. Blei. *Analysis in integer and fractional dimensions*. Cambridge University Press, 2001.

[13] H. Bohnenblust and E. Hille. On the absolute convergence of Dirichlet series. *Annals of Mathematics*, 32(3):600–622, 1931.

[14] A. Bonami. Étude des coefficients Fourier des fonctiones de Lp(G). *Ann. Inst. Fourier*, 20:335–402, 1970.

[15] J. Bourgain. On the distribution of the Fourier spectrum of Boolean functions. *Isr. J. Math.*, 131(1):269–276, 2002.

[16] J. Briët, H. Buhrman, T. Lee, and T. Vidick. Multiplayer XOR games and quantum communication complexity with clique-wise entanglement, 2009. `arXiv:0911.4007`.

[17] J. Briët and T. Vidick. Explicit lower and upper bounds on the entangled value of multiplayer XOR games, 2011. `arXiv:1108.5647`.

[18] H. Buhrman, O. Regev, G. Scarpa, and R. de Wolf. Near-optimal and explicit Bell inequality violations. In *Proc. 26th Annual IEEE Conf. Computational Complexity*, pages 157–166, 2011. `arXiv:1012.5043`.

[19] E. Carlen and E. Lieb. Optimal hypercontractivity for Fermi fields and related non-commutative integration inequalities. *Comm. Math. Phys.*, 155:27–46, 1993.

[20] F. Chung and P. Tetali. Communication complexity and quasi randomness. *SIAM J. Discrete Math.*, 6(1):110–125, 1993.

[21] J. Clauser, M. Horne, A. Shimony, and R. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23:880–884, 1969.

[22] R. Cleve, P. Høyer, B. Toner, and J. Watrous. Consequences and limits of nonlocal strategies. In *Proc. 19th Annual IEEE Conf. Computational Complexity*, pages 236–249, 2004. `quant-ph/0404076`.

[23] E. B. Davies, L. Gross, and B. Simon. Hypercontractivity: a bibliographic review. *Ideas and methods in quantum and statistical physics (Oslo, 1988)*, pages 370–389, 1992.

[24] A. Defant, D. Popa, and U. Schwarting. Coordinatewise multiple summing operators in Banach spaces. *J. Funct. Anal.*, 259:220–242, 2010.

[25] D. Diniz, G. Muñoz-Fernández, D. Pellegrino, and J. Seoane-Sepúlveda. The asymptotic growth of the constants in the Bohnenblust-Hille inequality is optimal. *J. Funct. Anal.*, 263:415–428, 2012. `arXiv:1108.1550`.

[26] I. Dinur and E. Friedgut. Analytical methods in combinatorics and computer-science, 2005. `http://www.cs.huji.ac.il/~analyt/`.

[27] I. Dinur, E. Friedgut, G. Kindler, and R. O'Donnell. On the Fourier tails of bounded functions over the discrete cube. *Israel Journal of Mathematics*, 160:389–412, 2007.

[28] J. Ford and A. Gál. Hadamard tensors and lower bounds on multiparty communication complexity. In *Proc. 32nd International Conference on Automata, Languages and Programming (ICALP'05)*, pages 1163–1175, 2005.

[29] J. Gallier. Notes on spherical harmonics and linear representations of Lie groups, 2009. `http://www.cis.upenn.edu/~cis610/sharmonics.pdf`.

[30] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf. Exponential separations for one-way quantum communication complexity, with applications to cryptography. In *Proc. 39th Annual ACM Symp. Theory of Computing*, pages 516–525, 2007. `quant-ph/0611209`.

[31] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975.

[32] L. Gross. Hypercontractivity, logarithmic Sobolev inequalities, and applications: a survey of surveys. In *Diffusion, Quantum Theory, and Radically Elementary Mathematics*, pages 45–74. Princeton University Press, 2006.

[33] A. Harrow and A. Montanaro. An efficient test for product states, with applications to quantum Merlin-Arthur games. In *Proc. 51st Annual Symp. Foundations of Computer Science*, pages 633–642, 2010. `arXiv:1001.0017`.

[34] M. Hartmann, G. Mahler, and O. Hess. Gaussian quantum fluctuations in interacting many particle systems. *Lett. Math. Phys.*, 68(103–112), 2004. `math-ph/0312045`.

[35] C. W. Helstrom. *Quantum detection and estimation theory*. Academic Press, New York, 1976.

[36] A. S. Holevo. Statistical decision theory for quantum systems. *Journal of Multivariate Analysis*, 3:337–394, 1973.

[37] S. Janson. *Gaussian Hilbert spaces*. Cambridge University Press, 1997.

[38] J. Kahn, G. Kalai, and N. Linial. The influence of variables on Boolean functions. In *Proc. 29th Annual Symp. Foundations of Computer Science*, pages 68–80, 1988.

[39] M. Kastoryano and K. Temme. Quantum logarithmic Sobolev inequalities and rapid mixing, 2012. `arXiv:1207.3261`.

[40] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM J. Comput.*, 37(1):319–357, 2007.

[41] C. King. Hypercontractivity for semigroups of unital qubit channels, 2012. `arXiv:1210.8412`.

[42] B. Klartag and O. Regev. Quantum one-way communication can be exponentially stronger than classical communication. In *Proc. 43rd Annual ACM Symp. Theory of Computing*, pages 31–40, 2011. `arXiv:1009.3640`.

[43] C. Lancien and A. Winter. Distinguishing multi-partite states by local measurements, 2012. `arXiv:1206.2884`.

[44] J. Littlewood. On bounded bilinear forms in an infinite number of variables. *Quarterly Journal of Mathematics*, os-1:164–174, 1930.

[45] W. Matthews, S. Wehner, and A. Winter. Distinguishability of quantum states under restricted families of measurements with an application to quantum data hiding. *Comm. Math. Phys.*, 291(3):813–843, 2009. `arXiv:0810.2327`.

[46] A. Montanaro and T. Osborne. Quantum boolean functions. *Chicago Journal of Theoretical Computer Science*, 2010. `arXiv:0810.2435`.

[47] E. Mossel, R. O'Donnell, and K. Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Ann. of Math.*, 171(1), 2010.

[48] E. Nelson. A quartic interaction in two dimensions. In *Mathematical Theory of Elementary Particles (Dedham, Massachusetts, 1965)*, pages 69–73. MIT Press, 1966.

[49] D. Nuñez-Alarcón and D. Pellegrino. On the growth of the optimal constants of the multilinear Bohnenblust-Hille inequality, 2012. `arXiv:1205.2385`.

[50] R. O'Donnell. 15-859S: Analysis of boolean functions, 2007. `http://www.cs.cmu.edu/~odonnell/boolean-analysis/`.

[51] R. O'Donnell. Some topics in analysis of boolean functions. In *Proc. 40th Annual ACM Symp. Theory of Computing*, pages 569–578, 2008.

[52] D. Pellegrino and J. Seoane-Sepúlveda. New upper bounds for the constants in the Bohnenblust-Hille inequality. *Journal of Mathematical Analysis and Applications*, 386(1):300–307, 2012. `arXiv:1010.0461v3`.

[53] D. Pérez-García, M. Wolf, C. Palazuelos, I. Villanueva, and M. Junge. Unbounded violation of tripartite Bell inequalities. *Comm. Math. Phys.*, 279(2):455–486, 2008. `quant-ph/0702189`.

[54] R. Roth and K. Viswanathan. On the hardness of decoding the gale-berlekamp code. *IEEE Trans. Inform. Theory*, 54(3):1050–1060, 2008.

[55] P. Sen. Random measurement bases, quantum state distinction and applications to the hidden subgroup problem. In *Proc. 21st Annual IEEE Conf. Computational Complexity*, page 287, 2006. `quant-ph/0512085`.

[56] D. Serrano-Rodríguez. A closed form for subexponential constants in the Bohnenblust-Hille inequality, 2012. `arXiv:1205.4735`.

[57] B. Simon and R. Høegh-Krohn. Hypercontractive semigroups and two dimensional self-coupled Bose fields. *J. Funct. Anal.*, 9(2):121–180, 1972.

[58] E. Stein and G. Weiss. *Fourier analysis on Euclidean spaces*. Princeton University Press, 1971.

[59] R. de Wolf. A brief introduction to Fourier analysis on the boolean cube. *Theory of Computing Library Graduate Surveys*, 1:1–20, 2008.