CHICAGO JOURNAL OF THEORETICAL COMPUTER SCIENCE 2010, Article 1, pages 1–45
http://cjtcs.cs.uchicago.edu/

Quantum boolean functions

Ashley Montanaro Tobias J. Osborne

Received: March 6, 2009; published: January 13, 2010.

Abstract: In this paper we introduce the study of *quantum boolean functions*, which are unitary operators f whose square is the identity: $f^2 = \mathbb{I}$. We describe several generalisations of well-known results in the theory of boolean functions, including quantum property testing; a quantum version of the Goldreich-Levin algorithm for finding the large Fourier coefficients of boolean functions; and two quantum versions of a theorem of Friedgut, Kalai and Naor on the Fourier spectra of boolean functions. In order to obtain one of these generalisations, we prove a quantum extension of the hypercontractive inequality of Bonami, Gross and Beckner.

Contents

DOI: 10.4086/cjtcs.2010.001

1 Introduction

Boolean functions stand at the crossroads between many areas of study, such as social science, combinatorics, computational complexity, and statistical mechanics, to name but a few (see, e.g., [44] or [17] for an introduction). The theory of boolean functions has reached a maturity of sorts, and the foundational results of the subject are now well established.

The advent of quantum algorithms (see, e.g., [42]) has added a twist to the computational complexity theory landscape of which boolean functions are an integral part, and there are many reasons to expect that this development will have more than superficial consequences. Indeed, quantum algorithmic techniques have already had important consequences within the field of classical complexity theory alone, leading to a simple proof of the closure of PP under intersection [1] as well as proofs of new lower bounds for locally decodable codes [34], amongst many others.

In this context it is natural to consider a generalisation of the notion of a boolean function to the quantum domain. While there is always a temptation to quantise an existing classical concept, there are other motivations for considering a quantum analogue of the theory of boolean functions. One reason is to develop lower bounds for quantum circuits. Another reason is to understand the propagation of correlations through multipartite quantum systems. Finally, and perhaps most importantly, we hope that a mature theory of quantum boolean functions will lead the way to a proof of a quantum generalisation of the PCP theorem [4]. We believe this should occur because the extant classical proofs (see, eg., [45, 18]) draw heavily on results from the property testing of boolean functions, which, as we'll see, generalise naturally to the quantum domain.

In this paper we pursue one particular generalisation of a boolean function to the quantum domain, namely a unitary operator f which squares to the identity. We provide several arguments for why our

definition is a natural generalisation of the classical definition, not least of which are several quantum generalisations of well-known classical results. These quantum generalisations often require new proof techniques, and since they reduce in each case to the classical results when f is diagonal in the computational basis, we sometimes obtain different proofs for the original classical results.

1.1 Summary of results

The main results we obtain can be summarised as follows.

- **Quantum property testing.** We give quantum tests that determine whether a unitary operator is a tensor product of Pauli operators, or far from any such tensor product; and similarly whether a unitary operator is a Pauli operator acting on only one qubit, or is far from any such operator. These are quantum generalisations of properties considered in the classical field of property testing of boolean functions. In particular, when applied to classical boolean functions, these tests have better parameters than the original classical tests.

- **Learning quantum boolean functions.** We develop a quantum analogue of the Goldreich-Levin algorithm, which is an important tool for approximately learning boolean functions [23, 38]. This algorithm allows the approximate learning of quantum dynamics, giving a natural counterpart to recent results of Aaronson on approximately learning quantum states [2].

- **Hypercontractivity and a quantum FKN theorem.** The Friedgut-Kalai-Naor (FKN) theorem [21] states that boolean functions whose Fourier transform is concentrated on the first level approximately depend on a single variable. We prove a quantum analogue of this statement. In order to obtain this result, we state and prove a quantum generalisation of the hypercontractive inequality of Bonami-Gross-Beckner [12, 24, 6] for functions $\{0,1\}^n \to \mathbb{R}$. This generalisation may be of independent interest and has several corollaries. Our result is an alternative generalisation of this inequality to that recently proven by Ben-Aroya et al [7].

- **Influences and progress towards a quantum KKL theorem.** The Kahn-Kalai-Linial (KKL) theorem [30] states that every balanced boolean function must have a variable with high influence (qv.). Defining a suitable quantum generalisation of the concept of influence, we prove the generalised theorem in several special cases, and conjecture that it holds in full generality. We also prove a weaker variant (a quantum Poincaré inequality).

Our presentation is based on the lecture notes [44], which are an excellent introduction to the field of the analysis of boolean functions.

1.2 Related work

This paper draws heavily on the classical field of the analysis of boolean functions, which for our purposes essentially began with the seminal paper of Kahn, Kalai and Linial [30], which proved that every balanced boolean function must have an influential variable (see Section 10). Since then, a substantial literature has developed, in which statements of interest about boolean functions are proven using mathematical

techniques of increasing sophistication, and in particular Fourier analysis. Other important works in this area include a result of Bourgain concerning the Fourier spectrum of boolean functions [13], and a result of Friedgut stating that boolean functions which are insensitive on average are close to depending on a small number of variables [20].

Ideas relating to the analysis of boolean functions have recently proven fruitful in the study of quantum computation. Indeed, the result of Bernstein and Vazirani giving the first super-polynomial separation between quantum and classical computation [8] is that quantum computers can distinguish certain boolean functions more efficiently than classical computers can. More recent cases where a quantum advantage is found for classical tasks relating to boolean functions include the quantum Goldreich-Levin algorithm of Adcock and Cleve [3]; the work of Buhrman et al [15] on quantum property testing of classical boolean functions; and the computational learning algorithms of Bshouty and Jackson [14], and also Atici and Servedio [5].

Fourier analysis of boolean functions has been used explicitly to obtain two recent results in quantum computation. The first is an exponential separation between quantum and classical one-way communication complexity proven by Gavinsky et al [22]. This separation uses the results of Kahn, Kalai and Linial [30] to lower bound the classical communication complexity of a particular partial function. The second result is a lower bound on the size of quantum random access codes obtained by Ben-Aroya et al [7], in which the key technical ingredient is the proof of a matrix-valued extension of the hypercontractive inequality on which [30] is based.

2 Preliminaries

In this section we set up our notation and describe the objects we'll work with in the sequel. We will sometimes use several notations for the same objects. We've made this decision for two reasons. The first is that the notations natural in the study of boolean functions are unnatural in quantum mechanics, and vice versa, and some fluidity with the notation greatly simplifies the transition between these two domains. Secondly, two notations means that practitioners in classical complexity theory and quantum mechanics can (hopefully) adapt to the tools of the other field.

In the classical domain we work with *boolean functions* of n variables, which are simply functions $f : \{0,1\}^n \to \{0,1\}$. We write elements of $\{0,1\}$ as regular letters, eg. y, and we write elements of $\{0,1\}^n$ – *strings* – as boldface letters, eg., $\mathbf{x} \equiv x_1 x_2 \cdots x_n$, $x_j \in \{0,1\}$, $j = 1,2,\ldots,n$. Similarly, we'll write strings in $\{0,1,2,3\}^n$ as boldface letters starting at $\mathbf{s}$. It is often convenient to exploit the isomorphism between the set $\{0,1\}^n$ of all strings and the set $\mathcal{P}([n])$ of all subsets of $[n] \equiv \{1,2,\ldots,n\}$ by letting $\mathbf{x}$ define the subset $S = \{j \in S \,|\, x_j \neq 0\}$. We write subsets of $[n]$ as capital letters beginning with S.

There are several elementary functions on $\{0,1\}^n$ and $\{0,1,2,3\}^n$ that will be useful in the sequel. Firstly, we define the *support* of a string $\mathbf{x} \in \{0,1\}^n$, written $\mathrm{supp}(\mathbf{x})$, via

$$\mathrm{supp}(\mathbf{x}) \equiv \{j \,|\, x_j \neq 0\}, \tag{2.1}$$

and for $\mathbf{s} \in \{0,1,2,3\}^n$ via

$$\mathrm{supp}(\mathbf{s}) \equiv \{j \,|\, s_j \neq 0\}. \tag{2.2}$$

We define the *Hamming weight* $|\mathbf{x}|$ of a string $\mathbf{x} \in \{0,1\}^n$ (respectively, $\mathbf{s} \in \{0,1,2,3\}^n$) via $|\mathbf{x}| \equiv |\mathrm{supp}(\mathbf{x})|$ (respectively, $|\mathbf{s}| \equiv |\mathrm{supp}(\mathbf{s})|$). The *intersection* of two strings $\mathbf{s}$ and $\mathbf{t}$, written $\mathbf{s} \cap \mathbf{t}$, is the set $\{i : s_i \neq 0, t_i \neq 0\}$.

We now make the notational switch to identifying $\{0,1\}$ with $\{+1,-1\}$ via $0 \equiv 1$ and $1 \equiv -1$. This takes a little getting used to, but is standard in the boolean function literature; it's often abundantly clear from the context which notation is being used. Unless otherwise noted, we'll use the $\{+1,-1\}$ notation. We identify $\{0,1\}$ with $\mathbb{Z}/2\mathbb{Z}$, with addition written $x \oplus y$. We identify $\{+1,-1\}$ with the multiplicative group of two elements, also isomorphic to $\mathbb{Z}/2\mathbb{Z}$, and write products as xy. This second identification still allows us to identify elements of the multiplicative group of two elements with elements of $\mathbb{Z}$ and to use *addition* defined as in $\mathbb{Z}$. Thus we say that a boolean function is *balanced* if $\sum_{\mathbf{x}} f(\mathbf{x}) = 0$ (here we've made the notational switch).

We denote by χ_S the *linear function* on subset S, defined by

$$\chi_S \equiv \prod_{j \in S} x_j. \tag{2.3}$$

Exploiting the connection between strings and subsets of $[n]$ allows us to write this as

$$\chi_S(\mathbf{x}) \equiv \chi_S(T) \equiv (-1)^{|S \cap T|}, \tag{2.4}$$

where T is the set defined by $\mathbf{x}$.

In the quantum domain we work with the Hilbert space of n qubits. This is the Hilbert space $\mathcal{H} \equiv (\mathbb{C}^2)^{\otimes n}$. We write elements of $\mathcal{H}$ as *kets* $|\psi\rangle$, and the inner product between two kets $|\phi\rangle$ and $|\psi\rangle$ is written $\langle\phi|\psi\rangle$. There is a distinguished basis for $\mathcal{H}$, called the *computational basis*, written $|\mathbf{x}\rangle$, $\mathbf{x} \in \{0,1\}^n$, with inner product $\langle \mathbf{x}|\mathbf{y}\rangle = \delta_{x_1,y_1}\delta_{x_2,y_2}\cdots\delta_{x_n,y_n}$. We'll also refer to another Hilbert space, namely the Hilbert space $\mathcal{B}(\mathcal{H})$ of (bounded) operators on $\mathcal{H}$. The inner product on $\mathcal{B}(\mathcal{H})$ is given by the *Hilbert-Schmidt* inner product $\langle M,N\rangle \equiv \frac{1}{2^n}\mathrm{tr}(M^\dagger N)$, $M,N \in \mathcal{B}(\mathcal{H})$.

We define the (normalised Schatten) p-norm of a d-dimensional operator f in terms of the singular values $\{s_j(f)\}$ of f as

$$\|f\|_p \equiv \left(\frac{1}{d}\sum_{j=1}^{d} s_j(f)^p\right)^{\frac{1}{p}}. \tag{2.5}$$

If we define $|f| \equiv \sqrt{f^\dagger f}$ (note the overload of notation) we can write the p-norm as

$$\|f\|_p \equiv \left(\frac{1}{d}\mathrm{tr}(|f|^p)\right)^{\frac{1}{p}}. \tag{2.6}$$

Note the useful equalities $\|f^p\|_q = \|f\|_{pq}^p$ and $\|f\|_q^p = \|f^p\|_{q/p}$. One unfortunate side effect of this normalisation is that $\|f\|_p$ is not a submultiplicative matrix norm (except at $p = \infty$). However, we do still have Hölder's inequality [9]: for $1/p + 1/q = 1$,

$$|\langle f,g\rangle| \leq \|f\|_p \|g\|_q. \tag{2.7}$$

3 Quantum boolean functions

The first question we attempt to answer is: what is the right quantum generalisation of a classical boolean function? We would expect the concept of a *quantum boolean function* to at least satisfy the following properties.

1. A quantum boolean function should be a unitary operator, so it can be implemented on a quantum computer.
2. Every classical boolean function should give rise to a quantum boolean function in a natural way.
3. Concepts from classical analysis of boolean functions should have natural quantum analogues.

Happily, we can achieve these requirements with the following definition.

Definition 3.1. A quantum boolean function of n qubits is a unitary operator f on n qubits such that $f^2 = \mathbb{I}$.

We note that it is immediate that f is Hermitian (indeed, all f's eigenvalues are ± 1), and that every unitary Hermitian operator is quantum boolean. Turning to the task of demonstrating the second desired property, there are at least two natural ways of implementing a classical boolean function $f : \{0,1\}^n \to \{0,1\}$ on a quantum computer. These are as the so-called *bit oracle* [32]:

$$|\mathbf{x}\rangle|y\rangle \mapsto U_f|\mathbf{x}\rangle|y\rangle \equiv |\mathbf{x}\rangle|y \oplus f(\mathbf{x})\rangle, \tag{3.1}$$

and as what's known as the *phase oracle*:

$$|\mathbf{x}\rangle \mapsto (-1)^{f(\mathbf{x})}|\mathbf{x}\rangle. \tag{3.2}$$

Making the previously discussed notational switch allows us to write the second action as

$$|\mathbf{x}\rangle \mapsto f(\mathbf{x})|\mathbf{x}\rangle. \tag{3.3}$$

When f acts as a bit oracle it defines a unitary operator U_f. When f acts as a phase oracle, we also obtain a unitary operator, which we simply write as f. As a sanity check of Definition 3.1 we have that these operators square to the identity: $U_f^2 = \mathbb{I}$ and $f^2 = \mathbb{I}$, and so are examples of quantum boolean functions. It turns out that these two actions are almost equivalent: the phase oracle can be obtained from one use of the bit oracle and a one-qubit ancilla, via

$$f(\mathbf{x})|\mathbf{x}\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \equiv U_f|\mathbf{x}\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle), \tag{3.4}$$

while the bit oracle can be similarly recovered, given access to a controlled-phase oracle

$$|\mathbf{x}\rangle|y\rangle \mapsto (-1)^{y \cdot f(\mathbf{x})}|\mathbf{x}\rangle|y\rangle, \tag{3.5}$$

where we introduce a one-qubit control register $|y\rangle$.

Having accepted Definition 3.1 for a quantum boolean function we should at least demonstrate that there is something quantum about it. This can be easily accomplished by noting that the theory of quantum boolean functions is at least as general as that of quantum error correction and quantum algorithms for decision problems. Firstly, any quantum error correcting code is a subspace P of $\mathcal{H}$. Using this code we define a quantum boolean function via $f = \mathbb{I} - 2P$. Secondly, note that any quantum algorithm for a decision problem naturally gives rise to a quantum boolean function in the *Heisenberg picture*: if U is the quantum circuit in question, and the answer (either 0 or 1, or some superposition thereof) is placed in some output qubit q, then measuring the observable $A = |0\rangle_q\langle 0| - |1\rangle_q\langle 1|$ (with an implied action as the identity on the remaining qubits) will read out the answer. However, this is equivalent to measuring the observable $f = U^\dagger AU$ on the initial state. It is easy to verify that f is quantum boolean.

4 Examples of quantum boolean functions

In this section we introduce several different examples of quantum boolean functions. We have already met our first such examples, namely quantisations of classical boolean functions $f(\mathbf{x})$. Our next example quantum boolean functions are aimed at generalising various classes of classical boolean functions.

First, we note the obvious fact that even in the single qubit case, there are infinitely many quantum boolean functions. By contrast, there are only four one-bit classical boolean functions.

Example 4.1. For any real θ, the matrix

$$\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \tag{4.1}$$

is a quantum boolean function.

Single qubit quantum boolean functions can of course be combined to form n qubit quantum boolean functions, as follows.

Definition 4.2. Let f be a quantum boolean function. Then we say that f is *local* if it can be written as

$$f = U_1 \otimes U_2 \otimes \cdots \otimes U_n, \tag{4.2}$$

where U_j is a 2×2 matrix satisfying $U_j^2 = \mathbb{I}$, $\forall j \in [n]$.

Local quantum boolean functions are the natural generalisation of linear boolean functions (indeed, every linear boolean function is a local quantum boolean function). One might reasonably argue that local quantum boolean functions aren't really quantum: after all, there exists a local rotation which diagonalises such an operator and reduces it to a classical linear boolean function. The next example illustrates that the generic situation is probably very far from this.

Example 4.3. Let $P^2 = P$ be a projector. Then $f = \mathbb{I} - 2P$ is a quantum boolean function. In particular, if $|\psi\rangle$ is an arbitrary quantum state, then $f = \mathbb{I} - 2|\psi\rangle\langle\psi|$ is a quantum boolean function.

If we set $|\psi\rangle = \frac{1}{\sqrt{2}}(|\mathbf{0}\rangle + |\mathbf{1}\rangle)$ then we see that there is no local rotation $V_1 \otimes V_2 \otimes \cdots \otimes V_n$ which diagonalises $f = \mathbb{I} - 2|\psi\rangle\langle\psi|$.

4.1 New quantum boolean functions from old

There are several ways to obtain new quantum boolean functions from existing quantum boolean functions. We summarise these constructions below.

Lemma 4.4. *Let f be a quantum boolean function and U be a unitary operator. Then $U^\dagger f U$ is a quantum boolean function.*

Lemma 4.5. *Let h be a Hermitian operator. Then $f = \mathrm{sgn}(h)$ is a quantum boolean function, where*

$$\mathrm{sgn}(x) = \begin{cases} 1, & x > 0 \\ -1, & x \leq 0, \end{cases} \tag{4.3}$$

and as usual the notation $f = g(h)$, for a Hermitian operator h, means the operator obtained by applying the function $g : \mathbb{R} \to \mathbb{R}$ to the eigenvalues of h.

Lemma 4.6. *Let $\alpha_j \in \mathbb{R}$, $j \in [m]$, satisfy $\sum_{j=1}^m \alpha_j^2 = 1$ and $\{f_j\}$, $j \in [m]$, be a set of* anticommuting *quantum boolean functions, i.e.*

$$\{f_j, f_k\} \equiv f_j f_k + f_k f_j = 2\delta_{jk}\mathbb{I}. \tag{4.4}$$

Then

$$f = \sum_{j=1}^m \alpha_j f_j \tag{4.5}$$

is a quantum boolean function.

Proof. Squaring f gives

$$f^2 = \sum_{j=1}^m \alpha_j^2 f_j^2 + \sum_{j<k} \alpha_j \alpha_k \{f_j, f_k\} = \mathbb{I}. \tag{4.6}$$

That f is Hermitian follows because it is a real-linear combination of Hermitian operators. □

5 Fourier analysis

It is well-known that every function $f : \{0,1\}^n \to \mathbb{R}$ can be expanded in terms of the characters of the group $(\mathbb{Z}/2\mathbb{Z})^n$. These characters are given by the set of linear functions $\chi_S(T) = (-1)^{|S \cap T|}$ (we are identifying input strings $\mathbf{x}$ with the subset T). This expansion is called the Fourier transform over $(\mathbb{Z}/2\mathbb{Z})^n$.

The use of Fourier analysis in the study of boolean functions was pioneered by Kahn, Kalai, and Linial, who were responsible for the eponymous KKL theorem [30], and has facilitated many of the core foundational results relating to boolean functions. We seek an analogous expansion for quantum boolean functions.

Our quantum analogues of the characters of $\mathbb{Z}/2\mathbb{Z}$ will be the Pauli matrices $\{\sigma^0, \sigma^1, \sigma^2, \sigma^3\}$[1]. These are defined as

$$\sigma^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \text{ and } \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.1}$$

It is clear that the Pauli operators are quantum boolean functions. A tensor product of Paulis (also known as a *stabilizer operator*) is written as $\sigma^{\mathbf{s}} \equiv \sigma^{s_1} \otimes \sigma^{s_2} \otimes \cdots \otimes \sigma^{s_n}$, where $s_j \in \{0,1,2,3\}$. We use the notation σ_i^j for the operator which acts as σ^j at the i'th position, and trivially elsewhere.

Where convenient we'll use one of two other different notations to refer to the operators $\sigma^{\mathbf{s}}$, namely as $\chi_{\mathbf{s}}$ and as $|\mathbf{s}\rangle$. We use the first of these alternate notations to bring out the parallels between the classical theory of boolean functions and the quantum theory, and the second to emphasise the Hilbert space structure of $\mathcal{B}(\mathcal{H})$. We write the inner products in each case as

$$\langle \mathbf{s}|\mathbf{t}\rangle \equiv \frac{1}{2^n} \operatorname{tr}((\sigma^{\mathbf{s}})^\dagger \sigma^{\mathbf{t}}) \equiv \langle \chi_{\mathbf{s}}, \chi_{\mathbf{t}} \rangle. \tag{5.2}$$

The set of stabilizer operators is orthonormal with respect to the Hilbert-Schmidt inner product, and thus forms a basis for the vector space $\mathcal{B}(\mathcal{H})$. So we can express any (bounded) operator f on n qubits in terms of stabilizer operators, with the explicit expansion

$$f = \sum_{\mathbf{s} \in \{0,1,2,3\}^n} \hat{f}_{\mathbf{s}} \chi_{\mathbf{s}}, \tag{5.3}$$

where $\hat{f}_{\mathbf{s}} = \langle \chi_{\mathbf{s}}, f \rangle$. In an abuse of terminology, we call the set $\{\hat{f}_{\mathbf{s}}\}$ indexed by $\mathbf{s} \in \{0,1,2,3\}^n$ the *Fourier coefficients* of f. Note that, if f is Hermitian, these coefficients are all real. This expansion is well-known in quantum information theory and, for example, was recently used by Kempe et al [33] to give upper bounds on fault-tolerance thresholds.

We extend the definition of support to operators via

$$\operatorname{supp}(f) \equiv \bigcup_{\mathbf{s} \,|\, \hat{f}_{\mathbf{s}} \neq 0} \operatorname{supp}(\mathbf{s}). \tag{5.4}$$

Similarly, we define the *weight* $\operatorname{wgt}(M)$ of an operator M via $\operatorname{wgt}(M) \equiv |\operatorname{supp}(M)|$.

Definition 5.1. Let f be a quantum boolean function. If $|\operatorname{supp}(f)| = k$ then we say that f is a *k-junta*. If $k = 1$ then we say that f is a *dictator*. If $k = 0$ then we say that f is *constant*.

We note that while, classically, there is a distinction between dictators (functions $f : \{0,1\}^n \to \{0,1\}$ such that $f(\mathbf{x}) = x_i$ for some i) and so-called *anti-dictators* (functions where $f(\mathbf{x}) = 1 - x_i$ for some i), there is no such distinction in our terminology here.

Our definition of $\{\hat{f}_{\mathbf{s}}\}$ as the Fourier coefficients for an operator f is no accident. Indeed, it turns out that when one expands a *boolean function* f (represented as a phase oracle) in terms of $\chi_{\mathbf{s}}$ then we recover the classical Fourier transform of f.

[1] The Pauli matrices are often written as $\sigma^0 \equiv \mathbb{I}$, $\sigma^1 \equiv \sigma^x$, $\sigma^2 \equiv \sigma^y$, and $\sigma^3 \equiv \sigma^z$.

Proposition 5.2. *Let f be a boolean function $f : \{0,1\}^n \to \{1,-1\}$. Then if f acts as $f|\mathbf{x}\rangle = f(\mathbf{x})|\mathbf{x}\rangle$, and if*

$$f = \sum_{\mathbf{s}} \hat{f}_{\mathbf{s}} \chi_{\mathbf{s}}, \tag{5.5}$$

then the set $\{\hat{f}_{\mathbf{s}}\}$ are given by

$$\hat{f}_{\mathbf{s}} = \begin{cases} 0, & \mathbf{s} \notin \{0,3\}^n, \\ \frac{1}{2^n} \sum_{T \subset [n]} (-1)^{|S \cap T|} f(T), & \mathbf{s} \in \{0,3\}^n, \end{cases} \tag{5.6}$$

where $S = \mathrm{supp}(\mathbf{s})$ and T is the support of the string $\mathbf{t} = t_1 t_2 \cdots t_n$, where $t_j = 0$ if $j \notin T$ and $t_j = 3$ if $j \in T$.

Proof. Because $f(\mathbf{x})$ is diagonal in the computational basis we immediately learn that $\langle \chi_{\mathbf{s}}, f \rangle = 0$ for any $\mathbf{s} \notin \{0,3\}^n$. When $\mathbf{s} \in \{0,3\}^n$ we have that

$$\begin{aligned} \hat{f}_{\mathbf{s}} = \langle \chi_{\mathbf{s}}, f \rangle &= \frac{1}{2^n} \mathrm{tr} \left(f \prod_{j \in \mathrm{supp}(\mathbf{s})} \sigma_j^3 \right) \\ &= \frac{1}{2^n} \sum_{\mathbf{x}} \left(\prod_{j \in \mathrm{supp}(\mathbf{s})} (-1)^{x_j} \right) f(\mathbf{x}) = \frac{1}{2^n} \sum_{T \subset [n]} (-1)^{|S \cap T|} f(T). \end{aligned} \tag{5.7}$$

□

We also immediately have the quantum analogues of Plancherel's theorem and Parseval's equality.

Proposition 5.3. *Let f and g be operators on n qubits. Then $\langle f, g \rangle = \sum_{\mathbf{s}} \hat{f}_{\mathbf{s}}^* \hat{g}_{\mathbf{s}}$. Moreover, $\|f\|_2^2 = \sum_{\mathbf{s}} |\hat{f}_{\mathbf{s}}|^2$. Thus, if f is quantum boolean, $\sum_{\mathbf{s}} \hat{f}_{\mathbf{s}}^2 = 1$.*

It is often convenient to decompose the Fourier expansion of an operator into different levels. An arbitrary n-qubit operator f can be expanded as

$$f = \sum_{\mathbf{s}} \hat{f}_{\mathbf{s}} \chi_{\mathbf{s}} = \sum_{k=0}^{n} f^{=k}, \tag{5.8}$$

where

$$f^{=k} \equiv \sum_{\mathbf{s}, |\mathbf{s}|=k} \hat{f}_{\mathbf{s}} \chi_{\mathbf{s}}. \tag{5.9}$$

(One can define $f^{<k}$, etc., similarly.) The *weight* of f at level k is then defined as $\|f^{=k}\|_2^2$. A natural measure of the complexity of f is its *degree*, which is defined as

$$\deg(f) \equiv \max_{\mathbf{s}, \hat{f}_{\mathbf{s}} \neq 0} |\mathbf{s}|. \tag{5.10}$$

We pause to note an important difference between quantum and classical boolean functions. In the classical case, it is easy to show that every non-zero Fourier coefficient of a boolean function on n bits is

at least 2^{1-n} in absolute value. However, this does not hold for quantum boolean functions. Consider the operator

$$f = \varepsilon\,\sigma^1 \otimes \sigma^1 + \sqrt{1-\varepsilon^2}\,\sigma^2 \otimes \mathbb{I}. \tag{5.11}$$

By Lemma 4.6, this is a quantum boolean function for any $0 \leq \varepsilon \leq 1$. Taking $\varepsilon \to 0$, we see that the coefficient of $\sigma^1 \otimes \sigma^1$ may be arbitrarily small.

6 Testing quantum boolean functions

The field of classical *property testing* solves problems of the following kind. Given access to a boolean function f that is promised to either have some property, or to be "far" from having some property, determine which is the case, using a small number of queries to f. Property testers are an important component of many results in classical computer science, e.g. [25, 18]; see the review [19] for an introduction to the area.

In this section we describe property testing for *quantum* boolean functions: we give quantum algorithms which determine whether a unitary operator, implemented as an oracle, has the property of being, variously, a stabilizer operator, or both a stabilizer operator and a dictator. These tests differ substantially from their classical counterparts and typically require fewer queries. However, as with their classical equivalents, we use Fourier analysis to bound their probabilities of success.

We note that Buhrman et al [15] have already shown that quantum computers can obtain an advantage over classical computers for the property testing of *classical* boolean functions.

6.1 Closeness

The tests that we define will, informally, output either that a unitary operator has some property, or is "far" from any operator with that property. In order to define the concept of property testing of quantum boolean functions, we thus need to define what it means for two operators to be close.

Definition 6.1. Let f and g be two operators. Then we say that f and g are ε-close if $\|f-g\|_2^2 \leq 4\varepsilon$.

In quantum theory, it is often natural to use the infinity, or *sup*, norm to measure closeness of operators (i.e., the magnitude of the largest eigenvalue). However, the 2-norm seems more intuitive when dealing with boolean functions; for example, if we produce a pair of quantum boolean functions f, g from *any* classical boolean functions that differ at any position, then $\|f-g\|_\infty = 2$. Intuitively, the 2-norm tells us how the function behaves *on average*, and the infinity norm tells us about the *worst case* behaviour. There is also the following relationship to unitary operator discrimination.

Proposition 6.2. *Given a unitary operator f promised to be one of two unitary operators f_1, f_2, there is a procedure that determines whether $f = f_1$ or $f = f_2$ with one use of f and success probability*

$$\frac{1}{2} + \frac{1}{2}\sqrt{1 - |\langle f_1, f_2\rangle|^2}. \tag{6.1}$$

Proof. The proof rests on the fundamental result of Holevo and Helstrom [29, 28] which says that the exact minimum probability of error that can be achieved when discriminating between two pure states $|\psi_1\rangle$ and $|\psi_2\rangle$ with a priori probabilities p and $1-p$ is given by

$$\mathbb{P}[\text{test succeeds}] = \frac{1}{2} + \frac{1}{2}\sqrt{1-4p(1-p)|\langle\psi_1|\psi_2\rangle|^2}. \tag{6.2}$$

We now apply f_1 and f_2 to halves of two maximally entangled states $|\Phi\rangle \equiv \frac{1}{2^{n/2}}\sum_{\mathbf{x}}|\mathbf{x}\rangle|\mathbf{x}\rangle$. This produces the states

$$|f_1\rangle \equiv f_1 \otimes \mathbb{I}|\Phi\rangle, \quad \text{and} \quad |f_2\rangle \equiv f_2 \otimes \mathbb{I}|\Phi\rangle. \tag{6.3}$$

The overlap between these two states can be calculated as follows:

$$\begin{aligned}\langle f_1|f_2\rangle &= \frac{1}{2^n}\sum_{\mathbf{x},\mathbf{y}}\langle\mathbf{y}|\langle\mathbf{y}|(f_1^\dagger f_2)\otimes\mathbb{I}|\mathbf{x}\rangle|\mathbf{x}\rangle \\ &= \frac{1}{2^n}\sum_{\mathbf{x},\mathbf{y}}\langle\mathbf{y}|f_1^\dagger f_2|\mathbf{x}\rangle\langle\mathbf{y}|\mathbf{x}\rangle \\ &= \frac{1}{2^n}\sum_{\mathbf{x}}\langle\mathbf{x}|f_1^\dagger f_2|\mathbf{x}\rangle = \frac{1}{2^n}\operatorname{tr}(f_1^\dagger f_2) = \langle f_1, f_2\rangle.\end{aligned} \tag{6.4}$$

The lemma follows when we apply the Holevo-Helstrom result to $|f_1\rangle$ and $|f_2\rangle$ and minimise over p. □

6.2 The quantum stabilizer test

In this subsection we describe a quantum test, the *quantum stabilizer test*, which decides, using only two queries, whether a unitary operator f is either ε-close to a stabilizer operator $\chi_{\mathbf{s}}$ up to a phase, or is far from any such operator. We also describe a test which is conjectured to decide whether a unitary operator is ε-close to local or not.

The idea behind our tests is very simple: suppose f is a unitary operator, and we want to work out if it is a stabilizer, i.e. if $f = \chi_{\mathbf{s}}$ for some $\mathbf{s}$. One way to do this is to apply f to the halves of n maximally entangled states resulting in a quantum state $f\otimes\mathbb{I}|\Phi\rangle$. If f is local then the result will just be a tensor product of n (possibly rotated) maximally entangled states, and if f is a stabilizer then it should be an n-fold product of one of four possible states. If not, then there will be entanglement between the n subsystems. The way to test this hypothesis is to create another identical state $f\otimes\mathbb{I}|\Phi\rangle$ by again applying f to another set of n maximally entangled states and separately apply an equality test to each of the n subsystems which are meant to be disentangled from each other.

Definition 6.3. Let f be a unitary operator on n qubits. The quantum stabilizer and locality tests proceed as follows.

1. Prepare $4n$ quantum registers in the state

$$|\Phi\rangle_{\mathbf{AA'}}|\Phi\rangle_{\mathbf{BB'}} \equiv \frac{1}{2^n}\sum_{\mathbf{x},\mathbf{y}}|\mathbf{x}\rangle_{\mathbf{A}}|\mathbf{x}\rangle_{\mathbf{A'}}|\mathbf{y}\rangle_{\mathbf{B}}|\mathbf{y}\rangle_{\mathbf{B'}} = |\phi\rangle_{\mathbf{AA'}}^{\otimes n}\otimes|\phi\rangle_{\mathbf{BB'}}^{\otimes n}, \tag{6.5}$$

where $|\phi\rangle \equiv \frac{1}{\sqrt{2}}\sum_{x\in\{\pm 1\}}|xx\rangle$, and $\mathbf{A} = A_1A_2\cdots A_n$, $\mathbf{A}' = A'_1A'_2\cdots A'_n$, $\mathbf{B} = B_1B_2\cdots B_n$, and $\mathbf{B}' = B'_1B'_2\cdots B'_n$.

2. Apply f to $\mathbf{A}$ and once more to $\mathbf{B}$ to give

$$|f\rangle|f\rangle = f_{\mathbf{A}}\otimes\mathbb{I}_{\mathbf{A}'}\otimes f_{\mathbf{B}}\otimes\mathbb{I}_{\mathbf{B}'}|\Phi\rangle|\Phi\rangle. \tag{6.6}$$

3. To test if f is a stabilizer measure the *equality* observable[2]:

$$\text{EQ} = \left(\sum_{s=0}^{3}|s\rangle_{AA'}\langle s|\otimes|s\rangle_{BB'}\langle s|\right)^{\otimes n}, \tag{6.7}$$

where

$$|s\rangle \equiv \chi_s\otimes\mathbb{I}|\phi\rangle, \quad s\in\{0,1,2,3\}, \tag{6.8}$$

giving

$$\mathbb{P}[\text{test accepts}] = \langle f|\langle f|\text{EQ}|f\rangle|f\rangle. \tag{6.9}$$

4. To test locality measure the *swap observable*

$$\text{SW} = \frac{1}{2^n}\left(\mathbb{I}_{AA':BB'} + \text{SWAP}_{AA':BB'}\right)^{\otimes n}, \tag{6.10}$$

where $\text{SWAP}_{X:Y}$ is the operator that swaps the two subsystems X and Y. This gives

$$\mathbb{P}[\text{test accepts}] = \langle f|\langle f|\text{SW}|f\rangle|f\rangle. \tag{6.11}$$

We now prove the following

Proposition 6.4. *Suppose that a unitary operator f passes the quantum stabilizer test with probability $1-\varepsilon$, where $\varepsilon < 1/2$. Then f is ε-close to an operator $e^{i\phi}\chi_{\mathbf{s}}$, $\mathbf{s}\in\{0,1,2,3\}^n$, for some phase ϕ. If f is a stabilizer operator then it passes the stabilizer test with probability 1.*

Proof. Expand f in the Fourier basis as

$$f = \sum_{\mathbf{s}}\hat{f}_{\mathbf{s}}\chi_{\mathbf{s}}. \tag{6.12}$$

Noting that

$$\langle f|\langle f|\text{EQ}|f\rangle|f\rangle = \sum_{\mathbf{s},\mathbf{t}}|\hat{f}_{\mathbf{s}}|^2|\hat{f}_{\mathbf{t}}|^2\langle\mathbf{tt}|\mathbf{ss}\rangle,$$

we see that

$$\mathbb{P}[\text{test accepts}] = \sum_{\mathbf{s}}|\hat{f}_{\mathbf{s}}|^4. \tag{6.13}$$

[2]This is not quite the same as measuring both subsystems and checking if the result is equal, as that would destroy coherent superpositions like $1/\sqrt{2}(|00\rangle+|11\rangle)$ which would otherwise be left undisturbed by measurement of EQ.

Now, thanks to Parseval's relation, we have that

$$1 = \sum_{\mathbf{s}} |\hat{f}_{\mathbf{s}}|^2, \tag{6.14}$$

and, given that the test passes with probability $1-\varepsilon$, we thus have

$$1-\varepsilon \leq \sum_{\mathbf{s}} |\hat{f}_{\mathbf{s}}|^4 \leq \left(\max_{\mathbf{s}} |\hat{f}_{\mathbf{s}}|^2\right) \sum_{\mathbf{s}} |\hat{f}_{\mathbf{s}}|^2 = \max_{\mathbf{s}} |\hat{f}_{\mathbf{s}}|^2. \tag{6.15}$$

So, according to Parseval, there is exactly one term $|\hat{f}_{\mathbf{s}}|^2$ in the expansion (6.12) which is at least $1-\varepsilon$, and the rest are each at most ε. Thus f is ε-close to $e^{i\phi}\chi_{\mathbf{s}}$ for some phase ϕ (we have that $|\langle f, \chi_{\mathbf{s}}\rangle| \geq \sqrt{1-\varepsilon}$). □

Remark 6.5. The stabilizer test is a quantum generalisation of the classical *linearity test* of Blum, Luby, and Rubenfeld [10] (the BLR test). When interpreted in quantum language the BLR test can be seen as a method to test if a quantum boolean function diagonal in the computational basis is close to a tensor product of σ^3s. (A task for which the stabilizer test can also be applied.) It is notable that the BLR test requires *three* queries to f in order to achieve the same success probability as its quantum counterpart, which only requires *two* queries. Thus, the stabilizer test can be said to have better parameters than its classical counterpart.

Now we turn to the quantum locality test.

Proposition 6.6. *The probability that the quantum locality test accepts when applied to an operator f is equal to*

$$\frac{1}{2^n}\left(\sum_{S\subset[n]} \mathrm{tr}(\rho_S^2)\right), \tag{6.16}$$

where ρ_S is the partial trace of $|f\rangle\langle f|$ over all subsystems $A_j A'_j$ with $j \notin S$, and we define $\mathrm{tr}(\rho_\emptyset^2) = 1$.

Proof. The proof proceeds via direct calculation:

$$\begin{aligned}
\mathbb{P}[\text{test accepts}] &= \langle f|\langle f| \left(\frac{\mathbb{I}_{AA':BB'} + \mathrm{SWAP}_{AA':BB'}}{2}\right)^{\otimes n} |f\rangle|f\rangle \\
&= \frac{1}{2^n} \sum_{S\subset[n]} \langle f|\langle f| \mathrm{SWAP}_{A_S A'_S : B_S B'_S} |f\rangle|f\rangle \\
&= \frac{1}{2^n} \sum_{S\subset[n]} \mathrm{tr}\left(\mathrm{SWAP}_{A_S A'_S : B_S B'_S} \rho_S \otimes \rho_S\right) \\
&= \frac{1}{2^n} \sum_{S\subset[n]} \mathrm{tr}(\rho_S^2),
\end{aligned} \tag{6.17}$$

where we use the notation A_S for an operator which is applied to the subsystems S, and acts as the identity elsewhere. □

It is easy to see that if f is local then the probability the quantum locality test accepts is equal to 1. We have been unable to show that if the test accepts with probability greater than $1-\varepsilon$ then f is close to being local. Thus we have the following

Conjecture 6.7. *Suppose f passes the quantum locality test with probability $\geq 1-\varepsilon$. Then there exist U_j, $j \in [n]$, with $U_j^2 = \mathbb{I}$ such that*

$$\langle f, U_1 \otimes U_2 \otimes \cdots \otimes U_n \rangle \geq 1 - 2\varepsilon. \tag{6.18}$$

6.3 Testing dictators

In this subsection we describe a quantum test — a quantisation of the Håstad test [25] — which tests whether a unitary operator f is ε-close to a dictator. In fact, we give two such tests. The first determines whether f is close, up to a phase, to a dictator which is also a stabilizer operator (a *stabilizer dictator*). The second is intended to determine whether f is close to a dictator. As with the situation for quantum locality testing, we are able to analyse the first test, but leave the second as a conjecture.

The dictator — or *quantum Håstad* — test is defined as follows.

Definition 6.8. Let f be a unitary operator and let $0 \leq \delta \leq 1$. Then the quantum Håstad test proceeds as follows.

1. Prepare $4n$ quantum registers in the state

$$|\Phi\rangle_{\mathbf{AA'}} |\Phi\rangle_{\mathbf{BB'}}. \tag{6.19}$$

   (Our notation is identical to that of the quantum locality test.)

2. Apply f to $\mathbf{A}$ and once more to $\mathbf{B}$ to give

$$|f\rangle|f\rangle = f_{\mathbf{A}} \otimes \mathbb{I}_{\mathbf{A'}} \otimes f_{\mathbf{B}} \otimes \mathbb{I}_{\mathbf{B'}} |\Phi\rangle|\Phi\rangle. \tag{6.20}$$

3. To test if f is close to a stabilizer dictator measure the δ-*equality* POVM given by the operators $\{\mathrm{EQ}_\delta, \mathbb{I} - \mathrm{EQ}_\delta\}$, where

$$\mathrm{EQ}_\delta = \left(|00\rangle\langle 00| + (1-\delta) \sum_{s=1}^{3} |s\rangle_{AA'}\langle s| \otimes |s\rangle_{BB'}\langle s| \right)^{\otimes n}. \tag{6.21}$$

   This measurement is easy to implement by flipping a δ-biased coin, and gives

$$\mathbb{P}[\text{test accepts}] = \langle f|\langle f|\mathrm{EQ}_\delta|f\rangle|f\rangle. \tag{6.22}$$

4. To test if f is a dictator measure the δ-*swap observable*

$$\mathrm{SW}_\delta = \frac{1}{2^n} \left(T(\delta)_{AA':BB'} + \sqrt{T(\delta)}\mathrm{SWAP}\sqrt{T(\delta)}_{AA':BB'} \right)^{\otimes n}, \tag{6.23}$$

where

$$T(\delta) = \sum_{s,t} (1-\delta)^{\frac{|s|+|t|}{2}} |s,t\rangle\langle s,t| \tag{6.24}$$

(recall that $|s| = 1$ if $s > 0$, for $s \in \{0,1,2,3\}$), giving

$$\mathbb{P}[\text{test accepts}] = \langle f|\langle f|\text{SW}|f\rangle|f\rangle. \tag{6.25}$$

We now prove the following

Proposition 6.9. *Suppose that a unitary operator f passes the stabilizer Håstad test with $\delta = \frac{3}{4}\varepsilon$ with probability $1-\varepsilon$. Then f is ε-close, up to a phase, to $\mathbb{I}$ or a stabilizer dictator. (We assume $\varepsilon \leq 0.01$.)*

Proof. Write the Fourier expansion of f as follows:

$$f = \sum_{\mathbf{s}} \hat{f}_{\mathbf{s}} \chi_{\mathbf{s}}. \tag{6.26}$$

It is easy to verify that

$$\mathbb{P}[\text{test accepts}] = \sum_{\mathbf{s}} (1-\delta)^{|\mathbf{s}|} |\hat{f}_{\mathbf{s}}|^4. \tag{6.27}$$

Now suppose that f is a stabilizer dictator, up to a phase, on some variable $j \in [n]$, i.e., f is $e^{i\phi}\sigma_j^s$, for some $s \in \{0,1,2,3\}$. Then f passes the quantum Håstad test with probability

$$\mathbb{P}[\text{test accepts}] = (1-\delta)|\hat{f}_{\mathbf{s}}|^4 = 1-\delta. \tag{6.28}$$

On the other hand, suppose that

$$1-\varepsilon \leq \mathbb{P}[\text{test accepts}] \leq \left(\sum_{\mathbf{s}} |\hat{f}_{\mathbf{s}}|^2\right) \max_{\mathbf{s}} (1-\delta)^{|\mathbf{s}|} |\hat{f}_{\mathbf{s}}|^2 = \max_{\mathbf{s}} \left(1-\frac{3}{4}\varepsilon\right)^{|\mathbf{s}|} |\hat{f}_{\mathbf{s}}|^2. \tag{6.29}$$

Since $(1-\delta)^{|\mathbf{s}|} \leq 1$ it follows that there exists some $\mathbf{s}$ such that $|\hat{f}_{\mathbf{s}}|^2 \geq 1-\varepsilon$. Using the fact that $(1-\frac{3}{4}\varepsilon)^{|\mathbf{s}|} < 1-\varepsilon$ for $\text{supp}(\mathbf{s}) \geq 2$, we know that this maximum occurs on a string $\mathbf{s}$ with support one or zero. That is, there exists a Fourier coefficient of magnitude at least $1-\varepsilon$ on a string with support at most one. □

We have been unable to prove the corresponding result for the full quantum dictator test, so we leave this as a conjecture.

Conjecture 6.10. *Suppose that a unitary operator f passes the quantum Håstad dictator test with $\delta = \frac{3}{4}\varepsilon$ with probability $1-\varepsilon$. Then f is ε-close, up to a phase, to $\mathbb{I}$ or a dictator. (Assume $\varepsilon \leq 0.01$.)*

7 Learning quantum boolean functions

The purpose of this section is to describe a family of results in the spirit of the *Goldreich-Levin algorithm* [23], an algorithm which was originally defined in a cryptographic context, but was shown by Kushilevitz and Mansour [38] to be a useful tool for learning boolean functions. Continuing the theme of the previous sections, we'll see that quantum computers are polynomially more efficient at learning tasks for boolean functions. Heuristically, this is because quantum computers can exploit quantum superposition to carry out "super-dense" coding [42], allowing us to pack more information in a single quantum query.

The presentation of the results in this section is based on [44].

7.1 Learning stabilizer operators and approximating Fourier coefficients

We begin by learning the class of stabilizer operators. It turns out that this can be done with only one quantum query, in contrast to the n queries required classically. This is a natural generalisation of the quantum algorithm of Bernstein and Vazirani [8], which learns linear boolean functions with one quantum query. For simplicity, the results in this subsection are given in terms of quantum boolean functions, but it should be clear how to extend them to general unitary operators.

Proposition 7.1. *If a quantum boolean function f is a stabilizer operator, then we can identify f with 1 quantum query, using $O(n)$ quantum measurements of Pauli operators.*

Proof. The idea behind the proof is simple: we apply f to one half of a collection of n maximally entangled states and then measure in a basis of maximally entangled states. More precisely, suppose that $f = \chi_{\mathbf{s}}$ for some $\mathbf{s}$. Then the first step of our algorithm queries f to produce the state (our notation is identical to that of the previous section)

$$f \otimes \mathbb{I} |\Phi\rangle_{\mathbf{A}\mathbf{A}'} \equiv |\mathbf{s}\rangle. \tag{7.1}$$

Since the set of states $\{|\mathbf{s}\rangle\}$ indexed by $\mathbf{s}$ forms an orthonormal basis for $\mathbf{A}\mathbf{A}'$ we can simply measure the state $|\mathbf{s}\rangle$ to find out $\mathbf{s}$. (The $O(n)$ measurements bit comes from the preparation step and the measurement step; one needs to measure each register $A_j A'_j$ separately.) □

The next proposition shows us that the previous result is robust against perturbations.

Proposition 7.2. *Suppose that a quantum boolean function f satisfies*

$$\hat{f}_{\mathbf{s}} \geq \frac{1+\varepsilon}{\sqrt{2}} \tag{7.2}$$

for some $\mathbf{s}$. Then $\chi_{\mathbf{s}}$ can be identified with probability $1-\delta$ with $O\left(\frac{1}{\varepsilon^2}\log\left(\frac{1}{\delta}\right)\right)$ uses of f.

Proof. Note that, by Parseval, there can only be one character $\chi_{\mathbf{s}}$ which satisfies (7.2); the rest of the characters must be further from f than $\chi_{\mathbf{s}}$.

The strategy of our proof is simple: we make q queries to f by applying it to sets of maximally entangled states $|\Phi\rangle$ and then measure each resulting state in the $\{|\mathbf{s}\rangle\}$ basis. We then take a majority vote. For each query, with probability $\mathbb{P}[\text{succ}] \geq \frac{1}{2}+\varepsilon$, we get the right answer. To work out the probability that the test fails we bound the failure probability by bounding the cumulative distribution function of the binomial distribution $B(q,p)$ with $p = \frac{1}{2}+\varepsilon$:

$$\begin{aligned} \mathbb{P}[\text{test fails}] = \mathbb{P}[\text{at least } q/2 \text{ failures}] &\leq e^{-2\frac{(qp-\frac{q}{2})^2}{q}} \\ &= e^{-2q\varepsilon^2} = \delta, \end{aligned} \tag{7.3}$$

so that choosing $q = O\left(\frac{1}{\varepsilon^2}\log\left(\frac{1}{\delta}\right)\right)$ gives us the desired result. □

Lemma 7.3. *Let $f = \sum_{\mathbf{s}\in\{0,1,2,3\}^n} \hat{f}_{\mathbf{s}}\chi_{\mathbf{s}}$ be a quantum boolean function. Then $\hat{f}_{\mathbf{s}}^2 \geq \gamma^2$ for at most $\frac{1}{\gamma^2}$ terms.*

Proof. This is a simple consequence of Parseval's relation. □

Lemma 7.4. *For any* $\mathbf{s} \in \{0,1,2,3\}^n$ *it is possible to estimate* $\hat{f}_{\mathbf{s}}$ *to within* $\pm\eta$ *with probability* $1-\delta$ *using* $O\left(\frac{1}{\eta^2}\log\left(\frac{1}{\delta}\right)\right)$ *queries.*

Proof. To prove this lemma we need access to the controlled-f quantum boolean function $U_f = |0\rangle_C\langle 0| \otimes \mathbb{I}_A + |1\rangle_C\langle 1| \otimes f_A$. (This can be easily implemented using f alone by adjoining n ancilla qubits and pre-applying a controlled-SWAP operation between the main qubits and ancilla qubits, applying f to the ancillas and post-applying a controlled-SWAP operation and then discarding the ancillas.)

The method takes place on a register consisting of the system $\mathbf{A}$ and a copy of the system $\mathbf{A}'$ and a control qubit. It proceeds as follows.

1. Prepare the control+system+copy in $|0\rangle_C|\Phi\rangle_{\mathbf{AA}'}$.
2. Apply a Hadamard operation $H = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix}1 & 1\\ 1 & -1\end{smallmatrix}\right)$ to C: the system is now in the state
$$\frac{1}{\sqrt{2}}|0\rangle_C|\Phi\rangle_{\mathbf{AA}'} + \frac{1}{\sqrt{2}}|1\rangle_C|\Phi\rangle_{\mathbf{AA}'}. \tag{7.4}$$
3. Apply the controlled-$\chi_{\mathbf{s}}$ operation $V_{\chi_{\mathbf{s}}} = |0\rangle_C\langle 0| \otimes \chi_{\mathbf{s}} + |1\rangle_C\langle 1| \otimes \mathbb{I}$ (implemented in the same way as U_f, above) to yield
$$\frac{1}{\sqrt{2}}|0\rangle_C\chi_{\mathbf{s}}|\Phi\rangle_{\mathbf{AA}'} + \frac{1}{\sqrt{2}}|1\rangle_C|\Phi\rangle_{\mathbf{AA}'}. \tag{7.5}$$
4. Next apply U_f to give
$$\frac{1}{\sqrt{2}}|0\rangle_C\chi_{\mathbf{s}}|\Phi\rangle_{\mathbf{AA}'} + \frac{1}{\sqrt{2}}|1\rangle_C f|\Phi\rangle_{\mathbf{AA}'}. \tag{7.6}$$
5. Apply a Hadamard operation once again to C. The system is now in the state
$$\frac{1}{2}|0\rangle_C(\chi_{\mathbf{s}}|\Phi\rangle_{\mathbf{AA}'} + f|\Phi\rangle_{\mathbf{AA}'}) + \frac{1}{2}|1\rangle_C(\chi_{\mathbf{s}}|\Phi\rangle_{\mathbf{AA}'} - f|\Phi\rangle_{\mathbf{AA}'}). \tag{7.7}$$
6. Now measure the control qubit in the computational basis. This gives "0" with probability
$$\mathbb{P}[0] = \langle\Phi|\frac{(\chi_{\mathbf{s}}+f)^2}{4}|\Phi\rangle = \frac{1}{2} + \frac{1}{2\cdot 2^n}\operatorname{tr}(\chi_{\mathbf{s}}f) = \frac{1}{2} + \frac{1}{2}\hat{f}_{\mathbf{s}} \tag{7.8}$$
and "1" with probability
$$\mathbb{P}[1] = \langle\Phi|\frac{(\chi_{\mathbf{s}}-f)^2}{4}|\Phi\rangle = \frac{1}{2} - \frac{1}{2\cdot 2^n}\operatorname{tr}(\chi_{\mathbf{s}}f) = \frac{1}{2} - \frac{1}{2}\hat{f}_{\mathbf{s}}. \tag{7.9}$$

An application of Hoeffding's inequality yields the desired result. □

Remark 7.5. Note that one may improve the performance of the procedures involved the proofs of Proposition 7.2 and Lemma 7.4 by exploiting quantum amplitude amplification. This achieves a square-root improvement of the dependence on ε and η, respectively.

7.2 The quantum Goldreich-Levin algorithm

In this subsection we describe the quantum Goldreich-Levin algorithm.

Theorem 7.6 (quantum Goldreich-Levin)**.** *Given oracle access to a unitary operator f on n qubits and its adjoint $f^\dagger$, and given $\gamma,\delta>0$, there is a $\mathrm{poly}\left(n,\frac{1}{\gamma}\right)\log\left(\frac{1}{\delta}\right)$-time algorithm which outputs a list $L=\{\mathbf{s}_1,\mathbf{s}_2,\ldots,\mathbf{s}_m\}$ such that with probability $1-\delta$: (1) if $|\hat{f}_{\mathbf{s}}|\geq\gamma$, then $\mathbf{s}\in L$; and (2) for all $\mathbf{s}\in L$, $|\hat{f}_{\mathbf{s}}|\geq\gamma/2$.*

This quantum algorithm can be understood as a kind of branch and bound algorithm: we initially assume that the set S_n of all 4^n strings $\mathbf{s}$ contributes significantly to the Fourier expansion of f. We then partition this set into four equal chunks and efficiently estimate (via Proposition 7.12) the total weight of the Fourier expansion on each of these chunks. We then throw away the chunks with low weight and repeat the process by successively partitioning the remaining chunks into four, etc. The reason the total number of remaining chunks doesn't blow up exponentially is because of Lemma 7.3.

Definition 7.7. Let f be a unitary operator on n qubits. Let $I\subset[n]$. For any $\mathbf{s}\in\{0,1,2,3\}^{|I|}$ with $S\equiv\mathrm{supp}(\mathbf{s})\subset I$ define

$$F_{\mathbf{s};I}\equiv\frac{1}{2^{|I|}}\,\mathrm{tr}_I(\chi_{\mathbf{s}}f). \tag{7.10}$$

Lemma 7.8. *Let f be a unitary operator on n qubits, then*

$$\frac{1}{2^{n-|I|}}\,\mathrm{tr}(\chi_{\mathbf{t}}F_{\mathbf{s};I})=\hat{f}_{\mathbf{s}\cup\mathbf{t}}, \tag{7.11}$$

for any $\mathbf{t}\in\{0,1,2,3\}^{n-|I|}$ with $T\equiv\mathrm{supp}(\mathbf{t})\subset I^c$, where I^c denotes the complement of the set I and $\mathbf{s}\cup\mathbf{t}$ denotes concatenation, i.e.,

$$[\mathbf{s}\cup\mathbf{t}]_j=\begin{cases}s_j, & j\in I\\ t_j, & j\notin I.\end{cases} \tag{7.12}$$

Proof. Note that

$$\begin{aligned}F_{\mathbf{s};I}&=\frac{1}{2^{n-|I|}}\sum_{\mathrm{supp}(\mathbf{t})\subset I^c}\mathrm{tr}(\chi_{\mathbf{t}}F_{\mathbf{s};I})\chi_{\mathbf{t}}\\&=\frac{1}{2^n}\sum_{\mathrm{supp}(\mathbf{t})\subset I^c}\mathrm{tr}(\chi_{\mathbf{s}}\otimes\chi_{\mathbf{t}}f)\chi_{\mathbf{t}},\end{aligned} \tag{7.13}$$

from which the result follows. □

Lemma 7.9. *Let f be a unitary operator on n qubits. Then*

$$\frac{1}{2^{n-|I|}}\,\mathrm{tr}(F_{\mathbf{s};I}^\dagger F_{\mathbf{s};I})=\sum_{\mathbf{t}\,|\,t_j=s_j,\forall j\in I}|\hat{f}_{\mathbf{t}}|^2. \tag{7.14}$$

Proof. Consider

$$\frac{1}{2^{n-|I|}}\operatorname{tr}(F_{\mathbf{s};I}^{\dagger}F_{\mathbf{s};I}) = \sum_{\operatorname{supp}(\mathbf{u})\subset I^c} |\hat{f}_{\mathbf{s}\cup\mathbf{u}}|^2 = \sum_{\mathbf{t}|t_j=s_j, j\in I} |\hat{f}_{\mathbf{t}}|^2. \tag{7.15}$$

□

It is convenient to write the set $\mathcal{S}$ of all $\mathbf{t}$ such that $t_j = s_j$, $j \in I$ as an *indicator string*:

Definition 7.10. Let $I \subset [n]$. The indicator string $\mathcal{S} = (s_{j_1}, s_{j_2}, \ldots, s_{j_{|I|}}, *, *, \ldots, *)$, where $s_{j_k} \in \{0,1,2,3\}$, is the set

$$\{\mathbf{t}|t_{j_k} = s_{j_k}, j_k \in I\}. \tag{7.16}$$

Definition 7.11. The *weight* $W(\mathcal{S})$ of an indicator string $\mathcal{S}$ is

$$W(\mathcal{S}) = \sum_{\mathbf{t}\in\mathcal{S}} |\hat{f}_{\mathbf{t}}|^2. \tag{7.17}$$

It turns out that we can efficiently estimate $W(\mathcal{S})$.

Proposition 7.12. *Let f be a unitary operator on n qubits. Then for any indicator string $\mathcal{S}$ it is possible to efficiently estimate the weight $W(\mathcal{S})$ to within $\pm\gamma^2$ with probability $1-\delta$ using $O\left(\frac{1}{\gamma^4}\log\left(\frac{1}{\delta}\right)\right)$ queries to f and $f^{\dagger}$.*

Proof. Our method takes place in a system consisting of one ancilla qubit C, four copies, called A_1, A_1', A_2, and A_2', of the qubits in I and two copies, called B and B', of the qubits in I^c. Thus the total system is $CA_1A_1'A_2A_2'BB'$. The algorithm proceeds as follows.

1. Initialise the system (by applying a Hadamard on C) into

$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)_C |\Phi\rangle_{A_1A_1'} |\Phi\rangle_{A_2A_2'} |\Phi\rangle_{BB'}. \tag{7.18}$$

   (From now on, for simplicity, we write $|\Phi\rangle \equiv |\Phi\rangle_{A_1A_1'} |\Phi\rangle_{A_2A_2'} |\Phi\rangle_{BB'}$.)

2. Apply the operation $U \equiv |0\rangle_C\langle 0| \otimes \chi_{\mathbf{s}} \otimes \chi_{\mathbf{s}} + |1\rangle_C\langle 1| \otimes \mathbb{I}_{A_1A_2}$ on CA_1A_2. The system is now in the state

$$\frac{1}{\sqrt{2}}|0\rangle_C(\chi_{\mathbf{s}} \otimes \chi_{\mathbf{s}}|\Phi\rangle) + \frac{1}{\sqrt{2}}|1\rangle_C|\Phi\rangle. \tag{7.19}$$

3. Apply the controlled-$(f, f^{\dagger})$ operation $V \equiv |0\rangle_C\langle 0| \otimes \mathbb{I}_{A_1A_2B} + |1\rangle_C\langle 1| \otimes f_{A_1B} f_{A_2B}^{\dagger}$ on CA_1A_2B. (This operation is easy to implement with two applications of controlled-f and controlled-$f^{\dagger}$ operations.) The system is now in the state

$$\frac{1}{\sqrt{2}}|0\rangle_C(\chi_{\mathbf{s}} \otimes \chi_{\mathbf{s}}|\Phi\rangle) + \frac{1}{\sqrt{2}}|1\rangle_C(f_{A_1B} f_{A_2B}^{\dagger}|\Phi\rangle). \tag{7.20}$$

4. Apply a Hadamard to the control register:

$$\frac{1}{2}|0\rangle_C(\chi_{\mathbf{s}}\otimes\chi_{\mathbf{s}}+f_{A_1B}f_{A_2B}^{\dagger})|\Phi\rangle+\frac{1}{2}|1\rangle_C(\chi_{\mathbf{s}}\otimes\chi_{\mathbf{s}}-f_{A_1B}f_{A_2B}^{\dagger})|\Phi\rangle. \tag{7.21}$$

5. Measure the control register in the computational basis. We get "0" with probability:

$$\begin{aligned}\mathbb{P}[0]&=\frac{1}{4}\langle\Phi|(\chi_{\mathbf{s}}\otimes\chi_{\mathbf{s}}+f_{A_2B}f_{A_1B}^{\dagger})(\chi_{\mathbf{s}}\otimes\chi_{\mathbf{s}}+f_{A_1B}f_{A_2B}^{\dagger})|\Phi\rangle\\&=\frac{1}{2}+\frac{1}{2}\frac{1}{2^{n+|I|}}\mathrm{Re}(\mathrm{tr}(\chi_{\mathbf{s}}\otimes\chi_{\mathbf{s}}f_{A_1B}f_{A_2B}^{\dagger}))\\&=\frac{1}{2}+\frac{1}{2}\left(\frac{1}{2^{n-|I|}}\mathrm{tr}(F_{\mathbf{s};I}^{\dagger}F_{\mathbf{s};I})\right)=\frac{1}{2}+\frac{1}{2}W(\mathcal{S}),\end{aligned} \tag{7.22}$$

with a similar formula for $\mathbb{P}[1]$; the second equality follows from noting that $\langle\Phi|M_{A_1A_2B}\otimes\mathbb{I}_{A_1'A_2'B'}|\Phi\rangle=\frac{1}{2^{n+|I|}}\mathrm{tr}(M)$, for any operator M on A_1A_2B (the extra $|I|$ comes from the fact that A appears twice). An application of Hoeffding's inequality gives the desired result.

□

We now describe the quantum Goldreich-Levin algorithm.

Algorithm 1 Quantum Goldreich-Levin algorithm

$L\leftarrow(*,*,\ldots,*)$
for $k=1$ to n **do**
  for each $\mathcal{S}\in L$, $\mathcal{S}=(s_1,s_2,\ldots,s_{k-1},*,*,\ldots,*)$ **do**
    Let $\mathcal{S}_{s_k}=(s_1,s_2,\ldots,s_{k-1},s_k,*,*,\ldots,*)$ then for $s_k=0,1,2,3$ estimate $W(\mathcal{S}_{s_k})$ to within $\pm\gamma^2/4$ with probability at least $1-\delta$.
    Remove $\mathcal{S}$ from L.
    Add $\mathcal{S}_{s_k}$ if the estimate of $W(\mathcal{S}_{s_k})$ is at least $\gamma^2/2$ for $s_k=0,1,2,3$.
  end for
end for

We now analyse the algorithm. To simplify the analysis we'll assume that all estimations are accurate. We'll remove this assumption later.

Lemma 7.13. *After 1 iteration of the algorithm, $W(\mathcal{S})\geq\frac{\gamma^2}{4}$ for all $\mathcal{S}\in L$.*

Proof. All the estimates are assumed to be correct, and for all $\mathcal{S}\in L$, $\mathcal{S}$ was entered into the list L because its estimated weight was at least $\frac{\gamma^2}{2}$, and the estimate is correct to within an additive $\frac{\gamma^2}{4}$. □

Lemma 7.14. *At any time $|L|\leq\frac{4}{\gamma^2}$.*

Proof. The result follows from Lemma 7.3. □

Lemma 7.15. *The quantum Goldreich-Levin algorithm requires at most a total of $\frac{16n}{\gamma^2}$ estimations.*

Proof. At each iteration there are at most $\frac{4}{\gamma^2}$ items in the list and the algorithm needs at most 4 estimations per iteration. There are only n iterations. □

Lemma 7.16. *For any* $\mathbf{s}$ *such that* $|\hat{f}_{\mathbf{s}}| \geq \gamma$*, there exists* $\mathcal{S} \in L$ *such that* $\mathbf{s} \in \mathcal{S}$.

Proof. It suffices to note that $|\hat{f}_{\mathbf{s}}|^2 \geq \gamma^2$, thus the weight of any string $\mathcal{S}$ containing $\mathbf{s}$ is greater than γ^2, hence at least once such $\mathcal{S}$ remains in the list L after the first step. □

To remove the accuracy assumption, given $\delta > 0$, we define $\delta' = \frac{\delta\gamma^2}{16n}$, and perform each estimation with confidence $1-\delta'$. By the union bound, if the algorithm performs $\frac{16n}{\gamma^2}$ estimations, they are all correct with probability at least $1-\delta$, so the algorithm is correct with probability at least $1-\delta$.

We now have all the ingredients to prove Theorem 7.6:

Proof of Theorem 7.6. The total running time is dominated by the estimations. There are at most $\frac{16n}{\gamma^2}$, and each takes $O(\log(\frac{1}{\delta})/\gamma^2)$ samples to estimate, so the overall running time is $\mathrm{poly}(n, \frac{1}{\delta})\log(\frac{1}{\delta})$. □

8 Noise and quantum hypercontractivity

One of the most useful tools in the classical analysis of boolean functions has been the hypercontractive (also known as "Bonami-Gross-Beckner") inequality [6, 11, 12, 24, 40, 41, 46]. Perhaps the most intuitive way to write down this inequality is in terms of a *noise operator* on functions, which can be defined in two equivalent ways.

Definition 8.1. For a bit-string $x \in \{0,1\}^n$, define the distribution $y \sim_\varepsilon x$ as follows: each bit of y is equal to the corresponding bit of x with probability $1/2 + \varepsilon/2$, and flipped with probability $1/2 - \varepsilon/2$. Then the noise operator with rate $-1 \leq \varepsilon \leq 1$, written T_ε, is defined via

$$(T_\varepsilon f)(x) = \mathbb{E}_{y\sim_\varepsilon x}[f(y)]. \tag{8.1}$$

Equivalently, T_ε may be defined by its action on Fourier coefficients, as follows.

$$T_\varepsilon f = \sum_{\mathbf{s}\in\{0,1\}^n} \varepsilon^{|\mathbf{s}|} \hat{f}_{\mathbf{s}} \chi_{\mathbf{s}}. \tag{8.2}$$

It is easy to see that noise rate 1 leaves the function as it is, whereas noise rate 0 replaces the function with the constant function $\hat{f}_{\mathbf{0}}\mathbb{I}$. The Fourier-analytic definition given above immediately extends to the quantum setting, giving a superoperator defined as follows.

Definition 8.2. The noise superoperator with rate[3] $-1/3 \leq \varepsilon \leq 1$, written T_ε, is defined as follows.

$$T_\varepsilon f = \sum_{\mathbf{s}\in\{0,1,2,3\}^n} \varepsilon^{|\mathbf{s}|} \hat{f}_{\mathbf{s}} \chi_{\mathbf{s}}. \tag{8.3}$$

[3]This restriction on ε is necessary for the map to be completely positive [37].

Perhaps surprisingly, just as the classical noise operator has an alternative natural definition in terms of "local smoothing" (eqn. (8.1)), its quantum generalisation has a natural definition too, in terms of the action of the qubit depolarising channel.

Proposition 8.3. *Let f be a Hermitian operator on n qubits. Then, for $-1/3 \leq \varepsilon \leq 1$, $T_\varepsilon f = \mathcal{D}_\varepsilon^{\otimes n} f$, where $\mathcal{D}_\varepsilon$ is the* qubit depolarising channel *with noise rate ε, i.e.*

$$\mathcal{D}_\varepsilon(f) = \frac{(1-\varepsilon)}{2}\operatorname{tr}(f)\mathbb{I} + \varepsilon f. \tag{8.4}$$

Proof. To verify the claim it is only necessary to check (8.4) on the characters $\chi_{\mathbf{s}}$ and then extend by linearity. Thus, since for a single qubit $\mathcal{D}_\varepsilon(\chi_0) = \chi_0$, and $\mathcal{D}_\varepsilon(\chi_j) = \varepsilon\chi_j$ for $j \neq 0$ we have that

$$\mathcal{D}_\varepsilon^{\otimes n}(\chi_{\mathbf{s}}) = \varepsilon^{|\mathbf{s}|}\chi_{\mathbf{s}} = T_\varepsilon(\chi_{\mathbf{s}}). \tag{8.5}$$

□

We have the following easy observations about the behaviour of the noise superoperator.

- $\|T_\varepsilon f\|_p \leq \|f\|_p$ for any $0 \leq \varepsilon \leq 1$. Follows because $T_\varepsilon f$ is a convex combination of conjugations by unitaries.
- The semigroup property $T_\delta T_\varepsilon = T_{\delta\varepsilon}$ is immediate.
- If $\delta \leq \varepsilon$, $\|T_\delta f\|_p \leq \|T_\varepsilon f\|_p$ for any $p \geq 1$. Follows from the previous two properties.
- For any f and g, $\langle T_\varepsilon f, g\rangle = \langle f, T_\varepsilon g\rangle$. Follows from Plancherel's theorem: $\langle T_\varepsilon f, g\rangle = \sum_{\mathbf{s}} \widehat{T_\varepsilon f}_{\mathbf{s}}^{*}\, \hat{g}_{\mathbf{s}} = \sum_{\mathbf{s}} \varepsilon^{|\mathbf{s}|} \hat{f}_{\mathbf{s}}^{*} \hat{g}_{\mathbf{s}} = \sum_{\mathbf{s}} \hat{f}_{\mathbf{s}}^{*} \widehat{T_\varepsilon g}_{\mathbf{s}} = \langle f, T_\varepsilon g\rangle$.

We are now ready to state the quantum hypercontractive inequality. By the identification of the noise superoperator with a tensor product of qubit depolarising channels, this inequality is really a statement about the properties of this channel, and can also be seen as a generalisation of a hypercontractive inequality of Carlen and Lieb [16].

Theorem 8.4. *Let f be a Hermitian operator on n qubits and assume that $1 \leq p \leq 2 \leq q \leq \infty$. Then, provided that*

$$\varepsilon \leq \sqrt{\frac{p-1}{q-1}}, \tag{8.6}$$

we have

$$\|T_\varepsilon f\|_q \leq \|f\|_p. \tag{8.7}$$

Just as in the classical case, the proof of this inequality will involve two steps: first, a proof for $n = 1$, and then an inductive step to extend to all n. The proof of the base case is essentially the same as in the classical setting. In the classical proof, the inductive step uses a quite general tensor product argument.

One might hope that this argument extended to the quantum setting. This would be true if it held that, for any channels (superoperators) C and D on n qubits,

$$\|C \otimes D\|_{q\to p} \leq \|C\|_{q\to p}\|D\|_{q\to p}, \tag{8.8}$$

where we define the $q \to p$ norm as

$$\|C\|_{q\to p} = \sup_f \frac{\|Cf\|_p}{\|f\|_q}. \tag{8.9}$$

However, this most general statement is actually *false*, as it would imply the so-called "maximal output p-norm multiplicativity conjecture" in the case $q = 1$, to which counterexamples have recently been found by Hayden and Winter [27] for all $p > 1$ and for $p = 1$ by Hastings [26]. Our proof must therefore be specific to the depolarising channel, and will turn out to rely on a non-commutative generalisation of Hanner's inequality recently proven by King [36], rather than the Minkowski inequality used in the classical proof. In fact, a corollary of our result is the proof of multiplicativity of the maximum output $q \to p$ norm for the depolarising channel for certain values of q and p.

Before we begin the proof of Theorem 8.4 in earnest, we will need some subsidiary lemmata.

Lemma 8.5. *Let f be a single qubit Hermitian operator and let $1 \leq p \leq q \leq \infty$. Then, provided that*

$$\varepsilon \leq \sqrt{\frac{p-1}{q-1}} \tag{8.10}$$

we have

$$\|T_\varepsilon f\|_q \leq \|f\|_p. \tag{8.11}$$

Proof. We diagonalise f as $U^\dagger \Lambda U$, where Λ is diagonal. Since the depolarising channel is symmetric we have that $T_\varepsilon(U^\dagger \Lambda U) = \mathcal{D}_\varepsilon(U^\dagger \Lambda U) = U^\dagger \mathcal{D}_\varepsilon(\Lambda)U = U^\dagger T_\varepsilon(\Lambda)U$ and since the q-norm is unitarily invariant we have that $\|T_\varepsilon(f)\|_q = \|T_\varepsilon(\Lambda)\|_q$. So we may as well focus on diagonal operators $\Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$. In terms of the eigenvalues of Λ the inequality (8.11) is just the two-point inequality established by Bonami [12], Gross [24] and Beckner [6]. □

Lemma 8.6. *It suffices to prove Theorem 8.4 for $p = 2$.*

Proof. The classical proof (see, e.g., [39, Lecture 12]) goes through unchanged. The first step is to prove that Theorem 8.4 holds for $q = 2$, and any $1 \leq p \leq 2$. Assume the theorem holds for $p = 2$ and any q, and take p' such that $1/p' + 1/p = 1$; then

$$\|T_{\sqrt{p-1}} f\|_2 = \sup_{\|g\|_2=1} |\langle g, T_{\sqrt{p-1}} f\rangle| = \sup_{\|g\|_2=1} |\langle T_{\sqrt{p-1}} g, f\rangle| \tag{8.12}$$

$$\leq \|f\|_p \sup_{\|g\|_2=1} \|T_{\sqrt{p-1}} g\|_{p'} \leq \|f\|_p, \tag{8.13}$$

where the first inequality is Hölder's inequality. Now, to prove the theorem when $1 \leq p < 2 < q$, we use the semigroup property:

$$\|T_{\sqrt{\frac{p-1}{q-1}}} f\|_q = \|T_{1/\sqrt{q-1}} T_{\sqrt{p-1}} f\|_q \leq \|T_{\sqrt{p-1}} f\|_2 \leq \|f\|_p. \tag{8.14}$$

□

Lemma 8.7. *For all $p \geq 1$, $\|T_\varepsilon f\|_p \leq \||T_\varepsilon| f\|_p$, where $|f|$ is the operator $\sqrt{f^2}$.*

Proof. This holds because $f \leq |f|$ (in a positive semidefinite sense) and applying T_ε doesn't change this ordering because it is a convex combination of conjugations by unitaries. □

We are now ready to prove Theorem 8.4. By Lemma 8.6, it suffices to prove the following statement.

Proposition 8.8. *Let f be a Hermitian operator. Then $\|T_\varepsilon f\|_{1+1/\varepsilon^2} \leq \|f\|_2$ for any $0 \leq \varepsilon \leq 1$.*

Proof. For readability, it will be convenient to switch to the un-normalised standard Schatten p-norm, so for the remainder of the proof $\|f\|_p = (\mathrm{tr}\,|f|^p)^{1/p}$. With this normalisation, what we want to prove is

$$\|T_\varepsilon f\|_{1+1/\varepsilon^2} \leq 2^{-n\left(\frac{1-\varepsilon^2}{2(1+\varepsilon^2)}\right)} \|f\|_2. \tag{8.15}$$

The proof is by induction on n. The theorem is true for $n = 1$ by Lemma 8.5, so assume $n > 1$ and expand f as follows.

$$f = \mathbb{I} \otimes a + \sigma^1 \otimes b + \sigma^2 \otimes c + \sigma^3 \otimes d = \begin{pmatrix} a+d & b-ic \\ b+ic & a-d \end{pmatrix}. \tag{8.16}$$

Then, by direct calculation, we have

$$T_\varepsilon f = \begin{pmatrix} T_\varepsilon(a+\varepsilon d) & \varepsilon\, T_\varepsilon(b-ic) \\ \varepsilon\, T_\varepsilon(b+ic) & T_\varepsilon(a-\varepsilon d) \end{pmatrix}, \tag{8.17}$$

where the operator T_ε on the left-hand side acts on n qubits, while the operator T_ε on the right-hand side acts on $n-1$ qubits. For brevity, set $q = 1 + 1/\varepsilon^2$, and assume that $\|T_\varepsilon f\|_q^q \leq 2^{-(n-1)(q/2-1)} \|f\|_2^q$ for any $(n-1)$-qubit Hermitian operator f. Using a non-commutative Hanner's inequality for positive block matrices [36], which holds for $q \geq 2$, and noting that we can assume that f is positive by Lemma 8.7, we obtain

$$\begin{aligned} \|T_\varepsilon f\|_q^q &\leq \left\| \begin{pmatrix} \|T_\varepsilon(a+\varepsilon d)\|_q & \|\varepsilon T_\varepsilon(b-ic)\|_q \\ \|\varepsilon T_\varepsilon(b+ic)\|_q & \|T_\varepsilon(a-\varepsilon d)\|_q \end{pmatrix} \right\|_q^q \\ &\leq 2^{-(n-1)(q/2-1)} \left\| \begin{pmatrix} \|a+\varepsilon d\|_2 & \varepsilon\|b-ic\|_2 \\ \varepsilon\|b+ic\|_2 & \|a-\varepsilon d\|_2 \end{pmatrix} \right\|_q^q, \end{aligned}$$

where we use the inductive hypothesis. The proposition will thus be proven if we can show that

$$\begin{aligned} \|g\|_q^q &\equiv \left\| \begin{pmatrix} \|a+\varepsilon d\|_2 & \varepsilon\|b-ic\|_2 \\ \varepsilon\|b+ic\|_2 & \|a-\varepsilon d\|_2 \end{pmatrix} \right\|_q^q \\ &\leq 2^{-(q/2-1)} \|f\|_2^q = 2\left(\|a\|_2^2 + \|d\|_2^2 + \|b-ic\|_2^2\right)^{q/2}, \end{aligned}$$

where we call the matrix on the left-hand side of the inequality g. Write $g = T_\varepsilon h$ for some h, where the elements of h are

$$\begin{aligned} h_{11} &= \frac{1}{2}\left(\left(1+\frac{1}{\varepsilon}\right)\|a+\varepsilon d\|_2 + \left(1-\frac{1}{\varepsilon}\right)\|a-\varepsilon d\|_2\right), \\ h_{12} &= \|b-ic\|_2, \\ h_{21} &= \|b+ic\|_2, \\ h_{22} &= \frac{1}{2}\left(\left(1-\frac{1}{\varepsilon}\right)\|a+\varepsilon d\|_2 + \left(1+\frac{1}{\varepsilon}\right)\|a-\varepsilon d\|_2\right). \end{aligned}$$

Note that h is indeed Hermitian; $\|b-ic\|_2 = \|b+ic\|_2$, using the cyclicity of trace. Now, by Lemma 8.5, $\|g\|_q^q \leq 2^{-(q/2-1)}\|h\|_2^q = 2\left(\frac{1}{2}\|h\|_2^2\right)^{q/2}$. An explicit expansion gives

$$\|h\|_2^2 = 2\|b-ic\|_2^2 + \left(1+\frac{1}{\varepsilon^2}\right)\left(\|a\|_2^2+\varepsilon^2\|d\|_2^2\right) + \left(1-\frac{1}{\varepsilon^2}\right)\|a+\varepsilon d\|_2\|a-\varepsilon d\|_2, \tag{8.18}$$

implying that, in order to prove the proposition, we need

$$\left(\frac{1}{2}+\frac{1}{2\varepsilon^2}\right)\left(\|a\|_2^2+\varepsilon^2\|d\|_2^2\right) + \left(\frac{1}{2}-\frac{1}{2\varepsilon^2}\right)\|a+\varepsilon d\|_2\|a-\varepsilon d\|_2 \leq \|a\|_2^2+\|d\|_2^2. \tag{8.19}$$

So, noting that $1/2-1/(2\varepsilon^2)$ is negative, it suffices to show that

$$\|a\|_2^2 - \varepsilon^2\|d\|_2^2 \leq \|a+\varepsilon d\|_2\|a-\varepsilon d\|_2. \tag{8.20}$$

This last inequality can be proven using the matrix Cauchy-Schwarz inequality:

$$\begin{aligned}
\|a\|_2^2 - \varepsilon^2\|d\|_2^2 &= \mathrm{tr}((a+\varepsilon d)(a-\varepsilon d)) && (8.21)\\
&\leq \sqrt{\mathrm{tr}((a+\varepsilon d)^2)}\sqrt{\mathrm{tr}((a-\varepsilon d)^2)} && (8.22)\\
&= \|a+\varepsilon d\|_2\|a-\varepsilon d\|_2, && (8.23)
\end{aligned}$$

so we are done. □

Theorem 8.4 has the following easy corollaries. The first says, informally, that low-degree quantum boolean functions are smooth.

Corollary 8.9. *Let f be a Hermitian operator on n qubits with degree at most d. Then, for any $q \geq 2$, $\|f\|_q \leq (q-1)^{d/2}\|f\|_2$. Also, for any $p \leq 2$, $\|f\|_p \geq (p-1)^{d/2}\|f\|_2$.*

Proof. The proof follows that of Corollary 1.3 in Lecture 16 of [44] with no changes required. Explicitly,

$$\begin{aligned}
\|f\|_q^2 &= \left\|\sum_{k=0}^d f^{=k}\right\|_q^2 = \left\|T_{1/\sqrt{q-1}}\left(\sum_{k=0}^d (q-1)^{k/2} f^{=k}\right)\right\|_q^2 && (8.24)\\
&\leq \left\|\sum_{k=0}^d (q-1)^{k/2} f^{=k}\right\|_2^2 = \sum_{k=0}^d (q-1)^k \sum_{\mathbf{s},|\mathbf{s}|=k} \hat{f}_{\mathbf{s}}^2 && (8.25)\\
&\leq (q-1)^d \sum_{\mathbf{s}} \hat{f}_{\mathbf{s}}^2 = (q-1)^d \|f\|_2^2. && (8.26)
\end{aligned}$$

The second part is proved using the first part and Hölder's inequality, following immediately from

$$\|f\|_2^2 = \langle f, f\rangle \leq \|f\|_p\|f\|_q \leq (q-1)^{d/2}\|f\|_p\|f\|_2, \tag{8.27}$$

where $1/p+1/q=1$. □

The second corollary is a quantum counterpart of the fundamental Schwartz-Zippel Lemma [47, 50]. This lemma states that any non-zero function $f : \{0,1\}^n \to \mathbb{R}$ of degree d must take a non-zero value on at least a 2^{-d} fraction of the inputs. By analogy with the classical lemma, we conjecture that the constant in the exponent of this corollary can be improved.

Corollary 8.10. *Let f be a non-zero Hermitian operator on n qubits with m non-zero eigenvalues and degree d. Then $m \geq 2^{n-(2\log e)d} \approx 2^{n-2.89d}$.*

Proof. Let f_1 denote the projector onto the subspace spanned by f's eigenvectors that have non-zero eigenvalues. Then, for any $q \geq p \geq 1$,

$$\|f\|_p = \langle f^p, f_1 \rangle^{1/p} \leq \left(\|f^p\|_{q/p} \|f_1\|_{q/(q-p)}\right)^{1/p} = \|f\|_q \left(\frac{m}{2^n}\right)^{1/p-1/q}, \tag{8.28}$$

where we use Hölder's inequality and the fact that f_1's non-zero eigenvalues are all 1. Thus, using Corollary 8.9, for any $q \geq 2$ we have

$$\left(\frac{m}{2^n}\right)^{1/2-1/q} \geq \frac{\|f\|_2}{\|f\|_q} \geq (q-1)^{-d/2}, \tag{8.29}$$

implying

$$m \geq \frac{2^n}{\left((q-1)^{q/(q-2)}\right)^d}. \tag{8.30}$$

Taking the limit of this expression as $q \to 2$ gives the desired result. □

The final corollary is a quantum generalisation of a lemma attributed to Talagrand [49], which bounds the weight of a projector on the first level in terms of its 1-norm (equivalently, its dimension). This lemma quantifies the intuition that low-rank projections (eg., pure states) on the Hilbert space of n qubits must have a Fourier spectrum whose support includes high-weight Fourier coefficients.

Corollary 8.11. *Let $P^2 = P$ be a projector. Then*

$$\|P^{=1}\|_2^2 \leq (q-1)\|P\|_1^{\frac{2}{p}}, \tag{8.31}$$

where $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. We begin by writing

$$P = \sum_{\mathbf{s}} \hat{P}_{\mathbf{s}} \chi_{\mathbf{s}} = P^{=1} + h, \tag{8.32}$$

where

$$P^{=1} = \sum_{\mathbf{s}\,|\,|\mathbf{s}|=1} \hat{P}_{\mathbf{s}} \chi_{\mathbf{s}}. \tag{8.33}$$

Note that $\|P^{=1}\|_2^2 = \langle P^{=1}, P \rangle$. Applying Hölder's inequality:

$$\|P^{=1}\|_2^2 = \langle P^{=1}, P \rangle \leq \|P\|_p \|P^{=1}\|_q, \tag{8.34}$$

where $\frac{1}{p}+\frac{1}{q}=1$. Next we use hypercontractivity to show that $\|P^{=1}\|_q \leq \sqrt{q-1}\|P^{=1}\|_2$, so that

$$\|P^{=1}\|_2^2 \leq \sqrt{q-1}\|P^{=1}\|_2\|P\|_p = \sqrt{q-1}\|P^{=1}\|_2\|P\|_1^{\frac{1}{p}}. \tag{8.35}$$

Dividing both sides by $\|P^{=1}\|_2$ and squaring both sides gives us

$$\|P^{=1}\|_2^2 \leq (q-1)\|P\|_1^{\frac{2}{p}}. \tag{8.36}$$

□

9 A quantum FKN theorem

The Friedgut-Kalai-Naor (FKN) theorem [21] states that, if a boolean function has most of its Fourier weight on the first level or below, then it is close to being a dictator. It has proven useful in social choice theory [31] and the study of hardness of approximation [35, 18]. In this section, we state and prove two quantum variants of the FKN theorem. The first is a direct generalisation of the classical result, and uses the quantum hypercontractive inequality. The second is a different generalisation, to the ∞-norm.

9.1 Balancing quantum boolean functions

It will be convenient for the later results in this section to deal only with quantum boolean functions which have no weight on level 0 (i.e. are traceless). We therefore describe a method to *balance* quantum boolean functions. That is, from a quantum boolean function f we produce another quantum boolean function g which satisfies $\mathrm{tr}(g)=0$ and has $\|f^{\leq 1}\|_2^2 = \|g^{=1}\|_2^2$.

Definition 9.1. The *spin flip* operation S is defined as

$$S(M) = \sigma^2 M^* \sigma^2 \tag{9.1}$$

for any single qubit operator M.

The spin flip operation is a superoperator but is not a CP map. Note that $S(\sigma^j) = -\sigma^j$ for all $j \in \{1,2,3\}$.

Definition 9.2. The *balancing operation* $\mathcal{B}$ is the superoperator

$$\begin{aligned}\mathcal{B}(f) &= |0\rangle\langle 0| \otimes f - |1\rangle\langle 1| \otimes S^{\otimes n}(f) \\ &= |0\rangle\langle 0| \otimes f - |1\rangle\langle 1| \otimes (\sigma^2 \otimes \cdots \otimes \sigma^2) f^* (\sigma^2 \otimes \cdots \otimes \sigma^2),\end{aligned} \tag{9.2}$$

where we attach an ancilla qubit, denoted A.

Lemma 9.3. *Let f be a quantum boolean function. Then $g = \mathcal{B}(f)$ is a quantum boolean function.*

Proof. Consider

$$\begin{aligned} g^2 &= |0\rangle\langle 0| \otimes f^2 + |1\rangle\langle 1| \otimes S^{\otimes n}(f^2) \\ &= |0\rangle\langle 0| \otimes \mathbb{I} + |1\rangle\langle 1| \otimes S^{\otimes n}(\mathbb{I}) = \mathbb{I}. \end{aligned} \tag{9.3}$$

□

Proposition 9.4. *Let f be a quantum boolean function. Then* $\mathrm{tr}(\mathcal{B}(f)) = 0$.

Proof. The proof is by direct calculation.

$$\mathrm{tr}(\mathcal{B}(f)) = \mathrm{tr}(f) - \mathrm{tr}(f^*) = 2\,\mathrm{Im}(\mathrm{tr}(f)) = 0, \tag{9.4}$$

where the first equality follows from an application of the cyclic rule of trace. □

Proposition 9.5. *Let f be a quantum boolean function, and $g = \mathcal{B}(f)$. Then $\|g^{\leq 1}\|_2^2 = \|f^{\leq 1}\|_2^2$.*

Proof. We can calculate the first level of $\mathcal{B}(f)$ by tracing against weight-1 operators on the system:

$$\begin{aligned} \mathrm{tr}(\sigma_j^s \mathcal{B}(f)) &= \mathrm{tr}(\sigma_j^s f) - \mathrm{tr}(\sigma_j^s (\sigma^2 \otimes \cdots \otimes \sigma^2)(f^*)(\sigma^2 \otimes \cdots \otimes \sigma^2)) \\ &= 2\,\mathrm{tr}(\sigma_j^s f), \end{aligned} \tag{9.5}$$

recalling that the notation σ_i^j is used for the operator which acts as σ^j at the i'th position, and trivially elsewhere. That is, $\langle \sigma_j^s, g\rangle = \langle \sigma_j^s, f\rangle$, so all the degree 1 terms in the Fourier expansion of g (on the system) are identical to those for f. The only non-zero weight-1 term of g on the ancilla A is given by

$$\mathrm{tr}(\sigma_A^3 g) = 2\,\mathrm{tr}(f). \tag{9.6}$$

□

Our balancing operation reduces to the previously known classical balancing operation on classical boolean functions [21].

9.2 Exact quantum FKN

To gain some intuition for the later results, we begin by sketching the (straightforward) proof of an exact variant of the FKN theorem.

Proposition 9.6. *Let f be a quantum boolean function on n qubits. If $\sum_{|s|>1} \hat{f}_s^2 = 0$, then f is either a dictator or constant.*

Proof. By the results of Section 9.1, we can assume that f is balanced. Expand f as

$$f = \sum_{i=1}^{n} U_i, \tag{9.7}$$

where U_i is a traceless operator that acts non-trivially on only the i'th qubit, with at most two distinct eigenvalues. Let the positive eigenvalue of the non-trivial component of U_i be λ_i, and the corresponding

eigenvector be $|e_i\rangle$. Then the tensor product $\bigotimes_i |e_i\rangle$ is an eigenvector of f with eigenvalue $\lambda = \sum_i \lambda_i \leq \|f\|_\infty$. Let $\mathbf{i} : \mathbf{x}$ denote the string of length n with value x at position i, and 0 elsewhere. Because $U_i = \mathbb{I}_{[i-1]} \otimes M_i \otimes \mathbb{I}_{[i+1,\ldots,n]}$ and $M_i = \hat{f}_{\mathbf{i:1}}\sigma_i^1 + \hat{f}_{\mathbf{i:2}}\sigma_i^2 + \hat{f}_{\mathbf{i:1}}\sigma_i^3 = \begin{pmatrix} \hat{f}_{\mathbf{i:3}} & \hat{f}_{\mathbf{i:1}} - i\hat{f}_{\mathbf{i:2}} \\ \hat{f}_{\mathbf{i:1}} + i\hat{f}_{\mathbf{i:2}} & -\hat{f}_{\mathbf{i:3}} \end{pmatrix}$, we have that

$$\lambda_i = \sqrt{\hat{f}_{\mathbf{i:1}}^2 + \hat{f}_{\mathbf{i:2}}^2 + \hat{f}_{\mathbf{i:3}}^2}. \tag{9.8}$$

Thus $\sum_i \lambda_i$ is strictly greater than 1 unless f is a dictator. □

9.3 Quantum FKN in the 2-norm

In this section, we will prove the following result.

Theorem 9.7. *There is a constant K such that, for every quantum boolean function f, if $\sum_{|s|>1} \hat{f}_s^2 < \varepsilon$, f is $K\varepsilon$-close to being a dictator or constant.*

The proof is essentially a quantisation of one of the two proofs of the classical theorem given in the original paper [21]; see also the exposition in Lecture 13 of [44]. We will require the following lemma.

Lemma 9.8. *Let q be a degree 2 Hermitian operator on n qubits such that $\Pr_i[|\lambda_i(q)| > \delta] = p$, where the probability is taken with respect to the uniform distribution on the eigenvalues of q. Then $\|q\|_2^2 \leq \frac{\delta^2(1-p)}{1-9\sqrt{p}}$.*

Proof. Expand $q = r \oplus s$, where r projects onto the eigenvectors of q with eigenvalues at most δ in absolute value. Thus $\mathrm{rank}(r) \leq (1-p)2^n$ and $\mathrm{rank}(s) = p2^n$. Then the lemma follows from

$$\begin{aligned} \|q\|_2^2 &= \|r\|_2^2 + \|s\|_2^2 \leq (1-p)\delta^2 + \sqrt{p}\|s\|_4^2 \\ &\leq (1-p)\delta^2 + \sqrt{p}\|q\|_4^2 \leq (1-p)\delta^2 + 9\sqrt{p}\|q\|_2^2, \end{aligned}$$

where the final inequality is Corollary 8.9. □

We now turn to the proof of Theorem 9.7.

Proof of Theorem 9.7. By the results of Section 9.1, we can assume that f has no weight on level 0, i.e. is traceless. Given $f = \sum_{|s|=1} \hat{f}_s \chi_{\mathbf{s}} + \sum_{|s|>1} \hat{f}_s \chi_{\mathbf{s}}$, call the first sum l and the second h ("low" and "high"). As f is quantum boolean, $(l+h)^2 = \mathbb{I}$. Thus $l^2 + hl + lh + h^2 = \mathbb{I}$. On the other hand, by explicit expansion of l, we have

$$l^2 = \sum_{\mathbf{s},|\mathbf{s}|=1} \hat{f}_{\mathbf{s}}^2 \mathbb{I} + \sum_{\substack{\mathbf{s},\mathbf{t}, \\ |\mathbf{s}|=|\mathbf{t}|=1, \\ |\mathbf{s}\cap\mathbf{t}|=0}} \hat{f}_{\mathbf{s}}\hat{f}_{\mathbf{t}}\chi_{\mathbf{s}}\chi_{\mathbf{t}} = (1-\varepsilon)\mathbb{I} + q, \tag{9.9}$$

where we define a new operator q. Our goal will be to show that $\|q\|_2^2$ is small, i.e. at most $K\varepsilon$ for some

constant K. The theorem will follow: if

$$K\varepsilon \geq \|q\|_2^2 = \sum_{\substack{\mathbf{s},\mathbf{t},\\ |\mathbf{s}|=|\mathbf{t}|=1,\\ |\mathbf{s}\cap\mathbf{t}|=0}} \hat{f}_{\mathbf{s}}^2\hat{f}_{\mathbf{t}}^2 = \left(\sum_{\mathbf{s},|\mathbf{s}|=1}\hat{f}_{\mathbf{s}}^2\right)^2 - \sum_{\substack{\mathbf{s},\mathbf{t},\\ |\mathbf{s}|=|\mathbf{t}|=1,\\ |\mathbf{s}\cap\mathbf{t}|=1}} \hat{f}_{\mathbf{s}}^2\hat{f}_{\mathbf{t}}^2 \tag{9.10}$$

$$= (1-\varepsilon)^2 - \sum_{\substack{\mathbf{s},\mathbf{t},\\ |\mathbf{s}|=|\mathbf{t}|=1,\\ |\mathbf{s}\cap\mathbf{t}|=1}} \hat{f}_{\mathbf{s}}^2\hat{f}_{\mathbf{t}}^2, \tag{9.11}$$

then, for some K',

$$1-K'\varepsilon \leq \sum_{\substack{\mathbf{s},\mathbf{t},\\ |\mathbf{s}|=|\mathbf{t}|=1,\\ |\mathbf{s}\cap\mathbf{t}|=1}} \hat{f}_{\mathbf{s}}^2\hat{f}_{\mathbf{t}}^2 = \sum_{\mathbf{s},|\mathbf{s}|=1}\hat{f}_{\mathbf{s}}^2\Big(\sum_{\substack{\mathbf{t},|\mathbf{t}|=1,\\ |\mathbf{s}\cap\mathbf{t}|=1}}\hat{f}_{\mathbf{t}}^2\Big) \tag{9.12}$$

$$\leq \max_{\mathbf{s},|\mathbf{s}|=1}\sum_{\substack{\mathbf{t},|\mathbf{t}|=1,\\ |\mathbf{s}\cap\mathbf{t}|=1}}\hat{f}_{\mathbf{t}}^2, \tag{9.13}$$

so there exists some index i such that all but $K'\varepsilon$ of the weight of f is on terms that only depend on the i'th qubit. Setting all the other terms in the Fourier expansion of f to zero and renormalising gives a quantum boolean function that is a dictator (on the i'th qubit) and is distance $K''\varepsilon$ from f, for some other constant K''.

It remains to show that $\|q\|_2^2$ is small. Expand q as

$$q = \varepsilon\mathbb{I} - hl - lh - h^2 = \varepsilon\mathbb{I} - h(f-h) - (f-h)h - h^2 = \varepsilon\mathbb{I} - hf - fh + h^2 \tag{9.14}$$

and consider the terms in this sum. By the hypothesis of the theorem, $\|h\|_2^2 = \frac{1}{2^n}\sum_i \lambda_i(h)^2 = \varepsilon$. We also have $\mathrm{tr}\,h = 0$. By Chebyshev's inequality, this implies that, for any $K > 0$, $\Pr_i[|\lambda_i(h)| > K\sqrt{\varepsilon}] \leq 1/K^2$ (taking the uniform distribution on the eigenvalues of h).

We also have $\sigma_i(hf) \leq \sigma_i(h)$ (see [9, Problem III.6.2] and note that $\|f\|_\infty = 1$), and similarly for $\sigma_i(fh)$. As $h^2 \leq \mathbb{I}$, for all i, $|\lambda_i| \leq 1$ and so $\Pr_i[|\lambda_i(h^2)| > K\sqrt{\varepsilon}] \leq \Pr_i[|\lambda_i(h)| > K\sqrt{\varepsilon}] \leq 1/K^2$. This implies that (see [9, Problem III.6.5])

$$\Pr_i[|\lambda_i(q)| > 3K\sqrt{\varepsilon} + \varepsilon] \leq 3/K^2. \tag{9.15}$$

Using Lemma 9.8, taking K to be a sufficiently small constant, the result follows. □

9.4 Quantum FKN in the ∞-norm

In this subsection we'll prove a quantum generalisation of the FKN theorem in terms of the *supremum* or infinity norm $\|\cdot\|_\infty$. Our proof doesn't make use of hypercontractivity: only standard results from matrix analysis are employed.

Theorem 9.9. *Let f be a quantum boolean function. If $\|f-g\|_\infty \leq \varepsilon$, where g is a Hermitian operator with $g = g^{=1}$ and $\varepsilon < \frac{1}{2}$, then f is close to a dictator h, i.e., $\|f-h\|_\infty \leq 2\varepsilon$.*

Remark 9.10. This result has no classical analogue as $\|f-g\|_\infty$ can never be small if f and g are different.

Proof. By the results of Section 9.1, we can assume that f is traceless. Our proof works by using the infinity-norm closeness of f and g in an application of Weyl's perturbation theorem [9] to force the eigenvalues of f to be close to those of g:

$$|\lambda_j^{\downarrow}(f) - \lambda_j^{\downarrow}(g)| \leq \varepsilon, \quad j = 1,2,\ldots,2^n, \tag{9.16}$$

where $\lambda_j^{\downarrow}(M)$ denote the eigenvalues of M, in descending order. Thus:

$$\lambda_j^{\downarrow}(f) = \lambda_j^{\downarrow}(g) + \varepsilon(j), \quad j = 1,2,\ldots,2^n. \tag{9.17}$$

Since g can be written as $g = \sum_{j\in[n]} g_j$, with $\mathrm{supp}(g_j) = \{j\}$ and $\mathrm{tr}(g_j) = 0$, the eigenvalues of g are given by $\sum_{j=1}^n x_j\mu_j$, where $x_j \in \{-1,+1\}$ and $\mu_j = \|g_j\|_\infty$. (This follows because g_j is acts on the jth qubit as a 2-dimensional matrix.)

So we can label the eigenvalues of g with $\mathbf{x} \in \{-1,+1\}^n$ and we rewrite the perturbation condition as

$$\lambda(\mathbf{x}) = \sum_{j=1}^n x_j\mu_j + \varepsilon(\mathbf{x}), \tag{9.18}$$

where $\lambda(\mathbf{x})$ is the eigenvalue of f corresponding to the eigenvalue $\sum_{j=1}^n x_j\mu_j$ of g and $\varepsilon(\mathbf{x})$ is a correction. Our strategy is to now show that $\varepsilon(\mathbf{x})$ is a linear function.

To this end, we differentiate the 2^n equations (9.18) with respect to x_j:

$$\mu_j = D_j\lambda - D_j\varepsilon(\mathbf{x}), \tag{9.19}$$

where, eg.,

$$D_j\lambda = \frac{\lambda(x_1,\ldots,x_j=+1,\ldots,x_n) - \lambda(x_1,\ldots,x_j=-1,\ldots,x_n)}{2}, \tag{9.20}$$

i.e., we take the difference between λ with the jth variable assigned to 1 and to -1. Note that $D_j\lambda : \{-1,+1\}^{n-1} \to \mathbb{R}$.

Since f is quantum boolean, by assumption, we have that $|\lambda(\mathbf{x})| = 1$, $\forall \mathbf{x}$. So, because μ_j is constant and $|D_j\varepsilon(\mathbf{x})| \leq \varepsilon < 1/2$, and (9.19) is true for all $(x_1,\ldots,x_{j-1},x_{j+1},\ldots x_n)$, we conclude that $D_j\varepsilon(\mathbf{x})$ is constant for all $(x_1,\ldots,x_{j-1},x_{j+1},\ldots x_n)$. (Otherwise we'd have a contradiction: as we run through all the assignments of $(x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)$ the value of $D_j\lambda$ can, in principle, take both the values 0 and 1. However, owing to the constancy of μ_j and the fact that $|D_j\varepsilon(\mathbf{x})| \leq \varepsilon < 1/2$ only one of two possible values can be taken.) This is true for all $j \in [n]$. So we learn that $\varepsilon(\mathbf{x})$ is linear:

$$\varepsilon(\mathbf{x}) = \sum_{j=1}^n x_j\varepsilon_j. \tag{9.21}$$

But the condition that $|\varepsilon(\mathbf{x})| \leq \varepsilon$, $\forall \mathbf{x}$, implies that

$$\sum_{j=1}^{n} |\varepsilon_j| \leq \varepsilon. \tag{9.22}$$

Summarising what we've learnt so far:

$$D_j\lambda = \mu_j + \varepsilon_j. \tag{9.23}$$

Since $D_j\lambda : \{-1,+1\}^{n-1} \to \{-1,0,1\}$ we must have that $(\mu_j + \varepsilon_j) \in \{-1,0,1\}$, $j \in [n]$. But this actually means that there is exactly one j for which $\mu_j = 1 - \varepsilon_j$ (or $-1 - \varepsilon_j$); the rest satisfy $|\mu_j| = \varepsilon_j$. The reason for this is as follows. Suppose there was more than one such j. This would then lead to a contradiction as one can always find an assignment of the variables x_k so that $|\lambda(\mathbf{x})| > 1$, contradicting the quantum booleanity of f.

We now need to show that f is in fact close to a dictator. We define our dictator to be $h = \mathrm{sgn}(g)$. An application of the triangle inequality gives us the result:

$$\begin{aligned} \|f - h\|_\infty &\leq \|f - g\|_\infty + \|g - h\|_\infty \\ &\leq \varepsilon + \sum_{j=1}^{n} |\varepsilon_j| \leq 2\varepsilon \end{aligned} \tag{9.24}$$

□

10 Influence of quantum variables

In this section we introduce the notion of *influence* for quantum boolean functions, and establish some basic properties of the influence.

The classical definition of the influence of variable j on a boolean function f is the probability that f's value is undefined if the value of j is unknown, formally defined as

$$I_j(f) = \mathbb{P}_x[f(x) \neq f(x \oplus e_j)], \tag{10.1}$$

where $x \oplus e_j$ flips the jth bit of x. In the quantum case, we define the influence in terms of *derivative operators*.

Definition 10.1. The jth derivative operator d_j is the superoperator

$$d_j \equiv \frac{1}{2}(\mathcal{I} - S_j), \tag{10.2}$$

where $\mathcal{I}$ is the identity superoperator and S_j is the spin flip operation (9.1) on the jth qubit. Note that

$$d_j(\chi_{\mathbf{s}}) = \begin{cases} \chi_{\mathbf{s}}, & s_j \neq 0 \\ 0, & s_j = 0. \end{cases} \tag{10.3}$$

The *gradient* of f is

$$\nabla f \equiv (d_1(f), d_2(f), \ldots, d_n(f)). \tag{10.4}$$

The *laplacian* of f is

$$\nabla^2 f = \|\nabla f\|_2^2 = \sum_{j=1}^{n} \|d_j(f)\|_2^2. \tag{10.5}$$

The following lemma is immediate from the definition of the derivative operator.

Lemma 10.2. *Let f be an operator on n qubits. Then the d_j operator acts as follows.*

$$d_j(f) = \sum_{\mathbf{s}|s_j \neq 0} \hat{f}_{\mathbf{s}} \chi_{\mathbf{s}}. \tag{10.6}$$

Definition 10.3. Let f be a quantum boolean function. We define the *influence* of the jth qubit to be

$$I_j(f) \equiv \|d_j(f)\|_2^2, \tag{10.7}$$

and the *total influence* $I(f)$ to be

$$I(f) \equiv \sum_{j=1}^{n} I_j(f). \tag{10.8}$$

Note that this definition reduces to the classical definition when f is diagonal in the computational basis. Intuitively the quantum influence of the jth qubit measures the extent to which the value of a quantum boolean function is changed when in the input the state of jth qubit is inverted through the origin of the Bloch sphere.

Proposition 10.4. *Let f be a quantum boolean function. Then*

$$I_j(f) = \sum_{\mathbf{s}|s_j \neq 0} \hat{f}_{\mathbf{s}}^2 \tag{10.9}$$

and

$$I(f) = \sum_{\mathbf{s}} |\mathbf{s}| \hat{f}_{\mathbf{s}}^2. \tag{10.10}$$

Proof. Both results follow immediately from the definition of influence. □

The next result provides two other characterisations of the derivative.

Lemma 10.5. *Let f be a quantum boolean function. Then*

$$\begin{aligned} d_j(f) &= f - \mathrm{tr}_j(f) \otimes \frac{\mathbb{I}_j}{2} \\ &= f - \int dU \, U_j^\dagger f U_j \end{aligned} \tag{10.11}$$

and

$$I_j(f) = \frac{1}{2} \int dU \, \|[U_j, f]\|_2^2, \tag{10.12}$$

where the commutator $[U_j, f] \equiv U_j f - f U_j$, dU is the Haar measure on $U(2)$ and $U_j \equiv \mathbb{I} \otimes \cdots \otimes U \otimes \cdots \otimes \mathbb{I}$ with $\mathrm{supp}(U_j) = \{j\}$.

Proof. Both of the first two identities can be established by checking d_j on single qubit operators and extending by linearity.

The alternative characterisation of the influence can be proven as follows.

$$\begin{aligned}
I_j(f) &= \left\| f - \int dU\, U_j^\dagger f U_j \right\|_2^2 \\
&= \frac{1}{2^n} \int dU dV \ \mathrm{tr}\left((f - U_j^\dagger f U_j)(f - V_j^\dagger f V_j) \right) \\
&= \frac{1}{2^n} \int dU dV \ \mathrm{tr}\left(\mathbb{I} - U_j^\dagger f U_j f - f V_j^\dagger f V_j + U_j^\dagger f U_j V_j^\dagger f V_j \right) \\
&= 1 - \frac{2}{2^n} \int dU \ \mathrm{tr}\left(U_j^\dagger f U_j f \right) + \frac{1}{2^n} \int dU dV \ \mathrm{tr}\left(f U_j V_j^\dagger f V_j U_j^\dagger \right) \\
&= 1 - \frac{2}{2^n} \int dU \ \mathrm{tr}\left(U_j^\dagger f U_j f \right) + \frac{1}{2^n} \int dU \ \mathrm{tr}\left(f U_j f U_j^\dagger \right) \\
&= 1 - \frac{1}{2^n} \int dU \ \mathrm{tr}\left(U_j^\dagger f U_j f \right) \\
&= \frac{1}{2^n} \int dU \ \mathrm{tr}\left(\mathbb{I} - U_j^\dagger f U_j f \right) \\
&= \frac{1}{2^{n+1}} \int dU \ \mathrm{tr}\left([U_j, f][f, U_j^\dagger] \right) = \frac{1}{2} \int dU \, \| [U_j, f] \|_2^2 .
\end{aligned} \tag{10.13}$$

□

Now we generalise the single-qubit influence to multiple qubits.

Definition 10.6. Let f be a quantum boolean function. Then the influence of a set $J \subset [n]$ on f, written $I_J(f)$, is the quantity

$$I_J(f) \equiv \| d_J(f) \|_2^2, \tag{10.14}$$

where

$$d_J(f) \equiv f - \mathrm{tr}_J(f) \otimes \frac{\mathbb{I}_J}{2^{|J|}}. \tag{10.15}$$

The next result is a straightforward generalisation of Lemma 10.5.

Corollary 10.7. *Let f be a quantum boolean function, $J \subset [n]$, and $m = |J|$. Then*

$$d_J(f) = f - \int dU_1 dU_2 \cdots dU_m \, (U_1 \otimes U_2 \otimes \cdots \otimes U_m)^\dagger f (U_1 \otimes U_2 \otimes \cdots \otimes U_m) \tag{10.16}$$

and

$$I_J(f) = \int dU_1 dU_2 \cdots dU_m \| [U_1 \otimes U_2 \otimes \cdots \otimes U_m, f] \|_2^2, \tag{10.17}$$

where $J = \bigcup_{j=1}^m \mathrm{supp}(U_j)$.

Unlike the definition of influence on a single qubit, the physical interpretation of multiple qubit influence is less clear, and we leave it as an open question.

Definition 10.8. Let f be a quantum boolean function. Then we define the *variance* of f to be

$$\mathrm{var}(f) = \frac{1}{2^n}\mathrm{tr}(f^2) - \left(\frac{1}{2^n}\mathrm{tr}(f)\right)^2. \tag{10.18}$$

Proposition 10.9 (Quantum Poincaré inequality for n qubits)**.** *Let f be a quantum boolean function. Then*

$$\mathrm{var}(f) \leq \nabla^2(f) = I(f). \tag{10.19}$$

Proof. The proof follows from writing left and right-hand sides in terms of the Fourier expansion:

$$\mathrm{var}(f) = \sum_{\mathbf{s}} \hat{f}_{\mathbf{s}}^2 - f_{\mathbf{0}}^2 \tag{10.20}$$

and

$$I(f) = \sum_{\mathbf{s}} |\mathbf{s}| \hat{f}_{\mathbf{s}}^2. \tag{10.21}$$

The inequality is now obvious. □

Corollary 10.10. *Let f be a quantum boolean function such that $\mathrm{tr}(f) = 0$. Then there is a $j \in [n]$ such that $I_j(f) \geq 1/n$.*

11 Towards a quantum KKL theorem

An influential paper of Kahn, Kalai and Linial [30] proved the following result, known as the KKL theorem. For every balanced boolean function $f : \{0,1\}^n \to \{1,-1\}$, there exists a variable x such that $I_j(f) = \Omega\left(\frac{\log n}{n}\right)$. By Corollary 10.10, for every balanced quantum boolean function f on n qubits there is a qubit j such that $I_j(f) \geq \frac{1}{n}$. It is thus natural to conjecture that a quantum analogue of the KKL theorem holds – and also a quantum analogue of Friedgut's theorem [20], which is based on similar ideas.

However, the immediate quantum generalisation of the classical proof does not go through. Intuitively, the reason for this is as follows. The classical proof shows that, if the influences are all small, then their sum is large. This holds because, if the derivative operator in a particular direction has low norm, then it has small support, implying that it has some Fourier weight on a high level, which must be included in derivatives in many different directions. In the quantum case, this is not true: there exist quantum boolean functions whose derivative is small in a particular direction, but which are also low-degree.

On the other hand, it is immediate that the KKL theorem holds for quantum boolean functions which can be diagonalised by local unitaries, as these do not change the influence on each qubit. In this section we describe three partial results aimed at generalising the KKL theorem further. The first result is a simple quantisation of one of the classical proofs. This serves to illustrate what goes wrong when we generalise to the quantum world. Our next result shows that the classical proof technique breaks down precisely for anticommuting quantum boolean functions (qv.). Our final result is then a stronger version of KKL for a class of anticommuting quantum booolean functions.

11.1 A quantum Talagrand's lemma for KKL

The purpose of this section is to prove the following quantum generalisation of a theorem of Talagrand [48].

Proposition 11.1. *Let f be a traceless Hermitian operator on n qubits. Then*

$$\|f\|_2^2 \leq \sum_{i=1}^{n} \frac{10\|d_i f\|_2^2}{(2/3)\log(\|d_i f\|_2/\|d_i f\|_1)+1}. \tag{11.1}$$

In the case of classical boolean functions, this result can be applied to give an essentially immediate proof of the KKL theorem, using the fact that the functions $\{d_i f\}$ take values in $\{-1,0,1\}$. However, this does not extend to the quantum case, as the operators $d_i f$ have no such constraint on their eigenvalues. The proof of Proposition 11.1, on the other hand, is essentially an immediate generalisation of the classical proof in [48] (alternatively, see the exposition in [17]).

Proof of Proposition 11.1. For any operator g, define

$$M^2(g) = \sum_{\mathbf{s}\neq 0} \frac{\hat{g}_{\mathbf{s}}^2}{|\mathbf{s}|}. \tag{11.2}$$

Then it is clear that

$$\|f\|_2^2 = \sum_{i=1}^{n} M^2(d_i(f)). \tag{11.3}$$

Our strategy will be to find upper bounds on $M^2(g)$ for any traceless operator g. For some integer $m \geq 0$, expand

$$\begin{aligned} M^2(g) &= \sum_{1\leq|\mathbf{s}|\leq m} \frac{\hat{g}_{\mathbf{s}}^2}{|\mathbf{s}|} + \sum_{|\mathbf{s}|>m} \frac{\hat{g}_{\mathbf{s}}^2}{|\mathbf{s}|} && (11.4)\\ &\leq \sum_{k=1}^{m} \left(\frac{2^k}{k}\right) \|g^{=k}\|_{3/2}^2 + \frac{1}{m+1} \sum_{|\mathbf{s}|>m} \hat{g}_{\mathbf{s}}^2 && (11.5)\\ &\leq \|g\|_{3/2}^2 \sum_{k=1}^{m} \frac{2^k}{k} + \frac{1}{m+1}\|g\|_2^2, && (11.6) \end{aligned}$$

where we use quantum hypercontractivity (Corollary 8.9) in the first inequality. In order to bound the first sum, we note that $\sum_{k=1}^{m} 2^k/k \leq 4\cdot 2^m/(m+1)$, which can be proved by induction, so

$$M^2(g) \leq \frac{1}{m+1}\left(4\cdot 2^m \|g\|_{3/2}^2 + \|g\|_2^2\right). \tag{11.7}$$

Now pick m to be the largest integer such that $2^m\|g\|_{3/2}^2 \leq \|g\|_2^2$. Then $2^{m+1}\|g\|_{3/2}^2 \geq \|g\|_2^2$, and also $m+1 \geq 1$. Thus

$$m+1 \geq \frac{1}{2}\left(2\log\left(\frac{\|g\|_2}{\|g\|_{3/2}}\right)+1\right), \tag{11.8}$$

implying

$$M^2(g) \leq \frac{10\,\|g\|_2^2}{2\log(\|g\|_2/\|g\|_{3/2})+1}. \tag{11.9}$$

Noting that $\|g\|_2/\|g\|_{3/2} \geq (\|g\|_2/\|g\|_1)^{1/3}$, which can be proven using Cauchy-Schwarz, and summing $M^2(d_i f)$ over i completes the proof of the proposition. □

The worst case for this inequality is where the 2-norms and 1-norms of the operators $\{d_i f\}$ are the same. We therefore address this case in the next section.

11.2 A KKL theorem for anticommuting quantum boolean functions

In this subsection we study the situation where the 2-norms and 1-norms of the operators $\{d_i f\}$ are the same. This situation is the "worst-case scenario" for a straightforward quantum generalisation of the classical proof. We show that if this is the case then f must be a sum of anticommuting quantum boolean functions and so we identify this class as the "most quantum" of quantum boolean functions. While this class might be expected to avoid a KKL-type theorem, we then study a subclass of such quantum boolean functions and the nevertheless provide a lower bound for the influence (which is better than a KKL-type bound).

Definition 11.2. Let f be a quantum boolean function. Then the set $J \subset [n]$ of variables is said to have *bad influence* on f if

$$\left\| f - \mathrm{tr}_J(f) \otimes \frac{\mathbb{I}}{2^{|J|}} \right\|_2 = \left\| f - \mathrm{tr}_J(f) \otimes \frac{\mathbb{I}}{2^{|J|}} \right\|_1. \tag{11.10}$$

Lemma 11.3. *Let M be an $n \times n$ Hermitian matrix. If $\|M\|_2 = \|M\|_1$ then the eigenvalues λ_j of M satisfy*

$$\lambda_j \in \{-\alpha, +\alpha\}, \quad j \in [n], \tag{11.11}$$

for some $\alpha \in \mathbb{R}$.

Proof. Diagonalising M and writing out the equality $\|M\|_2^2 = \|M\|_1^2$ gives us

$$\frac{1}{n}\sum_{j=1}^n |\lambda_j|^2 = \frac{1}{n^2}\left(\sum_{j=1}^n |\lambda_j|^2 + 2\sum_{j<k}^n |\lambda_j||\lambda_k|\right). \tag{11.12}$$

Multiplying through by n^2 and rearranging gives us

$$(n-1)\sum_{j=1}^n |\lambda_j|^2 - 2\sum_{j<k}^n |\lambda_j||\lambda_k| = 0 \tag{11.13}$$

But this is the same as

$$\sum_{j<k} (|\lambda_j| - |\lambda_k|)^2 = 0, \tag{11.14}$$

so that $|\lambda_j| = |\lambda_k| = \alpha$, $j < k$, for some constant α. □

The next lemma quantifies the structure of quantum boolean functions with bad influence. We will see that they are highly constrained: they must be a sum of anticommuting quantum boolean functions.

Lemma 11.4. *Let f be a quantum boolean function. Then $J \subset [n]$ has bad influence on f if and only if*

$$f = \sqrt{1-\alpha^2} f' \otimes \mathbb{I}_J + \alpha g \tag{11.15}$$

where f' and g are quantum boolean functions, $\{f' \otimes \mathbb{I}_J, g\} = 0$, and $\alpha^2 = I_J(f)$.

Proof. Write

$$f = f_{J^c} \otimes \mathbb{I}_J + \sum_{\substack{\mathrm{supp}(\mathbf{s}) \subset J \\ \mathbf{s} \neq \mathbf{0}}} f_{\mathbf{s}} \otimes \chi_{\mathbf{s}}. \tag{11.16}$$

The derivative of f with respect to J is given by

$$d_J(f) = \sum_{\substack{\mathrm{supp}(\mathbf{s}) \subset J \\ \mathbf{s} \neq \mathbf{0}}} f_{\mathbf{s}} \otimes \chi_{\mathbf{s}}. \tag{11.17}$$

Since J has bad influence we have, according to Lemma 11.3, that

$$|d_J(f)| = \alpha \mathbb{I}, \tag{11.18}$$

so that, defining $g = \frac{1}{\alpha} d_J(f)$, we have $g^2 = \mathbb{I}$. Thus we can write $f = f_{J^c} \otimes \mathbb{I}_J + \alpha g$.

We now square f to find

$$\mathbb{I} = f_{J^c}^2 \otimes \mathbb{I}_J + \alpha^2 \mathbb{I} + \alpha \{f_{J^c} \otimes \mathbb{I}_J, g\}. \tag{11.19}$$

This implies that $\{f_{J^c} \otimes \mathbb{I}_J, g\} = 0$. Thus, defining

$$f' = \frac{1}{\sqrt{1-\alpha^2}} f_{J^c}, \tag{11.20}$$

we have $f'^2 = \mathbb{I}$ and the result follows. □

Let f be quantum boolean with Fourier expansion $f = \sum_s \hat{f}_{\mathbf{s}} \chi_{\mathbf{s}}$. If $[\chi_{\mathbf{s}}, \chi_{\mathbf{t}}] = 0$ for all $s \neq t$ where $\hat{f}_{\mathbf{s}}$ and $\hat{f}_{\mathbf{t}}$ are both non-zero, then we call f *commuting*. Similarly, if $\{\chi_{\mathbf{s}}, \chi_{\mathbf{t}}\} = 0$ for all $s \neq t$ where $\hat{f}_{\mathbf{s}}$ and $\hat{f}_{\mathbf{t}}$ are both non-zero, then we call f *anticommuting*. It follows from the classical KKL theorem that those commuting quantum boolean functions that can be diagonalised by local unitaries have a qubit with influence at least $\Omega\left(\frac{\log n}{n}\right)$. In the remainder of this section, we will show that anticommuting quantum boolean functions also have an influential qubit. Indeed, the influence of this qubit must be very high.

Proposition 11.5. *Let f be an anticommuting quantum boolean function on n qubits. Then there exists a j such that $I_j(f) \geq \frac{1}{\sqrt{n}}$.*

Proof. Our approach will be to show that $\sum_{j=1}^{n} I_j(f)^2 \geq 1$, whence the theorem follows trivially. Write $f = \sum_{i=1}^{m} w_i f_i$ for some m, where w_i is real and f_i is an arbitrary stabilizer operator. Let S_j be the set of indices of the stabilizer operators that act non-trivially on qubit j, i.e. the set $\{i : d_j(f_i) \neq 0\}$. Then $I_j(f) = \sum_{i \in S_j} w_i^2$, so

$$\sum_{j=1}^{n} I_j(f)^2 = \sum_{j=1}^{n} \left(\sum_{i \in S_j} w_i^2 \right)^2 = \sum_{j=1}^{n} \sum_{i,k \in S_j} w_i^2 w_k^2.$$

Rearrange the sum as follows.

$$\sum_{j=1}^{n} I_j(f)^2 = \sum_{i,k=1}^{m} w_i^2 w_k^2 |\{j : i \in S_j, k \in S_j\}|.$$

For each pair of stabilizer operators f_i, f_k to anticommute, they must both act non-trivially on the same qubit in at least one place. Thus

$$\sum_{j=1}^{n} I_j(f)^2 \geq \sum_{i,k=1}^{m} w_i^2 w_k^2 = \left(\sum_{i=1}^{m} w_i^2 \right)^2 = 1$$

and the proof is complete. □

This result hints that quantum KKL may be true, as it holds for two "extremal" cases (classical boolean functions and anticommuting quantum boolean functions).

12 Conclusions and conjectures

We have introduced the concept of a quantum boolean function, and have quantised some results from the classical theory of boolean functions. However, there is still a hoard of interesting results which we have not yet been able to plunder. We list some specific conjectures in this vein below.

1. **Quantum locality and dictator testing.** We have candidate quantum tests for the properties of locality and being a dictator (Conjectures 6.7 and 6.10), but have not been able to analyse their probability of success.

2. **General hypercontractivity.** We have proven hypercontractivity of the noise superoperator (Theorem 8.4) only in the case where $1 \leq p \leq 2 \leq q \leq \infty$. We conjecture that, as with the classical case, this in fact holds for all $1 \leq p \leq q \leq \infty$.

3. **Every quantum boolean function has an influential variable.** The results of Section 11 prove a quantum generalisation of the KKL theorem in some special cases. We conjecture that it holds in general, but, as we argued, a proof of such a theorem appears to require quite different techniques to the classical case.

4. **Lower bounds on the degree of quantum boolean functions.** A classical result of Nisan and Szegedy [43] states that the degree of any boolean function that depends on n variables must be at least $\log n - O(\log\log n)$. We conjecture that the degree of any quantum boolean function that acts non-trivially on n qubits is also $\Omega(\log n)$. The classical proof does not go through immediately: it relies on the fact that the influence of each variable of a degree d boolean function is at least $1/2^d$, which is not the case for degree d quantum boolean functions (for a counterexample, see (5.11)).

It is interesting to note that the proofs in the classical theory of boolean functions which go through easily to the quantum case tend to be those based around techniques such as Fourier analysis, whereas proofs based on combinatorial and discrete techniques do not translate easily in general. There are many such combinatorial results which would be interesting to prove or disprove in the quantum regime.

Acknowledgements

AM was supported by the EC-FP6-STREP network QICS. TJO was supported by the University of London central research fund. We'd like to thank Koenraad Audenaert, Jens Eisert, and Aram Harrow for helpful conversations, and also Ronald de Wolf and a STOC'09 referee for helpful comments on a previous version.

References

[1] SCOTT AARONSON: Quantum computing, postselection, and probabilistic polynomial-time. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 461(2063):3473–3482, 2005. 2

[2] SCOTT AARONSON: The learnability of quantum states. *Proceedings of the Royal Society A*, 463:3089–3114, 2007. 3

[3] MARK ADCOCK AND RICHARD CLEVE: A quantum Goldreich-Levin Theorem with cryptographic applications. In G. GOOS, J. HARTMANIS, AND J. VAN LEEUWEN, editors, *19th Annual Symposium on Theoretical Aspects of Computer Science*, volume 2285 of *Lecture notes in computer science*, pp. 323–334, Heidelberg, Berlin, 2002. Springer-Verlag. 4

[4] DORIT AHARONOV: Pondering about QPCP's. 2008. 2

[5] ALP ATICI AND ROCCO A. SERVEDIO: Quantum algorithms for learning and testing juntas. *Quantum Information Processing*, 6:323–348, 2007. 4

[6] WILLIAM BECKNER: Inequalities in Fourier analysis. *Ann. of Math.*, 102(1):159–182, 1975. [MR:0385456] 3, 22, 24

[7] AVRAHAM BEN-AROYA, ODED REGEV, , AND RONALD DE WOLF: A hypercontractive inequality for matrix-valued functions with applications to quantum computing and LDCs. In *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science*, New York, 2008. IEEE Computer Society. held in Philadelphia, PA, October 25–28, 2008. 3, 4

[8] ETHAN BERNSTEIN AND UMESH VAZIRANI: Quantum complexity theory. *SIAM J. Comput.*, 26(5):1411–1473, 1997. [MR:1471988] 4, 17

[9] RAJENDRA BHATIA: *Matrix analysis*. Springer-Verlag, New York, 1997. [MR:1477662] 5, 31, 32

[10] MANUEL BLUM, MICHAEL LUBY, AND RONITT RUBINFELD: Self-testing/correcting with applications to numerical problems. In *Proceedings of the 22nd Annual ACM Symposium on Theory of Computing*, volume 47, pp. 549–595, Baltimore, MD, 1990, 1993. [MR:1248868] 14

[11] ALINE BONAMI: Ensembles $\Lambda(p)$ dans le dual de D^∞. *Ann. Inst. Fourier*, 18(fasc. 2):193–204, 1968. [MR:0249940] 22

[12] ALINE BONAMI: Étude des coefficients de Fourier des fonctions de $L^p(G)$. *Ann. Inst. Fourier*, 20(fasc. 2):335–402, 1970. [MR:0283496] 3, 22, 24

[13] JEAN BOURGAIN: On the distribution of the Fourier spectrum of Boolean functions. *Isr. J. Math.*, 131(1):269–276, 2002. 4

[14] N. H. BSHOUTY AND J. C. JACKSON: Learning DNF over the uniform distribution using a quantum example oracle. *SIAM J. Comput.*, 28:1136–1153, 1999. 4

[15] HARRY BUHRMAN, LANCE FORTNOW, ILAN NEWMAN, AND HEIN RÖHRIG: Quantum property testing. In *Proceedings of the fourteenth annual ACM-SIAM symposium on Discrete Algorithms*, pp. 480–488, New York, 2003. Association for Computing Machinery (ACM). held in Baltimore, Maryland, January 12–14, 2003. 4, 11

[16] E. CARLEN AND E. LIEB: Optimal hypercontractivity for Fermi fields and related non-commutative integration inequalities. *Comm. Math. Phys*, 155:27–46, 1993. 23

[17] RONALD DE WOLF: A brief introduction to Fourier analysis on the boolean cube. *Theory of Computing Library Graduate Surveys*, 1:1–20, 2008. 2, 37

[18] IRIT DINUR: The PCP theorem by gap amplification. *Journal of the ACM*, 54(3):241–250, 2007. 2, 11, 28

[19] E. FISCHER: The art of uninformed decisions: A primer to property testing. *Bulletin of the European Association for Theoretical Computer Science*, 75:97–126, 2001. 11

[20] EHUD FRIEDGUT: Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica*, 18(1):27–35, 1998. 4, 36

[21] EHUD FRIEDGUT, GIL KALAI, AND ASSAF NAOR: Boolean functions whose Fourier transform is concentrated on the first two levels. *Adv. in Appl. Math.*, 29(3):427–437, 2002. 3, 28, 29, 30

[22] DMITRY GAVINSKY, JULIA KEMPE, IORDANIS KERENIDIS, RAN RAZ, AND RONALD DE WOLF: Exponential separations for one-way quantum communication complexity, with applications to cryptography. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, pp. 516–525, New York, 2007. Association for Computing Machinery (ACM). held in San Diego, CA, June 11–13, 2007. 4

[23] ODD GOLDREICH AND LEONID A. LEVIN: A hard-core predicate for all one-way functions. In *Proceedings of the Twenty First Annual ACM Symposium on Theory of Computing*, pp. 25–32, 1989. 3, 16

[24] LEONARD GROSS: Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975. [MR:0420249] 3, 22, 24

[25] JOHAN HÅSTAD: Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001. [MR2144931 (2006c:68066)] 11, 15

[26] M. B. HASTINGS: A counterexample to additivity of minimum output entropy. *Nature Physics*, 5, 2008, 255. 24

[27] PATRICK HAYDEN AND ANDREAS WINTER: Counterexamples to the maximal p-norm multiplicativity conjecture for all $p > 1$. *Communications of Mathematical Physics*, 284(1), 2008. 24

[28] C. W. HELSTROM: *Quantum detection and estimation theory*. Academic Press, New York, 1976. 12

[29] A. S. HOLEVO: Statistical decision theory for quantum systems. *Journal of Multivariate Analysis*, 3:337–394, 1973. 12

[30] J. KAHN, G. KALAI, AND N. LINIAL: The influence of variables on Boolean functions. In *Proceedings of the 29th Annual IEEE Symposium on Foundations of Computer Science (FOCS'88)*, pp. 68–80. IEEE Computer Society, 1988. Los Alamitos, CA, USA. 3, 4, 8, 36

[31] G. KALAI: A Fourier-theoretic perspective on the Condorcet paradox and Arrow's theorem. *Advances in Applied Mathematics*, 29(3):412–426, 2002. 28

[32] E. KASHEFI, A. KENT, V. VEDRAL, AND K. BANASZEK: A comparison of quantum oracles. *Phys. Rev. A*, 65(050304), 2002. 6

[33] J. KEMPE, O. REGEV, F. UNGER, AND R. DE WOLF: Upper bounds on the noise threshold for fault-tolerant quantum computing. In *Proc. 35th International Colloquium on Automata, Languages and Programming (ICALP'08)*, p. 856, 2008. 9

[34] IORDANIS KERENIDIS AND RONALD DE WOLF: Exponential lower bound for 2-query locally decodable codes via a quantum argument. *J. Comput. System Sci.*, 69(3):395–420, 2004. 2

[35] S. KHOT, G. KINDLER, E. MOSSEL, AND R. O'DONNELL: Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pp. 146–154, New York, 2004. IEEE Computer Society. held in Rome, Italy, October 17–19, 2004. 28

[36] C. KING: Inequalities for trace norms of 2x2 block matrices. *Comm. Math. Phys.*, 242(3):531–545, 2003. 24, 25

[37] C. KING AND M. B. RUSKAI: Minimal entropy of states emerging from noisy quantum channels. *Comm. Math. Phys.*, 47(1):192–209, 2001. 22

[38] E. KUSHILEVITZ AND Y. MANSOUR: Learning decision trees using the Fourier spectrum. *Siam J. Comput.*, 22(6):1331–1348, 1993. 3, 16

[39] ELCHANAN MOSSEL: Stat 206a: Polynomials of random variables, 2005. class notes. 24

[40] EDWARD NELSON: A quartic interaction in two dimensions. In *Mathematical Theory of Elementary Particles*, pp. 69–73, Cambridge, Mass., 1966. M.I.T. Press. Proc. Conf., Dedham, Mass., 1965. [MR:0210416] 22

[41] EDWARD NELSON: Construction of quantum fields from Markoff fields. *J. Functional Analysis*, 12:97–112, 1973. [MR:0343815] 22

[42] MICHAEL A. NIELSEN AND ISAAC L. CHUANG: *Quantum computation and quantum information*. Cambridge University Press, Cambridge, 2000. [MR:1796805] 2, 16

[43] NOAM NISAN AND MARIO SZEGEDY: On the degree of Boolean functions as real polynomials. *Computational Complexity*, 4(4):301–313, 1994. 41

[44] RYAN O'DONNELL: 15-859S: Analysis of Boolean Functions, 2007. lecture notes. 2, 3, 16, 26, 30

[45] JAIKUMAR RADHAKRISHNAN AND MADHU SUDAN: On Dinur's proof of the PCP theorem. *Bull. Amer. Math. Soc.*, 44(1):19–61, 2006. 2

[46] WALTER RUDIN: Trigonometric series with gaps. *J. Math. Mech.*, 9:203–227, 1960. [MR:0116177] 22

[47] J. T. SCHWARTZ: Fast probabilistic algorithms for verification of polynomial identities. *J. Assoc. Comput. Mach.*, 27:701–717, 1980. 27

[48] MICHEL TALAGRAND: On Russo's approximate zero-one law. *Ann. Prob.*, 23(3):1576–1587, 1994. 37

[49] MICHEL TALAGRAND: How much are increasing sets positively correlated? *Combinatorica*, 16(2):243–258, 1996. 27

[50] R. ZIPPEL: Probabilistic algorithms for sparse polynomials. In *Proc. EUROSAM'79*, volume 72 of *Lecture Notes in Computer Science*, pp. 216–226. Springer-Verlag, 1979. 27

AUTHORS

Ashley Montanaro
Department of Computer Science
University of Bristol
Woodland Road, Bristol, BS8 1UB, UK
montanar [at] cs [dot] bris [dot] ac [dot] uk

Tobias J. Osborne
Department of Mathematics
Royal Holloway, University of London
Egham, TW20 0EX, UK
tobias [dot] osborne [at] rhul [dot] ac [dot] uk