Quantum PCP

Aareyan Manzoor

Contents

1 Preliminaries

1.1 notation

We will use braket notation. So $|\xi\rangle$ will denote a vector in some hilbert space $\mathcal{H}$. $\langle\xi|$ will denote its dual in $\mathcal{H}^*$. I.e it is a functional:

$$\langle\xi| : \mathcal{H} \to \mathbb{C} \quad \langle\xi|(|\eta\rangle) := \langle\xi \,|\, \eta\rangle.$$

Here $\langle\xi \,|\, \eta\rangle$ denotes the inner product. We require the inner product be conjugate linear in the first coordinate. We will write rank one operators like $|\eta\rangle\langle\xi|$. This sends $|\psi\rangle \mapsto |\eta\rangle\langle\xi \,|\, \psi\rangle$ as notation suggest.

$\|-\|_op$ will denote the operator norm on $B(\mathcal{H})$. $\|-\|_1$ will denote the trace norm, i.e $\|A\|_1 = \mathrm{Tr}(|A|)$ on $B(\mathcal{H})$ for finite dimensional Hilbert space $\mathcal{H}$.

1.2 Quantum Information

Since I am writing this for mathemeticians, I will quickly write down the quantum information framework.

A **state** is an element of $B((\mathbb{C}^2)^{\otimes n})$ that is positive and has trace 1. A **pure state** is a state of rank one (these are also precisely the extreme points of the set). These are all of the form $|\psi\rangle\langle\psi|$ for some unit vector $|\psi\rangle \in (\mathbb{C}^2)^{\otimes n}$. In this paper, when we say state we will mainly mean pure states, and instead of writing the corresponding matrix we will usually write just $|\psi\rangle \in (\mathbb{C}^2)^{\otimes n}$ to denote it. A general state then can be written as a convex combination of pure states, and will be thought of as a probabilistic mixture of the pure states.

We will label the standard basis of $\mathbb{C}^2$ as $|0\rangle$ and $|1\rangle$. For $x \in \mathbb{Z}_2^n$, we will also define

$$|x\rangle := |x_1\rangle \otimes \ldots |x_n\rangle \in (\mathbb{C}^2)^{\otimes n}.$$

This is also an orthonormal basis of $(\mathbb{C}^2)^{\otimes n}$, and will be called the **standard basis**.

A quantum operation is a **quantum channel** $\Phi : B((\mathbb{C}^2)^{\otimes n}) \to B((\mathbb{C}^2)^{\otimes m})$. These are completely positive trace preserving maps. Completely positive means that for each k,

$$\Phi_k : M_k(B((\mathbb{C}^2)^{\otimes n})) \to M_k(B((\mathbb{C}^2)^{\otimes m}))$$

is positive, i.e sends positive elements to positive elements. Here Φ_k is defined as applying Φ to

each entry of the matrix. An equivalent way to say this would be for each hilbert space $\mathcal{H}$,

$$\Phi \otimes I_{\mathcal{H}} : B((\mathbb{C}^2)^{\otimes n}) \otimes B(\mathcal{H}) \to B((\mathbb{C}^2)^{\otimes m}) \otimes B(\mathcal{H})$$

is positive. This makes sense, as we asked for states to be positive matrices, and adding extra qubits should not suddenly make a quantum operation not give a state.

A quantum operation on a state can be estimated arbritarily well by **quantum circuits**[Wat08, section III]. A quantum circuit consists of two parts, the first is a unitary $U : (\mathbb{C}^2)^{\otimes n} \to (\mathbb{C}^2)^{\otimes n}$, that will send a state $\rho \mapsto U\rho U^*$. The second part is measurement. we can write an arbritary state as:

$$|\psi\rangle = \sum_{x \in \mathbb{Z}_2^n} \alpha_x |x\rangle.$$

A measurement of $|\psi\rangle$ in the standard basis will give x with probability $|\alpha_x|^2$. We say that the state has collapsed to $|x\rangle$. One way to think about this is if Π_x is the projection onto the subspace generated by $|x\rangle$, then the probability of measuring x is:

$$\langle\psi|\Pi_x|\psi\rangle = \mathrm{Tr}(\Pi_x|\psi\rangle\langle\psi|).$$

We can also do measurements on only a part of the state. I.e lets say we have a pure state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ and let $\{|x\rangle\}$ be a orthonormal basis of $\mathcal{H}_A$. We can then write

$$|\psi\rangle = \sum_x \alpha_x |x\rangle \otimes |\psi_x\rangle$$

for some states $|\psi_x\rangle \in \mathcal{H}_B$. Measuring register A will yield x with probability $|\alpha_x|^2$, and the state will collapse to $|\psi_x\rangle$. Note that this is a probabilistic mixture of pure states, i.e a mixed state. So the post measurement state is

$$\sum_x |\alpha_x|^2 |\psi_x\rangle\langle\psi_x|.$$

There is a nice way to write this using partial trace:

Definition 1.2.1. *For finite dimensional Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B$ we define the* ***partial trace***

$$\mathrm{Tr}_A : B(\mathcal{H}_A \otimes \mathcal{H}_B) \to B(\mathcal{H}_B) \quad \mathrm{Tr}_A := \mathrm{Tr} \otimes \mathrm{id}\,.$$

Note that under this notion,

$$\mathrm{Tr}_A(|\psi\rangle\langle\psi|) = \sum_x |\alpha_x|^2 |\psi_x\rangle\langle\psi_x|.$$

We can extend this to mixed states too. So for a state $\rho \in B(\mathcal{H}_A \otimes \mathcal{H}_B)$, measuring register A gives post measurement state $\mathrm{Tr}_A(\rho)$.

We will make use of the following lemma in the main body of this paper:

Lemma 1.2.2. *Take a state $\rho \in B(\mathcal{H}_A \otimes \mathcal{H}_B)$. Let $M_B \in B(\mathcal{H}_B)$. Then*

$$\mathrm{Tr}(\rho I_A \otimes M_B) = \mathrm{Tr}(\mathrm{Tr}_A(\rho) M_B).$$

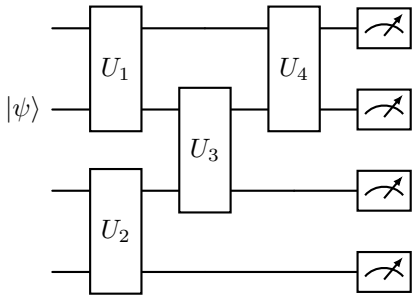
One way to think about this is as follows, if M_B is some projection, then $\mathrm{Tr}(\rho I_A \otimes M_B)$ is basically measuring the probability of some result on a measurment. The lemma is saying that measuring ρ on only register B is the same is first measuring it in register A and then in register B. I.e order of measurements do not matter.

Proof. Let $\{|a\rangle\}$ be a orthonormal basis of $\mathcal{H}_A$ and $\{|b\rangle\}$ of $\mathcal{H}_B$. Things of the form $|ab\rangle\langle a'b'|$ form a basis of $B(\mathcal{H}_A \otimes \mathcal{H}_B)$. Now note

$$\mathrm{Tr}(|ab\rangle\langle a'b'| I_A \otimes M_B) = \langle a'b'| I_A \otimes M_B |ab\rangle = \langle a' \,|\, a\rangle \langle b'| M_B |b\rangle = \mathrm{Tr}(\mathrm{Tr}_A(|ab\rangle\langle a'b'|) M_B).$$

By linearity this identity holds for everything. □

We can restrict the quantum circuit above to only using 2-qubit unitaries[Bar+95]. I.e $U = U_T \dots U_1$ where each U_i acts on only 2 qubits. By grouping together these 2 qubit unitaries that do not interact, we can write $U = U'_d \dots U'_1$ where each U'_j consists of non-interacting 2 qubit unitaries. We say d is the depth of this circuit.



The figure above is a circuit acting on 4 qubit, of depth 3. U_1 and U_4 are acting on qubits $1, 2$, U_2 on qubits $3, 4$ and U_3 on qubits $2, 3$. The meters at the end indicate that we are measuring out.

There is a nice characterization of quantum channels:

Theorem 1.2.3 (Krauss representation)**.** *Let $\Phi : B(\mathcal{H}_A) \to B(\mathcal{H}_B)$ be a a quantum channel. Then there are maps $K_i : \mathcal{H}_A \to \mathcal{H}_B$ so that*

$$\Phi(X) = \sum_i K_i X K_i^*$$

and $\sum_i K_i^* K_i = 1_A$.

Based on the description of how to measure in the computational basis, its easy to see how to do a projective measurement. I.e take projections on $(\mathbb{C}^2)^{\otimes n}$, say $\{\Pi_i\}$ with $\sum \Pi_i = 1$. $|\psi\rangle$ is defined measures to i with respect to the projective valued measurement (PVM) $\{\Pi_i\}$ with probability $\langle\psi|\Pi_i|\psi\rangle$. Basically, pick an appropriate orthonormal basis corresponding to subspaces of Π_i, then we can use unitaries that turn the standard $|x\rangle$ basis into them. So measuring $|\psi\rangle$ modified with this unitary will replicate this.

By a theorem of Naimark, we can extend this. We define a PVM as a collection of positive operators $\{A_i\}$ with $\sum A_i = 1$. $|\psi\rangle$ is defined measures to i with respect to the positive operator valued measurement (POVM) $\{A_i\}$ with probability $\langle\psi|A_i|\psi\rangle$. This can be replicated efficiently with circuits too due to Naimarks theorem, which says that these are unitary dilations of PVMs on a larger hilbert space[Wat18].

We will use the following lemma's in the main body:

Lemma 1.2.4 (Gentle measurement lemma)**.** *[Wat18] Let $\rho \in B(\mathcal{H})$ be a state and let Π be a projection on $\mathcal{H}$. If ρ has high measure with Π, i.e*

$$\mathrm{Tr}(\rho\Pi) \geq 1 - \varepsilon,$$

then the post-measurement state of ρ if we measure Π:

$$\rho' := \frac{\Pi\rho\Pi}{\mathrm{Tr}(\rho\Pi)}$$

is close to ρ:

$$\|\rho - \rho'\|_1 \leq 2\sqrt{\varepsilon}.$$

2 Introduction

2.1 Complexity classes

We will use the promise problem formulation of complexity classes. Let $\{0,1\}^*$ denote the set of all binary strings and $|x|$ denote the length of a string $x \in \{0,1\}^*$.

Definition 2.1.1. *Take $A = (A_{yes}, A_{no})$ with disjoint $A_{yes}, A_{no} \subset \{0,1\}^*$. We say this is a **promise problem**. We call all strings in A_{yes} a yes instance, and everything in A_{no} a no instance. We say A is a **language** if $A_{yes} \cup A_{no} = \{0,1\}^*$.*

Basically the idea is we are promising an input is in $A_{yes} \cup A_{no}$, so our algorithm doesnt need to worry about anything not in these. In the local Hamiltonian problem in the next section, we are promising that the lowest eigenvalue is not in a given range of (a, b). A promise problem is called a language if $A_{yes} \cup A_{no} = \{0,1\}^*$.

We will be using the standard definition of a turing machine, given in [AB09]. Recall the definition of the class NP:

Definition 2.1.2. *A promise problem $A = (A_{yes}, A_{no})$ is said to be in* NP *if there is a polynomial q and a polynomial time turing machine V so that:*

1. *If $x \in A_{yes}$, then there exists $c \in \{0,1\}^{p(|x|)}$ so that $V(x,c) = 1$.*
2. *If $x \in A_{n0}$, for all c we have $V(x,c) = 0$.*

These are problems whose solutions are easy to verify (but not-necessarily easy to solve). We should think of it as a prover sends a proof c, and the verifier uses c to verify that $x \in A_{yes}$. NP stands for non-deterministic polynomial time, as these can be modeled also by non-deterministic Turing machines [AB09].

One should think of the prover as adversarial: the prover will always try to make the verifier accept the proof. In a yes instance this is fine, but in a no instance, one has to take care to make sure the prover isn't lying.

Example 2.1.3. $3SAT$ is the language whose yes instance consists of 3CNF formulas with satisfying assignments. A formula $\phi : \{0,1\}^n \to \{0,1\}$ is said to be $3CNF$ if it can be written as:

$$\phi(x_1 \dots x_n) = \bigwedge_{i \leq m} \left(x_{i_1}^{e_{i_1}} \vee x_{i_2}^{e_{i_2}} \vee x_{i_3}^{e_{i_3}}\right).$$

where $e_j = \pm 1$, and $x_j^+ = x_j, x_j^- = \neg x_j$.

This problem is in NP. Indeed, if such a formula was satisfiable, then let $(x_1 \dots x_n)$ be the certificate, then the verifier can simply calculate ϕ on this input and verify that ϕ has a satisfying assignment. If there was no satisfying assignment, then all certificates would evaluate to 0.

Actually this is in a sense the hardest NP problem. This is the Cook-Levin theorem 2.2.5.

To make a quantum version of this first we would want a notion of Quantum polynomial time algorithm. One might be tempted to say a quantum polynomial algorithm is that which can be done with polynomial size circuits. However, the issue is the circuit at n doesnt need to be related to that of $n-1$, and so this is a huge class. In particular even classically, the class of problems solvable with polynomial size circuits has undecidable problems [AB09, sec 6.1]. The trick to fix this is to force some uniformity on the circuit:

Definition 2.1.4. *A promise problem $A = (A_{yes}, A_{no})$ is said to be in* BQP *(Bounded-error quantum polynomial time) if there is a polynomial time algorithim C that on input x outputs a description of a quantum circuit C_x and:*

1. *If $x \in A_{yes}$, then running C_x with all qubits initialized to $|0\rangle$ and measuring the first qubit gives 1 with probability $\geq 2/3$.*
2. *If $x \in A_{no}$, then running C_x with all qubits initialized to $|0\rangle$ and measuring the first qubit gives 1 with probability $\leq 1/3$.*

Note this is actually a quantum version of BPP, the class of problems that can be solved with high probability by a polynomial time random turing machine. This is because there is an inherent probabilistic measure to quantum: measuring a qubit will give for example 0 with some probability rather than always give 0 or 1.

We can define the quantum version of NP, or really of MA(Merlin-Arthur, Merlin sends proofs and Arthur verifies them), the probabilistic version of NP:

Definition 2.1.5. *A promise problem $A = (A_{yes}, A_{no})$ is said to be in* QMA(c, s)*(Quantum Merlin Arthur) if there is a polynomial q and a polynomial time turing machine that on input x gives a description of the quantum circuit C_x so that:*

1. *If $x \in A_{yes}$, then there exists $|\xi\rangle \in (\mathbb{C}^2)^{\otimes q(|x|)}$ so that $C_x(|0\rangle^{\otimes n} \otimes |\xi\rangle)$ measures to 1 in the first qubit with probability $\geq c(|x|)$*
2. *If $x \in A_{no}$, then for each state $|\xi\rangle$, $C_x(|0\rangle^{\otimes n} \otimes |\xi\rangle)$ measures to 1 in the first qubit with probability $\leq s(|x|)$.*

We will denote $\mathsf{QMA} = \mathsf{QMA}(2/3, 1/3)$.

Remark 2.1.6. *Note in the definitions of* QMA *we asked for a pure state as proof. We could have equally well asked for a mixed state* $\rho \in B((\mathbb{C}^2)^{\otimes q(|x|)})$*, and then the acceptence probability would be* $\mathrm{Tr}(\Pi_1 C_x(|0\rangle\langle 0| \otimes \rho)C_x^*)$ *where* Π_1 *is the projection onto the subspace generated by computational basis having* 1 *in the first qubit.*

Note that we can write ρ *as a convex combination of pure states:*

$$\rho = \sum_i p_i |\phi_i\rangle\langle\phi_i|.$$

Then note that

$$\mathrm{Tr}(\Pi_1 C_x(|0\rangle\langle 0| \otimes \rho)C_x^*) = \sum_i p_i \,\mathrm{Tr}(\Pi_1 C_x(|0\rangle\langle 0| \otimes |\phi_i\rangle\langle\phi_i|)C_x^*).$$

I.e the probability of accepting on ρ *is a convex combination of accepting on some pure states. So for each* ρ*, there is a pure state that the prover can send that makes the prover accept with higher or the same probability as* ρ*. Since the goal of the prover is to maximize this acceptance probability, pure states are enough.*

We will need the following later:

Proposition 2.1.7.

$$\mathsf{QMA}(c, s) = \mathsf{QMA}(2^{-\,\mathrm{poly}}, 1 - 2^{-\,\mathrm{poly}})$$

for any $c - s \geq 1/\,\mathrm{poly}$*,* $0 < s < c < 1$*.*

Proof. Clearly $\mathsf{QMA}(c, s) \supset \mathsf{QMA}(2^{-\,\mathrm{poly}}, 1 - 2^{-\,\mathrm{poly}})$. We will show the other inclusion.

Let $A = (A_{yes}, A_{no})$ be a promise problem in $\mathsf{QMA}(c, s)$ with $c - s \geq 1/\,\mathrm{poly}$. r will be a polynomial to be fixed later. Suppose an instance $x \in \{0, 1\}^n$ of A is verified with the circuit C_x and with proofs of size $p(n)$ for some polynomial p. We will look at $r(n)$ copies of the circuit C_x, i.e $C_x^{\otimes r(n)}$.

The prover will send over a state $|\Xi\rangle \in (\mathbb{C}^2)^{\otimes r(n)p(n)}$ and the circuit $C_x^{\otimes r(n)}$ will have $|\Xi\rangle$ distributed among the proof parts of each block. Measure the first qubit of each block, if the number of ones is greater than S than accept, otherwise reject. S will be chosen later.

For $x \in A_{yes}$, if $|\xi\rangle$ was the proof of the $\mathsf{QMA}(c, s)$ protocol, then let $|\Xi\rangle = |\xi\rangle^{\otimes r(n)}$. Note that each block measures to 1 with probability $\geq s(n)$. So the number of 1s is a binomial distribution,

and by a tail estimate[BLB04], the probability of getting less than S 1s is:

$$\leq \exp\left(-2r(n)\left(c(n) - \frac{S}{r(n)}\right)^2\right).$$

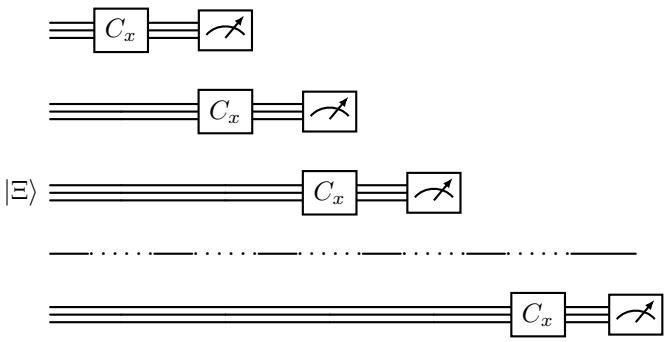
For a $x \in A_{no}$, first suppose the prover sends a state with no entanglement between the different blocks. That is:

$$|\Xi\rangle = |\xi_1\rangle \otimes \cdots \otimes |\xi_{r(n)}\rangle.$$

Note that each block measures to 1 with probability at most c, so the probability of getting more than S 1s is

$$\leq \exp\left(-2r(n)\left(s(n) - \frac{S}{r(n)}\right)^2\right).$$

The issue is the prover can lie and send an entangled state instead. The idea here is to do the circuit sequentially:



The idea is that we measure on block 1 first, get that this returns 1 with probability atmost $s(n)$. Then we do block 2 but now the input is the post measurement state of $|\Xi\rangle$, and get a probability of $s(n)$ again and so on. Formally:

Let Π be the projection onto the subspace generated by computational basis having 1 in the first qubit. We will be compressing any ancilla qubits and the $r(n)$ copies $|0\rangle^n$ for notations sake. Block i will be register i for $1 \leq i \leq r(n)$. The result of the first measurement is 1 with probability:

$$\mathrm{Tr}(C_x^* \Pi C_x \otimes I_{2,3\ldots r(n)} |\Xi\rangle\langle\Xi|) = \mathrm{Tr}(C_x^* \Pi C_x \, \mathrm{Tr}_{2,3\ldots r(n)}(|\Xi\rangle\langle\Xi|)) \leq s(n).$$

Here we used lemma 1.2.2, that if we are measuring only on the first block we can ignore the rest of the block. Ofc Note the second trace is the probability of measuring to 1 if $\mathrm{Tr}_{2,3\ldots r(n)}(|\Xi\rangle\langle\Xi|)$ was the proof state plugged into the circuit, so by soundness it has to be $\leq s(n)$.

Now the post measurement state is simply $\rho_2 = \mathrm{Tr}_1(|\Xi\rangle\langle\Xi|)$, i.e tracing out the first block. We now do the same thing and get the probability of measuring 1 on the second block is

$$\mathrm{Tr}(C_x^*\Pi C_x \otimes I_{3\dots r(n)}\rho_2) = \mathrm{Tr}(C_x^*\Pi C_x \,\mathrm{Tr}_{3\dots r(n)}(\rho_2) \leq s(n).$$

Inductively, we get each block measures to 1 on their first qubit with probability $\leq s(n)$. The Binomial tail bound gives that the probability of getting more than S 1s is

$$\leq \exp\left(-2r(n)\left(s(n) - \frac{S}{r(n)}\right)^2\right).$$

Finally, if we choose $2S = r(n)(c(n) + s(n))$ and $r(n)(c(n) - s(n))^2 \geq 2k(n)$, then we get a new completeness and soundness of:

$$c'(n) = 1 - e^{-k(n)}, \quad s'(n) = e^{-k(n)}.$$

This gives the result □

Note that the completeness-soundness gap being inverse polynomial is important, as that is what allows the number of trails to be polynomial. If the gap was exponentially small, then this class would just be PSPACE[FL16]. Also note that we required $0 < s < c < 1$. Indeed, it is open wether QMA(1), QMA with perfect completeness, is QMA[Aar08].

Basically, the prover can send quantum proves and the verifier can do quantum computations on it. However, we could also easily limit this and ask the prover only send classical proofs:

Definition 2.1.8. *A promise problem $A = (A_{yes}, A_{no})$ is said to be in* QCMA *if there is a polynomial q and a polynomial time turing machine that on input x gives a description of the quantum circuit C_x so that:*

1. *If $x \in A_{yes}$, then there exists $c \in \{0,1\}^{q(x)}$ so that $C_x(|0\rangle^{\otimes n} \otimes |c\rangle)$ measures to 1 in the first qubit with probability $\geq 2/3$*
2. *If $x \in A_{no}$, then for each string c, $C_x(|0\rangle^{\otimes n} \otimes |c\rangle)$ measures to 1 in the first qubit with probability $\leq 1/3$.*

It is widely believed that the inclusions are strict: NP $\subsetneq$ QCMA $\subsetneq$ QMA. There is an oracle seperation between the latter two as some evidence [NN24].

2.2 The PCP Theorem

The PCP (probabilistically checkable proof) theorem says that for any problem in NP, the proofs can be encoded in such a way that the verifier only needs to look at constantly many positions to decide with high probability if the proof is correct or not.

We will need a notion of a probabilistic turing machine, which is given in [AB09] once again. It will be a machine with two tapes, one to work one and one with a random string. It uses $r(n)$ bits of randomness if on input of size n it reads atmost $r(n)$ of the random tape. If on input x, $R(n)$ of the $2^{r(n)}$ random strings made it accept, then we say it accepts with probability $R(n)/2^{r(n)}$.

Definition 2.2.1. *We say a problem $A = (A_{yes}, A_{no})$ is in $\mathsf{PCP}[r(n), q(n)]$ if it has a randomized polynomial time verifier and a polynomial p so that:*

1. *If $x \in A_{yes}$, then there is a $c \in \{0,1\}^{p(|x|)}$ so that V uses $r(|x|)$ bits of randomness and reads $q(|x|)$ bits of the proof c and has $P(V(x,c) = 1) \geq 2/3$.*
2. *If $x \in A_{yes}$, then for all strings c, V uses $r(|x|)$ bits of randomness and reads $q(|x|)$ bits of the proof c and has $P(V(x,c) = 1) \leq 1/3$.*

Note that V can do whatever it wants to x, but only make $q(|x|)$ queries to the proof. The celebrated PCP theorem asserts:

Theorem 2.2.2 (PCP theorem, Proof checking variant)**.** $\mathsf{NP} = \mathsf{PCP}[\log(n), 1]$.

It turns out that a more useful way to formulate this into a quantum statement is by looking at constraint satisfiability problems:

Definition 2.2.3. *A $(m(n), q(n))-$CSP is a family $\mathfrak{C} = (\mathfrak{C}_n)_{n\in\mathbb{N}}$. Each $\mathfrak{C}_n$ consists of m constraints, i.e functions $\{0,1\}^n \longrightarrow \{0,1\}$, that only acts on q bits. We define the value of the CSP as:*

$$\omega(\mathfrak{C}) = \max_{x\in\{0,1\}^n} \frac{\#\{C \in \mathfrak{C}_n : C(x) = 1\}}{m}.$$

An important way to rank problems in a complexity class is by asking if one reduces to the other. We will use the following notion:

Definition 2.2.4. *Let $A = (A_{yes}, A_{no})$ and $B = (B_{yes}, B_{no})$ be promise problems. We say A **polynomial-time reduces** to B if:*

1. *There is a function $f : \{0,1\}^* \to \{0,1\}^*$ so that $f(A_{yes}) \subset B_{yes}$ and $f(A_{no}) \subset A_{no}$,*

2. *There is a polynomial time turing machine that on input $x \in \{0,1\}^*$ computes $f(x)$.*

We say a problem A is (polynomial-time) ***complete*** *for a complexity class* C *if every other problem in* C *polynomial-time reduces to A.*

It is well known that 3-SAT is NP-complete [AB09]. Note being satisfiable or not satisfiabile correspoonds to $\omega = 1$ or $\omega \leq 1 - 1/m$. This means:

Theorem 2.2.5 (Cook-Levin)**.** *For a $(m = \text{poly}(n), 3)$-CSP $\mathfrak{C}$, the problem of determining if*

$$\omega(\mathfrak{C}) = 1 \quad or \quad \omega(\mathfrak{C}) \leq 1 - \frac{1}{m}$$

is NP*-hard.*

Turns out we can weaken this. The a priori problem of deciding between $\omega = 1$ or $\omega < 1/2$ should be just as hard!

Theorem 2.2.6 (PCP theorem, CSP variant)**.** *[Aro+98] The problem of deciding for a $(\text{poly}(n), O(1))$-CSP $\mathfrak{C}$ weather*

$$\omega(\mathfrak{C}) = 1 \quad or \quad \omega(\mathfrak{C}) \leq \frac{1}{2}$$

is NP*-hard.*

It is simple enough to see the connection between the two versions of PCP:

Proposition 2.2.7. *The CSP Variant and the Proof checking variant of the PCP theorem are equivalent.*

Proof. ($\Longleftarrow$). Let $A = (A_{yes}, A_{no})$ be a problem in NP.Let x be an instance. Let V be a PCP verifier of it that has $r(|x|) = O(\log(n))$ bits of randomness and a proof of size $p(|x|) = \text{poly}(n)$ from which it reads $q(|x|) = O(1)$ bits. So the randomized turing machine takes as input a random string $R \in \{0,1\}^{r(|x|)}$, x, and a proof c. It selects $q(|x|)$ bits of c depending on x and R. So define a constraint $C_R : \{0,1\}^{2^{|c|}} \longrightarrow \{0,1\}$ that just does $C_R(y) = V(R, x, y)$. Note this acts only on $q(|x|)$ bits of y. Note also that there are $2^{|r(x)|} = \text{poly}(n)$ constraints. Now:

If $x \in A_{yes}$, then by the conditions of the PCP class, there is some proof that satisfies all the constraints. I.e $\omega = 1$.

If $x \in A_{no}$, then since atmost half the random strings will have the verifier return 1, atmost half the constraints are satisfied for all proofs. So $\omega \leq 1/2$.

So for $(\text{poly}(n), O(1))$-CSPs, it is NP hard tecide between $\omega = 1$ and $\omega \leq 1/2$.

($\Longrightarrow$) Just do the previous construction in reverse. □

Note the CSP variant of the PCP is ripe for generalizing to quantum, as there are quantum CSPs. These are the Local Hamiltonian problems of the next section. We shall see this in the next section.

2.3 Quantum Cook Levin

Definition 2.3.1. *Let* $\mathbf{H} : (\mathbb{C}^2)^{\otimes n} \longrightarrow (\mathbb{C}^2)^{\otimes n}$ *be a linear map, with* $\mathbf{H} = \sum_{i \leq m} \mathbf{H}_i$. *If each* $\mathbf{H}_i$ *is self adjoint and acts as identity on all but* k *of the qubits, then we say* $\mathbf{H}$ *is a* k*-local Hamiltonian.*

We will always assume each $0 \leq \mathbf{H}_i \leq 1$.

The idea is that it is easy to encode $\mathbf{H}$, only m $2^k \times 2^k$ matrices and the position of the qubits describe $\mathbf{H}$. This is opposed to a general matrix which would be a $2^n \times 2^n$ matrix. If we take $k = O(1)$ and let $m = \text{poly}(n)$, then we can encode $\mathbf{H}$ in polynomially many bits as opposed to exponentially many.

Note we could choose to describe this on qudits instead, i.e on local Hamoltonians on $\mathbb{C}^{\otimes n}$. For the purposes of this paper, this will not make a difference so we will just use qubits.

Definition 2.3.2. *The* $k - LH_{a,b}$ *is a promise problem with:*

- ***INPUT****: a* k*-local Hamiltonian* $\mathbf{H}$ *on* n *qubits.*
- ***YES INSTANCE****: There is a unit vector* $|\xi\rangle \in (\mathbb{C}^2)^{\otimes n}$ *with* $\langle \xi | \mathbf{H} | \xi \rangle \leq a(n)$.
- ***NO INSTANCE****: For all unit vectors* $|\xi\rangle \in (\mathbb{C}^2)^{\otimes n}$ *we have* $\langle \xi | \mathbf{H} | \xi \rangle > b(n)$.

Basically the question is: given a Hamiltonian, how hard is it to find its ground state energy (smallest eigenvalue)? From a physical perspective, this is a very sensible question. From a complexity perspective, this is the quantum version of Satisfiability:

Example 2.3.3. Take an instance of 3SAT:

$$\phi(x_1 \dots x_n) = \bigwedge_{i \leq m} \left(x_{i_1}^{e_{i_1}} \vee x_{i_2}^{e_{i_2}} \vee x_{i_3}^{e_{i_3}} \right).$$

We will construct a Hamiltonian problem equivalent to this. I.e we will construct a 3-local Hamiltonian on $(\mathbb{C}^2)^{\otimes n}$ who has a small eigenvalue iff ϕ is satisfiable. For each clause of ϕ, define $\mathbf{H}_i$ to act on qubit i_1, i_2, i_3. If the clause contains a x_j, then it will act on qubit j by $|0\rangle\langle 0|$. If it has $\neg x_j$, then it will act on qubit j by $|1\rangle\langle 1|$. Note that the eigenstates of each $\mathbf{H}_i$ are precisely the computational basis $|x\rangle, x \in \{0,1\}^n$. Moreover, by construction, $\mathbf{H}_i$ will have

eigenvalue 0 on $|x\rangle$ precisely when x satisfies the clause; otherwise the eigenvalue is 1. So the sum $\mathbf{H} = \sum \mathbf{H}_i$ has a 0 eigenvalue precisely if a single x satisfies all the clauses, i.e ϕ. Basically, 3SAT polynomial time reduces to $3 - LH_{0,1}$.

Maybe the most classical result of complexity theory is the cook-levin theorem, which says 3SAT is complete for NP . There is a quantum version of this:

Theorem 2.3.4. *[KKR05]* $2 - LH_{2^{-n},2^{-n}+1/\operatorname{poly}(n)}$ *is* QMA *complete.*

We will prove a weaker version of this.The Quantum PCP conjecture states that $O(1) - LH_{a,b+\gamma m}$ complete for QMA, which is a priori an easier problem. Basically the question is: is it as hard to estimate the ground state energy of a Hamiltonian to constant precision as it is to do it to a inverse polynomial precision?

We will formulate The Quantum PCP conjecture first as a CSP variant. Along the lines of what we previously did we will show the quantum CSPs, i.e local Hamiltonian problem, are QMA-complete. Then by changing the parameters we can state the Quantum PCP conjecture.

First we will show:

Theorem 2.3.5. $O(1) - LH_{a,b}$ *with* $b - a = 1/\operatorname{poly}(n)$ *is in* QMA.

Proof. Let $\mathbf{H} = \sum_{i \leq m} \mathbf{H}_i$ be an instance of the problem acting on n-qubits, i.e a $k = O(1)$ local Hamiltonian and we want to find its ground state to inverse polynomial precision. We can normalize so $0 \leq \mathbf{H}_i \leq 1$.

If the prover sends over state $|\xi\rangle$, then the verifier will first uniformly randomly pick some $\mathbf{H}_i$ and then do a POVM measurement of $|\xi\rangle$ with $\{\mathbf{H}_i, 1 - \mathbf{H}_i\}$. If it measures to $1 - \mathbf{H}_i$, then the verifier will accept. Otherwise, the verifier will reject.

If $\mathbf{H}_i$ was chosen, the probability of measuring $1 - \mathbf{H}_i$ is $1 - \langle\xi|H_i|\xi\rangle$. So the total probability will be

$$\frac{1}{m}\sum_{i=1}^{m}(1 - \langle\xi|H_i|\xi\rangle) = 1 - \frac{1}{m}\langle\xi|H|\xi\rangle.$$

If $x \in A_{yes}$, then there is a state $|\xi\rangle$ that the prover can send with $\langle\xi|H|\xi\rangle \leq a$. So the probability of accepting is $\geq 1 - a/m$.

If $x \in A_{no}$, then all states $|\xi\rangle$ have $\langle\xi|H|\xi\rangle \geq b$. So the probability of accepting is always $\leq 1 - b/m$.

The completeness soundness gap is $(b - a)/m \geq 1/poly$, and by 2.1.7 it is in QMA. □

Note that the gap being inverse polynomial was important here: it is what allowed the completeness-soundness gap of the QMA protocol to be big enough.

We will convert the verifier of any QMA problem into a Hamiltonian. This idea is originally due to Feynmen[KSV03], and we will essentially use that except with a more efficient proof due to [KKR05]. First, we will need a technical lemma:

Lemma 2.3.6. *Let $\mathbf{H} = \mathbf{H}_1 + \mathbf{H}_2$ act on a Finite dimensional hilbert space $\mathcal{H} = \mathcal{S} \oplus \mathcal{S}^\perp$. Suppose $\mathbf{H}_2$ has null space $\mathcal{S}$ and on $\mathcal{S}^\perp$ its eigenvalues are atleast $J > 2\|\mathbf{H}_1\|_{op}$. Then:*

$$\lambda(\mathbf{H}_1|_{\mathcal{S}}) - \frac{\|\mathbf{H}_1\|_{op}^2}{J - 2\|\mathbf{H}_1\|_{op}} \leq \lambda(\mathbf{H}) \leq \lambda(\mathbf{H}_1|_{\mathcal{S}}).$$

Here $\lambda(A)$ denotes the lowest eigenvalue of A. In particular, choosing $J \geq 8\|H_1\|_{op}^2 + 2\|H_1\|$ gives

$$\lambda(\mathbf{H}) \geq \lambda(\mathbf{H}_1|_{\mathcal{S}}) - \frac{1}{8}.$$

Proof. Let $|\eta\rangle \in \mathcal{S}$ be a ground state of $\mathbf{H}_1|_{\mathcal{S}}$. Note $\mathbf{H}_2|\eta\rangle = 0$. So

$$\langle\eta|\mathbf{H}|\eta\rangle = \lambda(\mathbf{H}_1|_{\mathcal{S}}).$$

Now let $|v\rangle$ be an arbitrary unit vector of $\mathcal{H}$, and decompose it as $\alpha_1|v_1\rangle + \alpha_2|v_2\rangle$ according to $\mathcal{S} \oplus \mathcal{S}^\perp$, and suppose α_1, α_2 are positive reals. Now a simple computation gives

$$\begin{aligned}
\langle v|\mathbf{H}|v\rangle &= \langle v|\mathbf{H}_1|v\rangle + \alpha_2^2\langle v_2|\mathbf{H}_2|v_2\rangle \\
&\geq \langle v|\mathbf{H}_1|v\rangle + J\alpha_2^2 \\
&= (1-\alpha_2^2)\langle v_1|\mathbf{H}_1|v_1\rangle + 2\alpha_1\alpha_2\Re(\langle V_2|\mathbf{H}_1|v_1\rangle) + \alpha_2^2\langle V_2|\mathbf{H}_1|v_2\rangle + J\alpha_2^2 \\
&\geq \langle v_1|\mathbf{H}_1|v_1\rangle - K\alpha_2^2 - 2K\alpha_2 - K\alpha_2^2 + J\alpha_2^2 \\
&\geq \lambda(\mathbf{H}_1|_{\mathcal{S}}) + (J-2K)\alpha_2^2 - 2K\alpha_2.
\end{aligned}$$

Plugging in $\alpha_2 = K/(J-2k)$ gives the desired lower bound. □

This will allow us to combine different hamiltonians and still have a good estimate for the ground-state energy. We now prove:

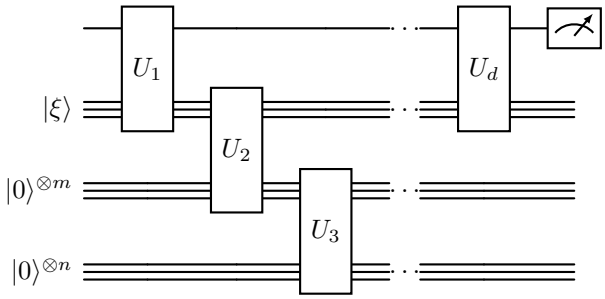
Theorem 2.3.7 (quantum Cook-Levin)**.** *The problem $5 - LH_{a,b}$ with $a(n) = 2^{-\operatorname{poly}(n)}$ and $b(n) = 1/\operatorname{poly}(n)$ is QMA-complete under polynomial time reductions.*

Setup: Let $A = (A_{yes}, A_{no})$ be a QMA problem. For each instance of the problem x, we want to create a local Hamiltonian $\mathbf{H}_x$ whose ground state energy is low precisely when $x \in A_{yes}$.

By the definition of QMA, there is a uniform family of quantum circuits C_x and a polynomial p so that:

- If $x \in A_{yes}$, then there is a state $|\xi\rangle \in (\mathbb{C}^2)^{\otimes p(n)}$ so that $C_x(|\xi\rangle)$ outputs 1 with probability $\geq 1 - \varepsilon$.
- If $x \in A_{no}$, then for each state $|\xi\rangle$, $C_x(|\xi\rangle)$ outputs 1 with probability $\leq \varepsilon$.

Lets say $C_x = U_T \dots U_1$ with U_i 2-local, and it takes as input $N = p(n) + n + m$ qubits. I.e, the input is the $p(n)$ qubits corresponding to $|\xi\rangle$, the n qubits corresponding to x (all initialized to $|0\rangle$), and m qubits for ancilla (also initialized to 0. The output of C_x is the result of measuring the first qubit after applying all the U_i.



Constructing the Hamiltonian: We will define $\mathbf{H}_x \in B((\mathbb{C}^2)^{\otimes n}) \otimes \mathbb{C}^{T+1}$ to be:

$$\begin{aligned}
\mathbf{H}_x &:= \mathbf{H}_{out} + J_{in}\mathbf{H}_{in} + J_{prop}\mathbf{H}_{prop} \\
\mathbf{H}_{out} &:= (T+1)|0\rangle\langle 0|_1 \otimes |T\rangle\langle T| \\
\mathbf{H}_{in} &:= \left(\sum_{i=p(n)+1}^{N} |1\rangle\langle 1|_i \right) \otimes |0\rangle\langle 0| \\
\mathbf{H}_{prop} &:= \sum_{j=1}^{T} (I \otimes (|j\rangle\langle j| + |j-1\rangle\langle j-1|) - U_j \otimes |j\rangle\langle j-1| - U_j^* \otimes |j-1\rangle\langle j|)
\end{aligned}$$

The $\mathbb{C}^{T+1}$ is a counter. So $|\psi\rangle \otimes |i\rangle$ should be thought of as "the circuit is at state $|\psi\rangle$ after applying i of the unitaries".

Essentially, think about each of these terms penalizing unwanted behaviour. For example, $\langle\phi|\mathbf{H}_{out}|\phi\rangle$ is basically measuring the probability of getting 0 on qubit 1 on the time T component of $|\psi\rangle$. We want to punish if we don't get 1 with high probability as the output of the final measurement, and this is exactly what the $\mathbf{H}_{out}$ term does.

The $\mathbf{H}_{in}$ term penalizes if any of the x qubits or the ancilla qubits are not initialized to $|0\rangle$ at step 0.

The individual terms of $\mathbf{H}_{prop}$ is essentially selecting out the step j and $j-1$ terms. It forces

that the state at j be the same as U_j applied to the one at $j-1$. Everytime this is not the case, the energy increases. I.e if

$$|\psi\rangle = \sum_{t=o}^{T} |\psi_t\rangle \otimes |t\rangle,$$

then $\mathbf{H}_{prop}|\Psi\rangle = 0 \iff U_{t+1}|\psi_t\rangle = |\psi_{t+1}\rangle$ for each $0 \leq t < T$. This is easily seen by induction.

Completeness: Suppose $x \in A_{yes}$, and let $|\xi\rangle$ be a certificate for it. Let

$$|\eta\rangle = \frac{1}{\sqrt{T+1}} \sum_{j=0}^{T} U_j \dots U_1 |\xi, 0\rangle \otimes |j\rangle,$$

where $|\xi, 0\rangle$ has $|\xi\rangle$ on the first $p(n)$ qubits and 0 for the x qubits and the ancilla qubits. By what was discussed in the setup, it is clear that

$$\langle\eta|\mathbf{H}_{in}|\eta\rangle = \langle\eta|\mathbf{H}_{prop}|\eta\rangle = 0$$

and

$$\langle\eta|\mathbf{H}_{out}|\eta\rangle = \frac{Pr(C_x(|\xi\rangle) = 0)}{T+1} \leq \frac{\varepsilon}{T+1}.$$

So the energy of this state is $\leq \dfrac{\varepsilon}{T+1}$.

Soundness: We will assume x is a no instance.

We will write $A_1 = J_{in}\mathbf{H}_{in} + \mathbf{H}_{out}$ and $A_2 = J_{prop}\mathbf{H}_{prop}$. Let

$$\mathcal{S}_{prop} := \ker(\mathbf{H}_{prop}).$$

This consists of states for which the circuit propogateed properly as discussed earlier. The idea is now to write $(\mathbb{C}^2)^{\otimes N} = \mathcal{S}_{prop} \oplus \mathcal{S}_{prop}^{\perp}$, and then apply lemma 2.3.6. Note that A_2 disappears on $\mathcal{S}_{prop}$ by definition, so we just want a bound for its smallest eigenvalue on $\mathcal{S}_{prop}^{\perp}$.

Let $W = \sum_{j=0}^{T} U_j \dots U_1 \otimes |j\rangle\langle j|$, this is easily seen to be unitary. A simple computation gives

$$W^*\mathbf{H}_{prop}W = I \otimes E, \quad E = \begin{bmatrix} 1/2 & -1/2 & 0 & 0 & \dots \\ -1/2 & 1 & -1/2 & 0 & \dots \\ 0 & -1/2 & 1 & -1/2 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix}.$$

The eigenvectors/values of E are precisely:

$$|\psi_k\rangle = \sum_{j=0}^{T} \cos\Big(\frac{\pi k}{T+1}\big(j + \frac{1}{2}\big)\Big)|j\rangle, \quad \lambda_k = 1 - \cos\Big(\frac{\pi k}{T+1}\Big)$$

so

$$\lambda_1(\mathbf{H}_{prop}) = \lambda_1(E) = 1 - \cos\Big(\frac{\pi k}{T+1}\Big) \geq \frac{c}{(T+1)^2}.$$

Here λ_1 is the second smallest eigenvalue, so smallest eigenvalue of $\mathbf{H}_{prop}|_{\mathcal{S}_{prop}^{\perp}}$ also.

Note that

$$\|A_1\|_{op} \leq \|\mathbf{H}_{out}\|_{op} + J_{in}\|H_{in}\|_{op} \leq T + 1 + J_{in}N.$$

So by picking $cJ_{prop}/T^2 \geq 8(T+1+J_{in}N)^2 + 2(T+1+J_{in}N)$, lemma 2.3.6 gives that

$$\lambda(H) \geq \lambda(A_1|_{\mathcal{S}_{prop}}) - \frac{1}{8}.$$

Now we will apply the lemma again, but this time on $A_1|_{\mathcal{S}_{prop}}$. Let

$$\mathcal{S}_{in} = \mathcal{S}_{prop} \cap \ker(\mathbf{H}_{in}).$$

I.e states in which the circuit propogates correctly and the qubits are initialized properly. This is spanned by elements of the form

$$\sum_{t=0}^{T} U_t \dots U_0|\xi, 0\rangle \otimes |j\rangle,$$

where the non-proof qubits are initialized to 0. Now by breaking up $\mathcal{S}_{prop} = \mathcal{S}_{in} \oplus \mathcal{S}_{in}^{\perp}$, we can use the lemma.

Note that H_{in} restricted to $\mathcal{S}_{in}^{\perp}$ will pick up atleast one penalty term because atleast one of the qubits is initialized to 1. So $J_{in}H_{in}$ has eigenvalue atleast $J_{in}/(T+1)$. Note also $\|H_{out}|_{S_{prop}}\| \leq T+1$. So once again choosing

$$J_{in}/(T+1) \geq 8(T+1)^2 + 2(T+1)$$

gives

$$\lambda(A_1|_{\mathcal{S}_{prop}}) \geq \lambda(H_{out}|_{\mathcal{S}_{in}}) - 1/8.$$

Finally note that for

$$|\Psi\rangle = \frac{1}{T+1}\sum_{t=0}^{T} U_t \dots U_0|\xi, 0\rangle \otimes |j\rangle \in \mathcal{S}_{in},$$

we get that

$$\langle\Psi|H_{out}|\Psi\rangle = \langle\xi, 0|C_x^*\langle 1|_1|1\rangle_1 C_X|\xi, 0\rangle = Pr(C_X \text{ measures to 1 on qubit 1}) \geq 1 - \varepsilon.$$

And now by combining the two bounds, we get

$$\lambda(H) \geq \frac{3}{4} - \varepsilon.$$

Locality: Note that the Hamiltonian constructed is not acting on just qubits, its also acting on this clock space that is a counter. But this is fixed easily, we define an embedding:

$$\iota : (\mathbb{C}^2)^{\otimes N} \otimes \mathbb{C}^{T+1} \hookrightarrow (\mathbb{C}^2)^{\otimes N} \otimes (\mathbb{C}^2)^{\otimes T}$$

where we embedd $|t\rangle \mapsto |\underbrace{1\dots1}_{t \text{ times}}00\dots0\rangle$.

We can then extend $\mathbf{H}_x$ on this space by setting it to 0 whenever something has a component in the last T qubit not arising from a $|t\rangle$. On a yes instance, we can still find a state with low energy, i.e image of the completeness state we found before under this embedding. The issue is the soundness, indeed by how we defined it, anything orthogonal to the images of the $|t\rangle$ will give 0 energy.

The fix to this is really easy, define:

$$\mathbf{H}_{clock} := \sum_{t=0}^{T-1} |01\rangle\langle 01|_{t,t+1}.$$

This checks if anything in the last T qubits isnt coming from a $|t\rangle$ (those terms' 0s comes after the 1s). So define

$$\mathbf{H}'_x = \mathbf{H}_x + \mathbf{H}_{clock},$$

and note the same state gives this the same low energy in the yes instance. In the no instance, the lemma 2.3.6 gives that

$$\lambda(\mathbf{H}'_x) \geq \lambda(\mathbf{H}_x|_{(\mathbb{C}^2)^{\otimes N} \otimes \mathbb{C}^{T+1}}) - \frac{1}{8} \geq \frac{5}{8} - \varepsilon.$$

Note the locality is atmost 5 now. Also note that we picked $J_{in} = \Theta(T^2)$ and so $J_{prop} = \Theta(T^4)$. So after normalizing, the promise gap is

$$\Theta(1/T^4)\left(\frac{5}{8} - \varepsilon - \frac{\varepsilon}{T+1}\right)$$

and since we can choose ε to be exponentially small, this is $\Theta(1/T^4)$.

2.4 Quantum PCP

We can now state the conjecture:

Conjecture 2.4.1. *$O(1) - LH_{a,b}$ with $b - a = \gamma m$ where m is the number of constraints and $\gamma = O(1)$ is* QMA*-complete.*

All known proofs of Hamiltonian families being complete are built upon in one way or the other the 5-local Hamiltonian being complete as in the previous theorem. That theorem can only ever give inverse polynomial gaps. Furtheremore, the proofs of classical PCP (like considering constraints on expanders) are known not to continue to the Quantum case. We will show the best known partial result in this survey, the NLTS theorem.

3 Quantum Error Correcting Codes

3.1 Classical error correcting codes

We will start with a simple question. Lets say Alice wants to send a message to Bob, and that there is a chance for one of the bits of this message to flip due to error. Can they recover the original data if such an error occurs? The answer is yes, simply take the map

$$\mathbb{Z}_2 \longrightarrow \mathbb{Z}_2^3 \quad x \mapsto (x, x, x)$$

and now if only one bit flips, we can easily just look at the majority and decide what x was.

This introduces the idea of an Error correcting code. We will in particular be looking at linear codes, which are as follows. We will encode k bits of data into n bits by a linear inclusion

$$G : \mathbb{Z}_2^k \hookrightarrow \mathbb{Z}_2^n.$$

We will think of G as a $k \times n$ matrix acting on a row $1 \times k$ vector of $\mathbb{Z}_2^k$ on the right. We call the subspace $\mathcal{C} := G(\mathbb{Z}_2^k)$ the codespace. We say $\mathbf{H}$ is a parity check matrix for $\mathcal{C}$ if $G\mathbf{H}^T = 0$. I.e, a word x is in the codespace $\mathcal{C}$ iff $Hx = 0$. The previous example of the repetition code was a linear code, with

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

Indeed, a vector (x, y, z) in $\mathbb{Z}_2^3$ has $\mathbf{H}(x, , y, z) = (x + y, y + z)$. This is zero precisely when $x = y = z$.

For $c \in \mathbb{Z}_2^n$, define $|c|$ the hamming weight to be the number of 1s in c. We say the linear code $\mathcal{C}$ has **distance**

$$d := \min_{c \in \mathcal{C} - \{0\}} |c|.$$

Basically, the code can detect upto $d - 1$ errors. Imagine if $d - 1$ bits were flipped from code word c to make c'. Then $c' - c$ has weight $d - 1$, and so is not in the codespace. So one could check against the parity matrix hence, that $c' \notin \mathcal{C}$ and conclude an error must have occurred.

If only $\leq (d-1)/2$ errors occurred, then the error is correctable. Indeed, to correct such an error on c', one just needs to find $c \in \mathcal{C}$ such that $|c - c'| \leq (d-1)/2$. Basically the idea is, since the distance between any two codewords is atleast d, balls of radius $(d-1)/2$ around codewords are all disjoint. So we can find a unique codeword which is within $(d-1)/2$ hamming distance of c'.

If e is an error on some codeword $c \in \mathcal{C}$, then we can detect it by applying $\mathbf{H}$:

$$\mathbf{H}(c+e) = \mathbf{H}e.$$

Since for any two distinct errors e, e' of weight less than $(d-1)/2$, $e+e'$ has weight less than d, we have $e+e' \notin \mathcal{C} = \ker(\mathbf{H})$. In particular, $\mathbf{H}e \neq \mathbf{H}e'$. So applying $\mathbf{H}$ uniquely identifies the error, and we call $\mathbf{H}e$ the **Syndrome** of the error e.

So gathering all of these:

Definition 3.1.1. *A classical linear code is a subspace $\mathcal{C} \subset \mathbb{Z}_2^n$. We say $\mathbf{H}$ is a parity check for $\mathcal{C}$ if $\mathcal{C} = \ker(\mathbf{H})$. If it has distance d and dimension k, we say this is a $[n,k,d]$ code.*

Example 3.1.2. The Hamming Code: Take the parity check matrix:

$$\mathbf{H} = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}.$$

The code is $\ker(\mathbf{H}) \subset \mathbb{Z}_2^7$. It's clear a codeword must have atleast 3 1s, and 1110000 is a codeword. So this code can correct any 1 bit-flip error. So this is a $[7,4,3]$ code. An easy computation shows that if e_i is the bit flip error on bit i, then He_i is i in binary.

3.2 The Shor Code

Correcting Quantum errors is much harder than classical ones. For one, the replication code will not work. Indeed, there is no unitary from $(\mathbb{C}^2)^{\otimes n} \longrightarrow (\mathbb{C}^2)^{\otimes 3n}$ that sends all states $|\psi\rangle \mapsto |\psi\rangle^{\otimes 3}$ by the no cloning theorem. And even if this were the case, another hard thing would be measuring. Measuring a qubit will destroy its state, so even if an error is detected it cannot be fixed as the state has collapsed.

Just like linear codes embedded k bits linearly into n bits, a quantum code will do the same for qubits. I.e a quantum code is an isometry

$$U : (\mathbb{C}^2)^{\otimes k} \hookrightarrow (\mathbb{C}^2)^{\otimes n}.$$

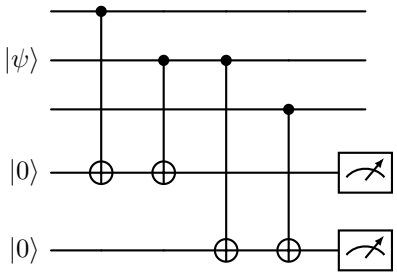
To fix the issue of measurement collapsing states, we will add ancillary qubits, say $|0\rangle^{\otimes \text{Syn}}$. The idea is that we will modify these depending on the original state $|\psi\rangle$, and errors on that will modify what we expect upon measuring these ancillary qubits. The measurement we will call **the syndrome** of the error. This will uniquely identify the error and hence we can correct the

error.

To see this in action, let our map be:

$$\alpha|0\rangle + \beta|1\rangle \mapsto (\alpha|000\rangle + \beta|111\rangle)|00\rangle.$$

Indeed this is a isometry from $(\mathbb{C}^2)^{\otimes 1} \longrightarrow (\mathbb{C}^2)^{\otimes 3} \otimes (\mathbb{C}^2)^{\otimes 2}$ and it will be able to correct bit flips. To see this, apply the map $|abc\rangle|00\rangle \mapsto |abc\rangle|a+b\rangle|b+c\rangle$, and measure the last 2 qubits.



The last two qubits are the ancillary qubits, and their measurement will be the syndrome. It is 00 unless a bit was flipped, 10 if first bit was flipped, 01 if 3rd bit was flipped and 11 if second bit was flipped. We can fix this easily hence.

The issue? we are in quantum world, so a lot more than bitflips can happen. A phase flip $(\alpha|000\rangle + \beta|111\rangle)|00\rangle \mapsto (\alpha|000\rangle - \beta|111\rangle)|00\rangle$ is undetectable. Infact, any one qubit operator can be considered as a possible error. An idea to fix a phase flip would be to write:

$$\alpha|0\rangle + \beta|1\rangle \mapsto \alpha\Big(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\Big)^{\otimes 3} + \beta\Big(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\Big)^{\otimes 3}.$$

Essentially a phase flip flips around $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$ and $|-\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}}$, and so the same idea as before applies. That is, say the error occurs on qubit 1, then the resulting state is $\alpha|-++\rangle - \beta|+--\rangle$. Note $\mathbf{H}|+\rangle = |0\rangle$, $\mathbf{H}|-\rangle = |1\rangle$, where $\mathbf{H}$ is the Hadamard matrix $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. So we can apply $\mathbf{H}^{\otimes 3}$ to make the state just $\alpha|100\rangle - \beta|011\rangle$ and then add two ancillary qubits as before to figure out where the error happened. But a bit flip is no longer protected.

Peter Shor came up with the first quantum error correcting[Sho95] code by combining these ideas to:

$$\alpha|0\rangle + \beta|1\rangle \mapsto \alpha\Big(\frac{|000\rangle + |111\rangle}{\sqrt{2}}\Big)^{\otimes 3} + \beta\Big(\frac{|000\rangle - |111\rangle}{\sqrt{2}}\Big)^{\otimes 3}.$$

One can check this protects against both phase and bit flips, or both. A bit flip can be corrected directly by looking at each group of 3 qubit, and doing the trick mentioned in the first instance

above. The phase flip will flip around a $\frac{|000\rangle+|111\rangle}{\sqrt{2}}$ and $\frac{|000\rangle-|111\rangle}{\sqrt{2}}$ and we can correct this by applying a unitary that sends $\frac{|000\rangle+|111\rangle}{\sqrt{2}} \mapsto |000\rangle$ and $\frac{|000\rangle+|111\rangle}{\sqrt{2}} \mapsto |111\rangle$ and then adding the ancilla to measure a change in these. Turns out this is enough to guarantee every 1-qubit error is fixable, as we will see below.

Note that to fix arbritary t-qubit errors, it suffices to fix Pauli errors. The Pauli matrices are:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}.$$

The idea is X is a bit flip, Z a phase flip and Y a combination of both. These form a basis for $B(\mathbb{C}^2)$, and tensors of I, X, Y, Z form a basis of $B((\mathbb{C}^2)^{\otimes n})$. We call tensors of Pauli matrices also Paulis. Suppose we could fix Pauli errors, i.e

$$P|\psi\rangle|0\rangle^{\text{Syn}} \mapsto P|\psi\rangle|P\rangle$$

under some error correcting operation (like in the Shor code above) and all the different $|P\rangle$ are orthogonal. In this case we can measure the ancilla qubits and decide what error P happened. Since these are a basis, we have a general t-qubit error will be off the form

$$\sum a_P P|\psi\rangle|0\rangle^{\text{Syn}} \mapsto \sum a_p P|\psi\rangle|P\rangle.$$

Now if measuring the ancillary qubit gives syndrome P, then the collapsed state is $P|\psi\rangle$ and now just apply P to get back $|\psi\rangle$.

3.3 General Quantum Code

Definition 3.3.1. *A **quantum code** is an isometry*

$$U : (\mathbb{C}^2)^{\otimes k} \hookrightarrow (\mathbb{C}^2)^{\otimes n}.$$

The image of U we will denote as $\mathcal{Q}$, the code space. An error is some matrix $E \in B((\mathbb{C}^2)^{\otimes n})$. We say a set of errors $\mathcal{E}$ is correctable if there exists a quantum channel $\mathcal{R} : B((\mathbb{C}^2)^{\otimes n}) \longrightarrow B((\mathbb{C}^2)^{\otimes n})$ so that for each codeword $|\psi\rangle \in \mathcal{Q}$ and $E \in \mathcal{E}$:

$$\mathcal{R}(E|\psi\rangle\langle\psi|E^*) \propto |\psi\rangle\langle\psi|.$$

Recall that a quantum channel is a trace Preserving completely positive map, and are precisely the things one can approximate by quantum circuits. This is saying $\mathcal{R}$ is a quantum operation

that can recover $|\psi\rangle$ from $E|\psi\rangle$.

Before characterizing correctable errors, we need a technical lemma:

Lemma 3.3.2. *Let $\mathcal{Q} \subset \mathbf{H}$ be finite dimensional Hilbert spaces, $\Phi : B(\mathbf{H}) \longrightarrow B(\mathbf{H})$ a completely positive map with Kraus representation $\Phi(\rho) = \sum_j K_j \rho K_j^*$. Suppose $\Phi(|\psi\rangle\langle\psi|) \propto |\psi\rangle\langle\psi|$ for each $|\psi\rangle \in \mathcal{Q}$, then $K_j|_{\mathcal{Q}} \propto \mathrm{Id}_{\mathcal{Q}}$.*

Proof. Let $|\psi\rangle \in \mathcal{Q}$ and $|\eta\rangle$ be orthogonal to it. Then

$$0 = \langle \eta \,|\, \psi\rangle\langle \psi \,|\, \eta\rangle \propto \sum_j \langle\eta|K_j|\psi\rangle\langle\psi|K_j^*|\eta\rangle = \sum_j |\langle\eta|K_j|\psi\rangle|^2.$$

In particular, $\langle\eta|K_j|\psi\rangle = 0$.

Note this means for any $|\eta\rangle \in \mathcal{Q}^\perp, |\psi\rangle \in \mathcal{Q}$, we have $\langle\eta|K_j|\psi\rangle = 0$. So $K_j|\psi\rangle \in \mathcal{Q}^{\perp\perp} = \mathcal{Q}$. So now $K_j|_{\mathcal{Q}} \in B(\mathcal{Q})$.

Note for any basis of $\mathcal{Q}$, $K_j|_{\mathcal{Q}}$ is diagonal by the inner product relation. The only operator that does this are scalar multiples of the identity. □

From this we easily get:

Lemma 3.3.3. *Let $\mathcal{E} = \{E_a\}_{a\in A} \subset B((\mathbb{C}^2)^{\otimes n})$ be a set of errors and $\mathcal{Q} \subset (\mathbb{C}^2)^{\otimes n}$ the codespace of some QECC. Let $\mathcal{R} : \rho \mapsto \sum_j R_j \rho R_j^*$ be a channel that corrects the errors. Then*

$$R_j E_a = \lambda_{ja} I$$

for some $\lambda_{ja} \in \mathbb{C}$. Conversely if such a channel exists, then that channel corrects $\mathcal{E}$.

Proof. Note that for $E_A \in \mathcal{E}$, $\rho \longrightarrow \mathcal{R}(E_a \rho E_a^*)$ is a completely positive map satisfying the hypothesis of lemma 3.3.2. Its Kraus representation is

$$\mathcal{R}(E_a \rho E_a^*) = \sum_j R_j E_a \rho (R_j E_a)^*$$

so we get $R_j E_a \propto I$ as desired. The converse is obvious by just computing $\mathcal{R}(E_a|\psi\rangle\langle\psi|E_a^*)$. □

Note that this means in particular if $\mathcal{Q}$ can correct $\mathcal{E}$, then it can correct $\mathrm{Span}(\mathcal{E})$. Now we can classify correctable errors:

Theorem 3.3.4. *Let $\mathcal{E} = \{E_a\}_{a\in A} \subset B((\mathbb{C}^2)^{\otimes n})$ be a set of errors and $\mathcal{Q} \subset (\mathbb{C}^2)^{\otimes n}$ the codespace of some QECC. Suppose $\{|\psi_i\rangle\}_{0\le i<2^k}$ is an orthonormal basis for $\mathcal{Q}$. Then $\mathcal{Q}$ can correct all errors in $\mathrm{Span}(\mathcal{E})$ iff*

$$\langle\psi_j|E_b^* E_a|\psi_i\rangle = C_{ab}\delta_{ij}.$$

The content of this theorem is that C_{ab} should not depend on i, j. Note that this characterization is basis independent. While c_{ab} itself will generally depend on the basis, the inner product will have to take this form.

Think about it like this: first if $i \neq j$ then one can run error correction of $E_b|\psi_j\rangle$ and $E_a|\psi_i\rangle$ and then perfectly distinguish them. This means these states have to be orthogonal. All the data about telling these states apart are in the inner product, and they should not reveal anything about the $|\psi_i\rangle$ if we want to recover $|\psi_i\rangle$, as otherwise our measurement would have destroyed something about it. Turns out this is enough.

Proof. ($\implies$) Let $\mathcal{R} : \rho \mapsto \sum_j R_j \rho R_j^*$ be a correcting channel for $\mathcal{E}$. Then $\sum_j R_j^* R_j = I$ due to the trace preserving condition. Hence we can write:

$$\langle\psi_j|E_b^* E_a|\psi_i\rangle = \sum_j \langle\psi_j|E_b^* R_j^* R_j E_a|\psi_i\rangle = \sum_j \langle\psi_j|\lambda_{bj}^* \lambda_{aj}|\psi_i\rangle = \Big(\sum_j \lambda_{bj}^* \lambda_{aj}\Big)\delta_{ij}.$$

($\impliedby$) Note that $C = [C_{ab}]_{a,b\in A}$ is a self adjoint matrix, and hence can be diagonalized unitarily. Say $C = V^* D V$ for unitary V and diagonal D. Now $F_a = \sum_{a'\in A} V_{aa'} E_{a'}$ has

$$\langle\psi_i|F_b^* F_a|\psi_i\rangle = \sum_{a',b'\in A} \langle\psi_i|\bar{V_{bb'}} E_{b'}^* E_{a'} V_{aa'}|\psi_i\rangle = \sum_{a',b'\in A} V_{aa'} C_{a'b'} \overline{V_{bb'}} = (VCV^*)_{ab}.$$

So the different $F_a|\psi_i\rangle$ are orthogonal.

It is easy to see now if $|\psi\rangle = \sum_i \alpha_i|\psi_i\rangle \in \mathcal{Q}$, then

$$\langle\psi|F_b^* F_a|\psi\rangle = \sum_{i,j} \overline{\alpha}_i \alpha_j \langle\psi_i|F_b^* F_a|\psi_j\rangle = \sum_{ij} \overline{\alpha}_i \alpha_j D_{ab}\delta_{ij} = D_{ab}.$$

So the $F_a|\psi\rangle$ for any codeword $|\psi\rangle$ are orthogonal, so we should be able to perfectly distinguish which error has occured. Note: orthogonal states are distinguishable as one can simply take a unitary that map them to the computational basis and then measure.

The explicit recovery channel is

$$\mathcal{R}(\rho) = \sum_{a\in A} F_a^* \rho F_a.$$

Indeed,

$$\mathcal{R}(F_b|\psi_i\rangle\langle\psi_j|F_b^*) = \sum_a F_a^* F_b|\psi_i\rangle\langle\psi_j|F_b^* F_a.$$

By our previous component, inner producting it against $\langle\psi_k|-|\psi_l\rangle$ only has a non-zero component when $k = i$ and $j = l$. I.e it is $\propto |\psi_i\rangle\langle\psi_j|$. Since these form a basis for $B(\mathcal{Q})$, we are done. $\square$

Example 3.3.5. We can check this for the **Shor code**, which has $|\psi_0\rangle = \left(\frac{|000\rangle+|111\rangle}{\sqrt{2}}\right)^{\otimes 3}$ and $|\psi_1\rangle = \left(\frac{|000\rangle-|111\rangle}{\sqrt{2}}\right)^{\otimes 3}$. We want to verify any weight 1 error is correctable, so we can take $\mathcal{E}$ to be the set of 1-qubit Pauli errors. Note then that $E_b^* E_a$ is two qubit, and $|\psi_j\rangle = P_{ab}|\psi_i\rangle$ where P is a 3-qubit Pauli consisting only of Zs. The Zs act on only one of each group of 3 qubits, so we can choose these to never interact with $E_b^* E_a$. In other words,

$$\langle\psi_j|E_b^* E_a|\psi_j\rangle = \langle\psi_i|P_{ab}E_b^* E_a P_{ab}|\psi_i\rangle = \langle\psi_i|P_{ab}P_{ab}E_b^* E_a|\psi_i\rangle = \langle\psi_i|E_b^* E_a|\psi_i\rangle.$$

The case $i \neq j$ is easily checked.

Now that we can talk about error correction without specifying a recovery channel, we can define distance:

Definition 3.3.6. *Let $\mathcal{Q} \subset (\mathbb{C}^2)^{\otimes n}$ be a k-qubit codespace. Let $|\psi_i\rangle$ be an orthonormal basis of $\mathcal{Q}$. We say it has distance*

$$d := \min\{\mathrm{Wt}(A) : A \in B((\mathbb{C}^2)^{\otimes n}) \text{ and } \langle\psi_j|A|\psi_i\rangle \neq C_A\delta_{ij} \text{ for any constant } C_A\}.$$

We say this code is a $[[n, k, d]]$ code. Here $\mathrm{Wt}(A)$ is the weight of A, i.e the number of qubits its acting on.

Note we can just check against Pauli's as they form a basis for $B((\mathbb{C}^2)^{\otimes n})$. Its clear from Theorem 3.3.4 that if a code has distance d, then it can correct $\lfloor (d-1)/2 \rfloor$-qubit errors.

3.4 Stabilizer Codes

Let $\mathcal{P}_n$ denote the subgroup of $U((\mathbb{C}^2)^{\otimes n})$ generated by the Paulis. Elements of this group are simply tensors of Pauli matrices with a sign from $\pm 1, \pm i$. Note that for $P, P' \in \mathcal{P}_n$, $PP' = \pm P'P$, i.e every 2 element either commute or anti-commute. Define $\hat{\mathcal{P}}_n$ to be $\mathcal{P}_n/\{\pm 1, \pm i\}$ the projective Pauli group. Note that clearly $|\mathcal{P}_n| = 4^{n+1}$, and $|\hat{\mathcal{P}}_n| = 4^n$.

When writing a tensor of Pauli matrices, we will omit the tensors. So for example, XI denotes $X \otimes I$ on $B((\mathbb{C}^2)^{\otimes 2})$. We will also use subscripts to label which qubit an operator works on. So X_2 is the Pauli that does X to the 2nd qubit.

Definition 3.4.1. *Let $S \subset \mathcal{P}_n$ be an abelian subgroup, and $\mathcal{Q} = \{|\psi\rangle \in (\mathbb{C}^2)^{\otimes n} : S|\psi\rangle = |\psi\rangle\}$. We say $\mathcal{Q}$ is a stabilizer code, generated by S.*

Note we can restrict to a commuting subset $S \subset \mathcal{P}_n$ and define $\mathcal{Q}$ interms of S, as the subgroup generated by S would still fix $\mathcal{Q}$. To have non-triviality in this result, we will hence

forth assume the subgroup generated by S should not contain -1 or $\pm i$, as clearly those only fix the 0 vector.

We will use this technical lemma:

Lemma 3.4.2. *Let $S \subset \mathcal{P}_n$ be an abelian subgroup of cardinality 2^k that contains neither -1 or i. Then the Codespace it generates, $\mathcal{Q}(S)$ has dimension 2^{n-k}.*

One way to think of this is like this: S is generated by k Paulis. Each Pauli M splits the space equally into its $+1$ eigenspace and -1 eigenspace. This lemma is saying if we do this recursively, then after adding each Pauli and restricting to the $+1$ eigenspace of that, it is also split in half.

Proof. Define $\Pi := \frac{1}{2^k}\sum_{M\in S} M$. First note that since S is closed under multiplication, for any $M \in S$, $M\Pi|\xi\rangle = \Pi|\xi\rangle$. So the image of Π is contained in $\mathcal{Q}(S)$. For any $|\psi\rangle \in \mathcal{Q}(S)$, clearly Π fixes it, so the image of Π is precisely $\mathcal{Q}(S)$. Also in particular, $\Pi^2 = \Pi$.

Note for a Pauli M, if $M \neq M^*$, then $M^2 = -I$. Since we excluded this in our hypothesis, Π is self adjoint. Hence Π is the projection onto $\mathcal{Q}(S)$. Now

$$\dim(\mathcal{Q}(S)) = \mathrm{Tr}(\Pi) = \frac{1}{2^k}\mathrm{Tr}(I) + \sum_{M\in S-\{I\}} \mathrm{Tr}(M) = 2^{n-k}$$

as desired. □

In particular this means that if $S < S'$, then $\mathcal{Q}(S) > \mathcal{Q}(S')$.

Example 3.4.3. The Shor Code. Note that the Shor code is precisely the subspace of $(\mathbb{C}^2)^{\otimes n}$ fixed by

$$S = \{Z_1Z_2, Z_2Z_3, Z_4Z_5, Z_5Z_6, Z_7Z_8, Z_8Z_9, X_1X_2X_3X_4X_5X_6, X_1X_2X_3X_7X_8X_9\}.$$

Clearly these are independent so it generates a subgroup of size 2^8. So the fixed subspace has dimension 2, and both $\left(\frac{|000\rangle+|111\rangle}{\sqrt{2}}\right)^{\otimes 3}$ and $\left(\frac{|000\rangle-|111\rangle}{\sqrt{2}}\right)^{\otimes 3}$ are fixed by S. This gives the result. This code is $[[9, 1, 3]]$.

Example 3.4.4. The Five Qubit Code. Consider $S \subset \mathcal{P}_5$ with

$$S = \{XZZXI, IXZZX, XIXZZ, ZXIXZ\}.$$

Then this gives a codespace of dimension 2, which one can work out to see the basis looks like:

$$\begin{aligned}|0_L\rangle = \frac{1}{4}(&+|00000\rangle - |00011\rangle + |00101\rangle - |00110\rangle + |01001\rangle + |01010\rangle - |01100\rangle - |01111\rangle \\ &- |10001\rangle + |10010\rangle + |10100\rangle - |10111\rangle - |11000\rangle - |11011\rangle - |11101\rangle - |11110\rangle)\end{aligned}$$

$$|1_L\rangle = \frac{1}{4}(+|11111\rangle - |11100\rangle + |11010\rangle - |11001\rangle + |10110\rangle + |10101\rangle - |10011\rangle - |10000\rangle$$
$$- |01110\rangle + |01011\rangle + |01101\rangle - |01000\rangle - |00001\rangle - |00010\rangle - |00100\rangle - |00111\rangle).$$

We will show now that this code has distance 3, so this is a $[[5, 1, 3]]$ code. I.e this can correct any one qubit error.

Theorem 3.4.5. *Let $\mathcal{Q}$ be a stabilizer code generated by abelian group $S \subset \mathcal{P}_n$. Let $N(S)$ be the subgroup of $\mathcal{P}_n$ of elements that commute with everything in S. Then the distance of $\mathcal{Q}$ is*

$$d = \min\{\mathrm{Wt}(A) : A \in N(S) - S\}.$$

Proof. Take an orthonormal basis $|\psi_i\rangle$ of $\mathcal{Q}$. Then remember from 3.3.6 that

$$d := \min\{\mathrm{Wt}(A) : A \in \mathcal{P}_n \text{ and } \langle\psi_j|A|\psi_i\rangle \neq C_A\delta_{ij} \text{ for some constant } C_A\}.$$

($\leq$) Note if $A \in S$, then $\langle\psi_j|A|\psi_i\rangle = \langle\psi_j \,|\, \psi_i\rangle = \delta_{ij}$. If $A \notin N(S)$, then there is some $M \in S$ so that $MS = -SM$ (as everything in $\mathcal{P}_n$ either commutes or anti-commutes). So

$$\langle\psi_j|A|\psi_i\rangle = \langle\psi_j|AM|\psi_i\rangle = -\langle\psi_j|MA|\psi_i\rangle = -\langle\psi_j|A|\psi_i\rangle,$$

i.e the inner product is 0. So $N(S) - S$ contains the set defining d.

($\geq$) Let $A \in N(S) - S$, and look at the subgroup generated by $S \cup \{A\}$. This fixes precisely a 2^{k-1} subspace of $\mathcal{Q}$ by Lemma 3.4.2, and by doing an eigenbasis decomposition, we can select $|\psi_1\rangle \ldots |\psi_{2^{k-1}}\rangle$ to be fixed by A and the rest to be a -1 eigenvector of A. Now note

$$1 = \langle\psi_i|A|\psi_i\rangle \neq \langle\psi_j|A|\psi_j\rangle = -1, \quad i \leq 2^{k-1}, j > 2^{k-1}.$$

This shows the set defining d is just $N(S) - S$ and gives the result. □

Example 3.4.6. For the five qubit code, note every one and two-qubit Pauli will anti commute with atleast one of $\{XZZXI, IXZZX, XIXZZ, ZXIXZ\}$. Note $IXXIZ$ commutes with everything, and is not in the group generated by these (since such elements have an even number of Zs). So the distance is 3.

One can construct stabilizer codes from Classical codes. For a vector $v \in \mathbb{Z}_2^n$, we will denote

$$X^v = \prod_{i=1}^{n} X_i^{v_i}, \; Z^v = \prod_{i=1}^{n} Z_i^{v_i}.$$

I.e X^v changes 0s to Is and 1s to Xs. E.g $(0, 1, 0, 0, 1) \mapsto IXIIX$. Same for Z^v.

Definition 3.4.7. *Let $\mathbf{H}_X \in M_{(n-k_X)\times n}(\mathbb{Z}_2)$ and $\mathbf{H}_Z \in M_{(n-k_Z)\times n}(\mathbb{Z}_2)$ act as parity check matrices for classical codes $\mathcal{C}_X, \mathcal{C'}_Z$, and lets say $\mathbf{H}_X\mathbf{H}_Z^T = 0$. Define:*

$$S := \{X^v : v \text{ row of } \mathbf{H}_X\} \cup \{Z^v : v \text{ row of } \mathbf{H}_Z\}$$

*Then the stabilizer code generated by $\langle S\rangle$ is called a **CSS code**(named after Calderbank-Shor-Steane) denoted as* $\mathrm{CSS}(\mathbf{H}_X, \mathbf{H}_Z)$.

Note that $H_X H_Z^T = 0$ means every row r in $\mathbf{H}_X$ and r' in $\mathbf{H}_Z$ has $r \cdot r' = 0$. So

$$X^r Z^{r'} = \prod_i X_i^{r_i} Z_i^{r'_i} = \prod_i (-1)^{r_i r'_i} Z_i^{r'_i} X_i^{r_i} = (-1)^{r\cdot r'} Z^{r'} X^r = Z^{r'} X^r.$$

Where we note that X_i^a, Z_i^b commute in every instance except when $a = b = 1$, so $X_i^a Z_i^b = (-1)^{ab} Z_i^b X_i^a$. Basically the rows having even overlap means the corresponding stabilizer will anticommute only on an even number of registers, i.e commute. So this is indeed a stabilizer code.

Definition 3.4.8. *For a classical code $\mathcal{C} \subset \mathbb{Z}_2^n$, define the **dual code** as the code*

$$\mathcal{C}^\perp := \{v \in \mathbb{Z}_2^n : v \cdot c = 0 \forall c \in \mathcal{C}\}.$$

We will call its distance $d^\perp$. Note this is a $[n, n-k, d^\perp]$ code. Also note that if $\mathbf{H}$ is a parity check for $\mathcal{C}$, then $\mathcal{C} = \ker(\mathbf{H})$ and $\mathcal{C}^\perp = \mathrm{image}(\mathbf{H}^T)$.

We can characterize CSS codes:

Theorem 3.4.9. *Let $\mathbf{H}_X$ define a $[n, k_X, d_X]$ code and $\mathbf{H}_Z$ a $[n, k_Z, d_Z]$ code. Then* $\mathrm{CSS}(\mathbf{H}_X, \mathbf{H}_Z)$ *is a $[[n, k_x + k_z - n, d]]$ code with*

$$d = \min_{w \in \mathcal{C}_Z - \mathcal{C}_X^\perp, \mathcal{C}_X - \mathcal{C}_Z^\perp} |w|.$$

Furthermore, elements of $\mathrm{CSS}(\mathbf{H}_X, \mathbf{H}_Z)$ *are spanned by $|z + \mathcal{C}_X^\perp\rangle$ for $z \in \mathcal{C}_Z$ where*

$$|z + \mathcal{C}_X^\perp\rangle = \frac{1}{\sqrt{|\mathcal{C}_X^\perp|}} \sum_{x \in \mathcal{C}_X^\perp} |z + x\rangle.$$

Proof. (i) First we will compute the number of encoded qubits.

Let S be the stabilizer group of the code. Then clearly S is a $\mathbb{Z}_2$-vector space. A X stabilizer and a Z stabilizer can never multiply to the identity, i.e $S = S_X \oplus S_Z$ where S_X are the X stabilizers and S_Z the Z stabilizers. Note that $v \mapsto X^v$ sends the rowspace of $\mathbf{H}$ to S_X

isomorphically, and same for $v \mapsto Z^v$. The rowspace has dimension $n - \dim(\ker(\mathbf{H}_X)) = n - k_x$, so $S = S_X \oplus S_Z$ has size $2^{2n-k_x-k_y}$. Hence by lemma 3.4.2. the code stabilized by it encodes

$$n - (2n - k_x - k_y) = k_x + k_y - n$$

qubits.

(ii) We will now prove that the $|z + \mathcal{C}_X^\perp\rangle$ span $\mathrm{CSS}(\mathbf{H}_X, \mathbf{H}_Z)$. Since $\mathcal{C}_X^\perp$ is the rowspace of $\mathbf{H}_X$, and since $H_Z H_X^T = 0$, we have $\mathcal{C}_X^\perp \subset \ker \mathbf{H}_Z = \mathcal{C}_Z$. Since $|z + \mathcal{C}_X^\perp\rangle = |z' + \mathcal{C}_X^\perp\rangle$ precisely when $z - z' \in \mathcal{C}_X^\perp$, we get that

$$\dim(\mathrm{Span}\{|z + \mathcal{C}_X^\perp\rangle : z \in \mathcal{C}_Z\}) = k_Z - (n - K_x).$$

We claim that $|z + \mathcal{C}_X^\perp\rangle$ are indeed in the codespace. We test against the X generators, i.e for row r of $\mathbf{H}_X$:

$$X^r|z + \mathcal{C}_X^\perp\rangle = \frac{1}{\sqrt{|\mathcal{C}_X^\perp|}} \sum_{x \in \mathcal{C}_X^\perp} X^r|z + x\rangle = \frac{1}{\sqrt{|\mathcal{C}_X^\perp|}} \sum_{x \in \mathcal{C}_X^\perp} |r + z + x\rangle = |z + \mathcal{C}_X^\perp\rangle.$$

Here we used that Xs perform bitflips, and that r is a row of $\mathbf{H}_X$, so in particular is a member of the rowspace $= \mathcal{C}_X^\perp$. We test the Z-stabilizer for a row r' of $\mathbf{H}_Z$:

$$Z^{r'}|z + \mathcal{C}_X^\perp\rangle = \frac{1}{\sqrt{|\mathcal{C}_X^\perp|}} \sum_{x \in \mathcal{C}_X^\perp} Z^{r'}|z + x\rangle = \frac{1}{\sqrt{|\mathcal{C}_X^\perp|}} \sum_{x \in \mathcal{C}_X^\perp} (-1)^{r' \cdot (z+x)}|z + x\rangle = |z + \mathcal{C}_X^\perp\rangle.$$

Here we used that Z are phase flips and that $z + x \in \mathcal{C}_Z$, so that dot product against rows of $\mathbf{H}_Z$ annihalate them. So now dimension counting gives us what we want.

(iii) Take a pauli of $N(S) - S$ of weight

$$t < \min_{w \in \mathcal{C}_Z - \mathcal{C}_X^\perp, \mathcal{C}_X - \mathcal{C}_Z^\perp} |w|,$$

then it is (upto a sign) equal to $P = X^a Z^b$ for some $a, b \in \mathbb{Z}_2^n$ with $|a|, |b| < t$. P commuting with all the X generators means Z^b commutes with all the X generators. The X generators are X^r for rows of $\mathbf{H}_X$, so this means

$$X^r Z^b = (-1)^{r \cdot b} Z^b X^r = Z^b X^r.$$

So $r \cdot b = 0$ for each row of $\mathbf{H}_X$, i.e $H_X b = 0$, so $b \in \mathcal{C}_X$. Since $|b| < t$, this forces $b \in \mathcal{C}_Z^\perp$. Similarly, $a \in \mathcal{C}_X^\perp$. But note that $\mathcal{C}_Z^\perp$ is precisely the rowspace of H_Z and same for $\mathcal{C}_X^\perp$.

This means $X^a Z^b$ is a product of the generators, i.e $P \in S$. This is a contradiction.

Let W be what minimizes

$$\min_{w \in \mathcal{C}_Z - \mathcal{C}_X^\perp, \mathcal{C}_X - \mathcal{C}_Z^\perp} |w|.$$

Without loss of generality take $w \in \mathcal{C}_Z - \mathcal{C}_X^\perp$, then $Z^w \in N(S) - S$. So this is precisely the distance.

□

In particular, note that $\mathrm{CSS}(\mathbf{H}_X, \mathbf{H}_Z)$ doesn't depend on the choice of parity check matrix, so we might as well say $\mathrm{CSS}(\mathcal{C}_X, \mathcal{C}_Z)$. Note also the distance calculated is atleast $\min(d_X, d_Z)$.

Example 3.4.10. Steane Code. Take both $\mathcal{C}_X$ and $\mathcal{C}_Z$ to be the Hamming code from example 3.1.2. That is:

$$\mathbf{H}_X = \mathbf{H}_Z = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}.$$

Note that indeed $H_X H_Z^T = 0$. So by Theorem 3.4.9, we have this is a $[[7, 1, 3]]$ code. One can check by using codewords of the Hamming code that the codewords of the Steane code is:

$$\begin{aligned} |0\rangle_L = \frac{1}{\sqrt{8}}\Big(&|0000000\rangle + |1010101\rangle + |0110011\rangle + |1100110\rangle \\ &+ |0001111\rangle + |1011010\rangle + |0111100\rangle + |1101001\rangle\Big) \\ |1\rangle_L = \frac{1}{\sqrt{8}}\Big(&|1111111\rangle + |0101010\rangle + |1001100\rangle + |0011001\rangle \\ &+ |1110000\rangle + |0100101\rangle + |1000011\rangle + |0010110\rangle\Big) . \end{aligned}$$

3.5 qLDPC codes

For a stabilizer code, we have an analog of the parity check matrix from classical coding theory. Indeed, let $\mathcal{Q}$ be a quantum code and S a generating set for its stabilizers. Then define the code Hamiltonian:

$$H_\mathcal{Q} := \sum_{P \in S} \frac{1 - P}{2}.$$

Note $(1 - P)/2$ is the projection onto the -1 eigenspace of P. So a state $|\psi\rangle$ is in the ground space precisely when it is in the shared $+1$ space of all the Ps, i.e is in $\mathcal{Q}$. So $\mathcal{Q} = \ker(H_\mathcal{Q})$. If we want to apply this to local hamiltonian problems, we would want each parity check $(1 - P)/2$ to have low locality. This would require a qLDPC (quantum low density parity check) code.

Definition 3.5.1. *Let $(\mathcal{C}_n)_{n\in\mathbb{N}}$ be a family of classical codes with parity check matrix $H_n \in M_{k_n\times n}(\mathbb{C})$. If each row and column of H_n has weight $O(1)$, then we say this family is LDPC (low density parity check).*

Definition 3.5.2. *Take a family of stabilizer codes $(\mathcal{Q}_n)_{n\in\mathbb{N}}$, and a distguished set of generators $S^n \subset \mathcal{P}_n$. We say this family is qLDPC if each of the generators have weight $O(1)$, and there are constantly many generators acting on a given qubit.*

The goal is to find good codes, which have recently been discovered in the quantum realm too:

Theorem 3.5.3. *There exist good LDPC codes, i.e $[n, \Theta(n), \Theta(n)]$ codes.*

Theorem 3.5.4. *[PK22] There exist good qLDPC CSS codes, i.e $[[n, \Theta(n), \Theta(n)]]$ codes.*

We will not detail these constructions as they take us too far off field, one thing we will say is that we will be using the Quantum tanner codes of [LZ22] instead of the codes of [PK22]. These codes have the clustering property, which is:

Definition 3.5.5. *Let $\mathcal{Q} = \mathrm{CSS}(\mathcal{C}_X, \mathcal{C}_Z)$ where $\mathcal{C}_X$ is a $[n, k_x, d_x]$ code and $\mathcal{C}_Z$ a $[n, k_Z, d_Z]$ code, then we say it has **clustering for approximate codewords** if there is a $\delta_0, c_1, c_2 > 0$ so that for each $0 < \delta < \delta_0$:*

1. *If $|\mathbf{H}_Z y| \leq \delta(n - k_Z)$ then either $d_{hamm}(y, \mathcal{C}_X^{\perp}) \leq c_1\delta n$ or $\geq c_2 n$.*
2. *If $|\mathbf{H}_X y| \leq \delta(n - k_X)$ then either $d_{hamm}(y, \mathcal{C}_Z^{\perp}) \leq c_1\delta n$ or $\geq c_2 n$.*

We are saying approximate codewords of e.g $\mathcal{C}_Z$ are either close to its subspace $\mathcal{C}_X^{\perp}$ or very far from it. Note by considering e.g $y \in \mathcal{C}_Z - \mathcal{C}_X^{\perp}$, we see that the distance is atleast $c_2 n$.

3.6 Locally testable codes

Classical locally testable codes has rich connections to classical PCP.

Definition 3.6.1. *A linear code $\mathcal{C} = \ker(\mathbf{H})$ defined by parity check matrix $\mathbf{H} \in \mathbb{F}_2^{(n-m)\times n}$ is **locally testable with soundness** ρ if for each $x \in \mathbb{Z}_2^n$:*

$$\frac{1}{m}|Hx| \geq \rho \frac{d_{hamm}(x, \mathcal{C})}{n}$$

Basically, the distance of a word from the codespace is proportional to how many parity checks it failed. So for none-code words, if it is not correctable, the syndrome will atleast inform us how far the word is from a codeword.

There are good classical LDPC with constant soundness, as proven in [PK22]. It turns out good enough qLDPC codes with constant soundness will give good results towards QPCP, so our goal is now to define these.

Let $\mathcal{Q} \subset (\mathbb{C}^2)^{\otimes n}$ be a quantum code correcting some errors $\mathcal{E}$. We want a notion of states being t-weight errors away from $\mathcal{Q}$, so we define:

$$\mathcal{Q}_t := \{E|\phi\rangle \,|\, |\phi\rangle \in \mathcal{Q}\,, \mathrm{Wt}(E) \leq t\,, E \in \mathcal{E}\}.$$

We then define the distance operator:

$$D_{\mathcal{Q}} := \sum_{t \geq 1} t(\Pi_{\mathcal{Q}_t} - \Pi_{\mathcal{Q}_{t-1}}).$$

Essentially, this is breaking up every state into portions that have a wieght t error from a codeword, and then doing a weighted sum. So $\langle\psi|D_{\mathcal{Q}}|\psi\rangle$ measures how many errors, on average, the state differs from a codeword.

Definition 3.6.2. *Let $\mathcal{Q} \subset (\mathbb{C}^2)^{\otimes n}$ be a quantum code defined as the ground space of $\mathbf{H} = \sum_{i \leq m} \Pi_i$, where Π_i are projections. Then we say $\mathcal{Q}$ is locally testable with soundness ρ if*

$$\frac{1}{m}\mathbf{H} \geq \frac{\rho}{n} D_{\mathcal{Q}}$$

For stabilizer codes, note that if $S = \{S_1 \dots S_m\}$ is a generating set, then $\Pi_i = (I - S_1)/2$ will define the code Hamiltonian. We can characterize local testibility here nicely interms of stabilizers, but first we need a lemma:

Lemma 3.6.3. *Let $\mathcal{Q}$ be a stabilizer code, then for each Paulis E, E', either $E\mathcal{Q} = E'\mathcal{Q}$ or they are orthogonal. The $E\mathcal{Q}$ also span the entirety of $(\mathbb{C}^2)^{\otimes n}$. In particular, for each state $|\psi\rangle \in (\mathbb{C}^2)^{\otimes n}$, we can decompose it as*

$$|\psi\rangle = \sum_i E_i|\psi_i\rangle$$

with E_i being a Pauli, $|\psi_i\rangle \in \mathcal{Q}$ and $E_i|\psi_i\rangle$ are all orthogonal to each other.

Proof. Let $S = \{S_1 \dots S_m\}$ be a generating set for $\mathcal{Q}$. Then for any Pauli E we have $S_i E = \omega E S_i$ for $\omega \in \{\pm\}$. So $E\mathcal{Q}$ is some simultanious eigenspace of S. So for two Paulis E, E', either

$E\mathcal{Q} = E'\mathcal{Q}$ or they are orthogonal.

Note that the same arguement as lemma 3.4.2, we get that each simultanous eigenspace is dimension 2^{n-m}. I.e if we care about the eigenspace that has eigenvalue $(-1)^{\lambda(i)}$ for S_i, then the projector onto this is

$$\prod_{i\leq m}\left(\frac{1+(-1)^{\lambda(i)}S_i}{2}\right) = \frac{I}{2^m} + \frac{1}{2^m}\sum_{M\in\langle S\rangle - \{I\}} \pm M.$$

This has trace 2^{n-m}. So there is some unitary that sends $\mathcal{Q}$ to the eigenspace. Since the unitary can be written as a linear combination of Paulis, we get the eigenspace is $\sum_i E_i\mathcal{Q}$. since these are all equal or orthogonal, and by counting dimensions, we get the eigenspace is some $E\mathcal{Q}$ for Pauli E.

Since $(\mathbb{C}^2)^{\otimes n}$ decomposes orthogonally into simultaneous eigenspaces of S, we can choose Paulis $E_1 \dots E_{2^m}$ so that

$$(\mathbb{C}^2)^{\otimes n} = \bigoplus_{i\leq 2^m} E_i\mathcal{Q}$$

is an orthogonal decomposition. This shows the result. □

Theorem 3.6.4 ([AE13])**.** *A stabilizer code $\mathcal{Q}$ generated by stabilizers $S = \{g_1 \dots g_m\}$ is ρ-locally testable iff for every possible error $E \in \mathcal{P}_n$:*

$$\frac{\#\{g \in S \,|\, gE = -Eg\}}{m} \geq \rho\frac{\mathrm{Wt}_{N(S)}(E)}{n}.$$

Here

$$\mathrm{Wt}_{N(S)}(E) := \min_{P\in N(S)} \mathrm{Wt}(EP).$$

Proof. First we will show for Pauli E and non-zero codeword $|\psi\rangle$:

$$\frac{\#\{g \in S \,|\, gE = -Eg\}}{m} \geq \rho\frac{\mathrm{Wt}_{N(S)}(E)}{n} \iff \frac{1}{m}\langle\psi|E^*HE|\psi\rangle \geq \frac{\rho}{n}\langle\psi|E^*D_{\mathcal{Q}}E|\psi\rangle.$$

And then we will use lemma 3.6.3 to patch together these to a general state.

Let $|\psi\rangle \in \mathcal{Q}$ be a codeword, and E a Pauli. Then

$$E^*HE = \frac{1}{2}\sum_{i\leq m} E^*(I - S_i)E = \frac{1}{2}\sum_{i\leq m}(I - E^*S_iE) = \#\{g \in S \,|\, gE = -Eg\}.$$

I.e, if E and S_i commute, the $I - E^*S_iE$ cancels out, and otherwise it gives $2I$. So

$$\langle\psi|E^*HE|\psi\rangle = \#\{g \in S \,|\, gE = -Eg\}\|\psi\|^2.$$

Suppose $E|\psi\rangle$ has non-zero projection onto $\Pi_{\mathcal{Q}_t}$, i.e there is a $A|\eta\rangle$ with $\mathrm{Wt}(A) \leq t$ and $|\eta\rangle \in \mathcal{Q}$ with

$$\langle\psi|E^*A|\eta\rangle \neq 0.$$

Note that we can split up A as Paulis of weight less than t, and hence there is some Pauli P of weight $\leq t$ with

$$\langle\psi|E^*P|\eta\rangle \neq 0.$$

Now if $E^*P \notin N(S)$, then there is some stabilizer S_i that anti-commutes with it. I.e

$$\langle\psi|E^*P|\eta\rangle = \langle\psi|S_i^*E^*PS_i|\eta\rangle = -\langle\psi|E^*P|\eta\rangle.$$

So $E^*P \in N(S)$ and so the smallest $\mathcal{Q}_t$ that contains E has $t = \mathrm{Wt}_{N(S)}(E)$. Now note that since $P\mathcal{Q} = \mathcal{Q}$ for any $P \in N(S)$, we have

$$E|\psi\rangle \in EP\mathcal{Q} \implies E|\psi\rangle \in \mathcal{Q}_{\mathrm{Wt}_{N(S)}(E)}.$$

So we have

$$\langle\psi|E^*D_{\mathcal{Q}}E|\psi\rangle = \mathrm{Wt}_{N(S)}(E)\|\psi\|^2,$$

showing the result.

Note we have already shown the forward implication. For the backwards implication, take an arbitrary state $|\psi\rangle \in (\mathbb{C}^2)^{\otimes n}$ and decompose as in lemma 3.6.3:

$$|\psi\rangle = \sum_i E_i|\psi_i\rangle$$

with E_i Pauli, $|\psi_i\rangle \in \mathcal{Q}$ and the $E_i|\psi_i\rangle$ are orthogonal. Note that if two states are orthogonal, then this remains true even if a projection is applied to both. So:

$$\langle\psi|\mathbf{H}|\psi\rangle = \sum_{i,j}\langle\psi_j|E_j^*HE_i|\psi_i\rangle = \sum_i\langle\psi_i|E_i^*HE_i|\psi_i\rangle = \sum_i\||\psi_i\rangle\|^2\#\{g \in S \,|\, gE_i = -E_ig\}$$

Similarly we get

$$\langle\psi|D_{\mathcal{Q}}|\psi\rangle = \sum_i\||\psi_i\rangle\|^2\,\mathrm{Wt}_{N(S)}(E_i).$$

So by the assumptions if the theorem,

$$\frac{1}{m}\langle\psi|\mathbf{H}|\psi\rangle \geq \frac{\rho}{n}\langle\psi|D_{\mathcal{Q}}|\psi\rangle$$

and since $|\psi\rangle$ was arbitrary, we are done. □

Finally, we will see how this works for CSS codes.

Theorem 3.6.5. *Let $\mathcal{Q} = \mathrm{CSS}(\mathcal{C}_X, \mathcal{C}_Z)$ be a CSS code. If both $\mathcal{C}_X$ and $\mathcal{C}_Z$ has soundness ρ, then $\mathcal{Q}$ has soundness*

$$\min\left(\frac{m_X}{m_X+m_Z}, \frac{m_Z}{m_X+m_Z}\right)\rho,$$

where m_X are the number of parity checks for $\mathcal{C}_X$ and same for m_Z.

Conversely, if $\mathcal{Q}$ has soundness ρ, then both $\mathcal{C}_X$ and $\mathcal{C}_Z$ have soundness ρ.

Proof. ($\implies$) Let P be an arbritary Pauli, and write it as $P = X^a Z^b$ for $a, b \in \mathbb{Z}_2^n$ (ignoring signs). Note P fails to commute with a X generator precisely when the corresponding row doesnt annihalate b. I.e P doesnt commute with $|H_X b|$ generators.

Note that for $c \in \mathcal{C}_X$, we have $d_{hamm}(b, c) = |b+c|$ and $X^b = X^{b+c}X^c$, with $X^c \in S \subset N(S)$. So $d_{hamm}(b, c) \geq \mathrm{Wt}_{N(S)}(X^b)$. So from the local testibility of $\mathcal{C}_X$ we get:

$$\frac{\#\{g \in S_X \,|\, gP = -Pg\}}{m_X} \geq \rho \frac{\mathrm{Wt}_{N(S)}(Z^b)}{n}.$$

where S_X is the x generators, and m_X is the number of X generators. We get the same result for the Z generators, and hence:

$$\begin{aligned}
\frac{\#\{g \in S \,|\, gP = -Pg\}}{m_X+m_Z} &\geq \frac{\rho}{n}\left(\frac{m_X}{m_X+m_Z}\mathrm{Wt}_{N(S)}(Z^b) + \frac{m_Z}{m_X+m_Z}\mathrm{Wt}_{N(S)}(X^a)\right) \\
&\geq \frac{\rho}{n}\min\left(\frac{m_X}{m_X+m_Z}, \frac{m_Z}{m_X+m_Z}\right)(\mathrm{Wt}_{N(S)}(X^a) + \mathrm{Wt}_{N(S)}(Z^b)) \\
&\geq \frac{1}{n}\rho\min\left(\frac{m_X}{m_X+m_Z}, \frac{m_Z}{m_X+m_Z}\right)\mathrm{Wt}_{N(S)}(P)
\end{aligned}$$

($\impliedby$) Just test against X^a for $\mathcal{C}_X$ and Z^b for $\mathcal{C}_Z$. □

In the literature, this lemma is stated wrong. It is stated that if both $\mathcal{C}_X, \mathcal{C}_Z$ have soundness ρ then so does ρ. Here is a simple counterexample: take the Steane code 3.4.10. Note that since the Hadamard code is a $[[7, 4, 3]]$ code, every string is distance 1 from a codeword. Also note that $x = (0, 0, 0, 0, 1, 1, 1)$ fails exactly 1 parity check. So

$$\frac{|Hx|}{3} \geq \frac{\rho}{7}d(x, \mathcal{C}) \implies \rho \leq \frac{7}{3},$$

and by what was discussed, $\rho = 7/3$ works as soundness.

For the Steane code, simply take $P = IIIIXXX$. The generators of the Steane code are

$$\{IIIXXXX, IXXIIXX, XXIIXXI, IIIZZZZ, IZZIIZZ, ZZIIZZI\}.$$

P only anticommutes with $IIIZZZZ$, so if the code had soundness 7/3, we would have

$$\frac{1}{6} \geq \frac{7}{3}\frac{1}{7}\,\mathrm{Wt}_{N(S)}(P) \geq \frac{1}{3}$$

which isnt the case. Note the main issue is that the number of parity checks are different between the classical and quantum case. However, in most cases this is a non-issue, as both m_X and m_Z are typically linear in n, so we only gain a constant factor.

4 The NLTS Thoerem

4.1 Introduction

The goal of QPCP is to show estimating the ground state of local hamiltonians to constant precision is QMA-hard. This means low energy states of the Hamiltonian should be hard. Take an instances of $O(1) - LH_{a,a+\varepsilon}$, and let $|\psi\rangle$ be a low energy state $\langle\psi|\mathbf{H}|\psi\rangle < a + \varepsilon$. If $|\psi\rangle$ can be prepared by a polynomial depth quantum circuit, then the description of that circuit is a classical witness for a quantum verifier. I.e this instance would be in QCMA. Hence assuming QCMA $\neq$ QMA, QPCP implies:

Conjecture 4.1.1. *There is a family of $O(1)$-local Hamiltonian $(\mathbf{H}^n)_{n\in\mathbb{N}}$ and a $\gamma > 0$ such that for any polynomial p, there is a some N so that for each $n > N$ any quantum circuit that prepares a $|\psi\rangle$ with*

$$\langle\psi|\mathbf{H}^n|\psi\rangle \leq \gamma,$$

has depth atleast $p(n)$.

This is probably not a very fruitful avenue to prove QPCP, as finding circuit lower bounds is ridiculously hard. Even classically the best know circuit depth lower bound for any family of formulas is linear. Actually the best we have been able to do so far is find Hamiltonians whose low energy states require more than constant depth:

Definition 4.1.2. *We say a family of Hamiltonians $(\mathbf{H}^n)_{n\in\mathbb{N}}$ with $0 \leq \mathbf{H}^n \leq 1$ is ε-**NLTS** if for every constant $c > 0$, there is some N so that for each $n > N$ any quantum circuit that prepares a state $|\psi\rangle$ with*

$$\langle\psi|\mathbf{H}^n|\psi\rangle \leq \varepsilon$$

has depth atleast c.

*If a family of Hamitlonians is ε-NLTS for some $\varepsilon > 0$, then we say it is **NLTS**.*

Theorem 4.1.3. *[ABN23] There exist an ε and a family of $O(1)$-local normalized hamiltonians that is ε-NLTS. In particular, the code hamiltonians of any qLDPC code family satisfying the clustering property 3.5.5 will be NLTS.*

The way to show this would be first to note that measurements of states coming from constant depth circuits are well spread out, wheres the ones that are low energy for a qLDPC code will be supported in clusters. Actually a similar method will also show:

Corollary 4.1.4. *[EH17] Let $(\mathcal{Q}_n)_{n\in\mathbb{N}}$ be a sequence of CSS codes with $\mathcal{Q}_n$ being $[[n, k \geq 1, d = \Theta(n)]]$ with soundness $\rho = \Theta(1)$. Then the code Hamiltonian will be NLTS.*

Proof. We will show this code satisfies the clustering property. First take $w \in \mathcal{C}_Z - \mathcal{C}_X^{\perp}$. Then by theorem 3.4.9, we have

$$d(w, \mathcal{C}_X^{\perp}) = \min_{x\in\mathcal{C}_X^{\perp}} |w + x| \geq d.$$

We will suppose $d \geq Cn$.

Let $\delta_0 = \min\{C\rho/2, C/4\}$. By theorem 3.6.5, we know $\mathcal{C}_Z$ has soundness ρ. So if $|H_Z y| \leq \delta m_Z$ for $\delta < \delta_0$, then

$$d(y, \mathcal{C}_Z) \leq \frac{1}{\rho} n\delta.$$

Now there are two option, if the $z \in \mathcal{C}_Z$ that minimizes this distance is in $\mathcal{C}_X^{\perp}$, then

$$d(y, \mathcal{C}_X^{\perp}) \leq \frac{1}{\rho} n\delta.$$

Or else, we would have

$$d(y, \mathcal{C}_X^{\perp}) \geq d(z, \mathcal{C}_X^{\perp}) - d(y, z) \geq \left(C - \frac{\delta_0}{\rho}\right) n.$$

So with parameters $\delta_0, c_1 = 1/\rho, c_2 = C - \delta_0/\rho$ we get this has the clustering property. □

4.2 Circuit Lower Bounds

As mentioned above, circuits with low depth produce distributions that are well spread out:

Theorem 4.2.1. *Let D be the probability distribution on $\mathbb{Z}_2^n$ that corresponds to measuring the output of a quantum circuit in the standard basis. If two sets $S_1, S_2 \subset \mathbb{Z}_2^n$ satisfy $D(S_1), D(S_2) \geq \mu$, then the depth of the curcuit is atleast*

$$\frac{1}{3} \log_2 \left(\frac{d(S_1, S_2)^2}{n \log(2/\mu)} \right)$$

To prove this, we will need some tools. Namely we will construct an approximate groundstate projector for self adjoint matrices bounded by 0 and I.

Lemma 4.2.2. *For every $s \in \mathbb{R}_{\geq 0}$ and $\eta \in (0, 1)$, there is a polynomial $T_{\eta,s}$ of degree $\lceil s \rceil$ such that $T_{\eta,s}(0) = 1$ and*

$$|T_{\eta,s}(x)| \leq 2e^{-2s\sqrt{\eta}} \quad \eta \leq x \leq 1$$

Proof. Let $T_k : \mathbb{R} \longrightarrow \mathbb{R}$ be a Chebyshev polynomial of degree k, i.e $T_k(\cos(x)) = \cos(kx)$. Then

define for $s \in \mathbb{N}$ and $\eta \in (0,1)$ a new polynomial $T_{\eta,s} : [0,1] \longrightarrow \mathbb{R}$ as

$$T_{\eta,s}(x) := \frac{T_{\lceil s \rceil}\left(\frac{2(1-x)}{1-\eta} - 1\right)}{T_{\lceil s \rceil}\left(\frac{2}{1-\eta} - 1\right)}.$$

Note in the range $\eta \leq x \leq 1$, $-1 \leq \frac{2(1-x)}{1-\eta} - 1 \leq 1$. So it is the cosine of a real number, and so the chebyshev polynomial is also the cosine of a real number, i.e absolute value bounded by 1.

Note in the range of $\eta \in (0,1)$, $\frac{2}{1-\eta} - 1$ is > 1. So it is equal to some $\cos(it) = \cosh(t)$ with $t > 0$. I.e

$$\cosh(t) = \frac{e^t + e^{-t}}{2} = \frac{1+\eta}{1-\eta} \implies e^t = \frac{1+\eta+2\sqrt{\eta}}{1-\eta}.$$

Note

$$\frac{1+\eta+2\sqrt{\eta}}{1-\eta} = 1 + 2\sum_{k=1}^{\infty} \eta^{k/2} \geq \sum_{k=0}^{\infty} \frac{2^k \eta^{k/2}}{k!} = e^{2\sqrt{\eta}}.$$

We used that $2 \geq 2^k/k!$. Also note that $\cosh(t) \geq e^t/2$. This means

$$T_{\lceil s \rceil}\left(\frac{2}{1-\eta} - 1\right) = \cosh(\lceil s \rceil t) \geq \frac{1}{2} e^{st} \geq \frac{1}{2} e^{2s\sqrt{\eta}}.$$

Combining these two, we get that

$$|T_{\eta,s}(x)| \leq 2e^{-2s\sqrt{\eta}} \quad \eta \leq x \leq 1.$$

□

Now we can prove the theorem:

Proof of theorem 4.2.1. Let $|\xi\rangle = U|0\rangle^{\otimes m}$ be prepared by a depth t circuit $U : (\mathbb{C}^2)^{\otimes m} \longrightarrow (\mathbb{C}^2)^{\otimes m}$ where $m \geq n$. We will look at the distribution that comes from measuring the first n qubits. Note by the light cone argument [AN22, fact 6], only $2^t n$ qubits of the input effect the resulting distribution. Hence we may assume $m = 2^t n$.

Take the local Hamiltonian

$$G = \sum_{i=1}^{m} U|1\rangle_i\langle 1|_i U^*.$$

Note that $U|1\rangle_i\langle 1|_i U^*$ is a projection, and since by the light cone, U acting on qubit i only changes 2^t qubits, the locality is 2^t. Also note that $U|x\rangle$ for $x \in \mathbb{Z}_2^n$ are eigenvectors with eigenvalue $|x|/m$. So $|\xi\rangle$ is the unique ground state, and the eigenvectors are of the form $j/m, 0 \leq j \leq m$.

Pick disjoint sets $S_1, S_2 \subset \mathbb{Z}_2^n$. with distance $u = d(S_1, S_2)$ and measure $D(S_1), D(S_2) \geq \mu$.

Let $P(x) = T_{s,\eta}(x)$ as in lemma 4.2.2 with $\eta = 1/m$ and $s = u/2^{t+1}$. Then P has degree $\lceil s \rceil$ with

$$P(0) = 1, \quad |P(j/m)| \leq 2\exp\left(-\frac{u}{\sqrt{2^{3t}n}}\right) \quad 1 \leq j \leq m.$$

Note if $s \leq 1$, then $2^t \geq u/2$ and we are already done. So we can assume $s > 1$, so $\lceil s \rceil < 2s = u/2^t$. So if we look at $P(G)$ now, it has locality $\lceil s \rceil 2^t < u$. The reason to do all this is that $P(G)$ is an approximate ground space projector:

$$\| |\xi\rangle\langle\xi| - P(G) \|_{op} \leq 2\exp\left(-\frac{u}{\sqrt{2^{3t}n}}\right).$$

I.e if we spectral decompose $G = 0|\xi\rangle\langle\xi| + \sum_{j=1}^{m} j/m\Pi_j$, then

$$P(G) = |\xi\rangle\langle\xi| + \sum_{j=1}^{m} P(j/m)\Pi_j$$

and the contributions of the non-zero eigenspaces are small.

Let Π_{S_i} be the projection on to the subspace spanned by $|x\rangle, x \in \mathbb{Z}_2^n$ with the first n bits of x belonging to S_i. Then note first that Π_{S_1} and Π_{S_2} are orthogonal as $u > 0$. Note that if $|s_i\rangle$ are standard basis elements in the image of Π_{S_i} as described above, then

$$\langle s_2|P(G)|s_1\rangle = 0.$$

This is because $P(G)$ will change less than u qubits of $|s_1\rangle$. I.e $P(G)|s_1\rangle = \sum|x\rangle$ with each $|x\rangle$ being a distance $< u$ from s_1. Since S_2 has distance u from S_1, these are all orthogonal to $|s_2\rangle$. So now

$$\begin{aligned} D(S_1)D(S_2) &\leq \langle\xi|\Pi_{S_2}|\xi\rangle\langle\xi|\Pi_{S_1}|\xi\rangle \\ &\leq \|\Pi_{S_2}|\xi\rangle\langle\xi|\Pi_{S_1}\|_{op}^2 \\ &\leq \|\Pi_{S_1}\|_{op}^2 \| |\xi\rangle\langle\xi| - P(G)\|_{op}^2 \|\Pi_{S_1}\|_{op}^2 \\ &\leq 4\exp\left(-2\frac{u}{\sqrt{2^{3t}n}}\right) \end{aligned}$$

This gives the bound, as $D(S_1)D(S_2) \geq \mu$. □

This was originally proved in [EH17], but their proof was a lot more complicated than this simplified proof in [ABN23].

4.3 The NLTS theorem

We will now prove the theorem. Suppose we have a CSS code $\mathcal{Q} = \mathrm{CSS}(\mathbf{H}_X, \mathbf{H}_Z)$ satisfying the hypothesis of theorem 4.1.3. That is, there are $\delta_0, c_1, c_2 > 0$ so that for each $0 < \delta < \delta_0$:

1. If $|\mathbf{H}_Z y| \leq \delta(n - k_Z)$ then either $d(y, \mathcal{C}_X^{\perp}) \leq c_1 \delta n$ or $\geq c_2 n$.
2. If $|\mathbf{H}_X y| \leq \delta(n - k_X)$ then either $d(y, \mathcal{C}_Z^{\perp}) \leq c_1 \delta n$ or $\geq c_2 n$.

Let $\mathbf{H}$ be its code Hamiltonian.

Note our Hamiltonian will be unnormalized. Let $|\psi\rangle$ denote a low energy state

$$\langle\psi|\mathbf{H}|\psi\rangle \leq \varepsilon n.$$

Let D_Z be the measure on $\mathbb{Z}_2^n$ corresponding to measuring the state in the standard basis, and D_X the measure corresponding to measuring it in the X-basis (apply n Hadamards and then measure). we will show for each low energy state, either D_x or D_z are clustered as in Theorem 4.2.1, and hence the circuit must be $\Theta(\log(n))$.

We will need an uncertainty lemma to show that either D_Z or D_X are highly concentrated, so we can use the circuit lower bound of theorem 4.2.1.

Lemma 4.3.1. *Let $S, T \subset \mathbb{Z}_2^n$ and let D_X, D_Z be the distributions coming from measuring a pure state $|\psi\rangle$. Then*

$$D_X(T) \leq 2\sqrt{1 - D_Z(S)} + \sqrt{|S| \cdot |T|/2^n}.$$

Proof. Let $|\psi\rangle = \sum_{y \in \mathbb{Z}_2^n} \alpha_y |y\rangle$. Then $D_Z(S) = \sum_{y \in S} |\alpha_y|^2$. Let

$$|\psi'\rangle := \frac{1}{\sqrt{D_Z(S)}} \sum_{y \in S} \alpha_y |y\rangle = \sum_{y \in \mathbb{Z}_2^n} \beta_y |y\rangle.$$

Then note that by the gentle measurement lemma:

$$\| |\psi\rangle\langle\psi| - |\psi'\rangle\langle\psi'| \|_1 \leq 2\sqrt{1 - D_Z(S)}.$$

Note that if we want to measure $|\psi'\rangle$ in the X basis, we will be measuring

$$H^{\otimes n}|\psi'\rangle = \frac{1}{\sqrt{2^n}} \sum_{a \in \mathbb{Z}_2^n} \sum_{y \in \mathbb{Z}_2^n} (-1)^{a \cdot y} \beta_y |a\rangle$$

in the standard basis. That is, the probability of measuring string x is

$$p(x) = \frac{1}{2^n} \left| \sum_{y \in \mathbb{Z}_2^n} (-1)^{x \cdot y} \beta_y \right|^2 = \frac{1}{2^n} \sum_{z, w \in \mathbb{Z}_2^n} (-1)^{x \cdot (z+w)} \beta_z \overline{\beta_w}$$

Now we will compute:

$$\begin{aligned}
\sum_{x\in\mathbb{Z}_2^n} p(x)^2 &= \frac{1}{2^{2n}} \sum_{x\in\mathbb{Z}_2^n}\sum_{x\in\mathbb{Z}_2^n}\sum_{s,t,z,w\in\mathbb{Z}_2^n} (-1)^{x\cdot(z+w+s+t)}\beta_z^*\beta_w\beta_t^*\beta_s \\
&= \frac{1}{2^n} \sum_{\substack{s+t+w+z=0 \\ s,t,z,w\in\mathbb{Z}_2^n}} \beta_z^*\beta_w\beta_t^*\beta_s \\
&= \frac{1}{2^n} \sum_{z,w\in\mathbb{Z}_2^n} \beta_z^*\beta_w \sum_{t\in\mathbb{Z}_2^n} \beta_t^*\beta_{z+w+t} \\
&\leq \frac{1}{2^n} \sum_{z,w\in\mathbb{Z}_2^n} |\beta_z||\beta_w| \left(\sqrt{\sum_{w\in\mathbb{Z}_2^n}|\beta_t|^2}\sqrt{\sum_{w\in\mathbb{Z}_2^n}|\beta_{s+t+w}|^2}\right) \\
&= \frac{1}{2^n} \sum_{z,w\in\mathbb{Z}_2^n} |\beta_z||\beta_w| \\
&= \frac{1}{2^n} \left(\sum_{z\in\mathbb{Z}_2^n} |\beta_z|\right)^2 \\
&\leq \frac{1}{2^n}|S| \sum_{z\in\mathbb{Z}_2^n} |\beta_z|^2 = \frac{|S|}{2^n}.
\end{aligned}$$

Last line used Cauchy-Schwartz and the fact that only $|S|$ of the β are non-zero. So now,

$$\sum_{x\in T} p(x) \leq \sqrt{|T| \sum_{x\in\mathbb{Z}_2^n} p(x)^2} \leq \sqrt{|T|\cdot|S|/2^n}.$$

And hence, we can finally compute:

$$\begin{aligned}
D_X(T) &= \sum_{y\in T} |\langle y|H^{\otimes n}|\psi\rangle|^2 \\
&= \sum_{y\in T} \langle y|H^{\otimes n}(|\psi\rangle\langle\psi| - |\psi'\rangle\langle\psi'|)H^{\otimes n}|y\rangle + \sum_{y\in T}\langle y|H^{\otimes n}|\psi'\rangle\langle\psi'|H^{\otimes n}|y\rangle \\
&\leq \||\psi\rangle\langle\psi| - |\psi'\rangle\langle\psi'|\|_1 + \sum_{y\in T} |\langle y|H^{\otimes n}|\psi'\rangle|^2 \\
&= \||\psi\rangle\langle\psi| - |\psi'\rangle\langle\psi'|\|_1 + \sum_{y\in T} p(y)^2 \\
&\leq 2\sqrt{1 - D_Z(S)} + \sqrt{|S|\cdot|T|/2^n}
\end{aligned}$$

□

Let

$$G_X^\delta := \{y \in \mathbb{Z}_2^n : |H_X y| \leq \delta m_x\}, \quad G_Z^\delta := \{y \in \mathbb{Z}_2^n : |H_Z y| \leq \delta m_Z\}.$$

I.e strings with low X and Z errors respectively. Then D_X is almost entirely supported on $G_X^{O(\varepsilon)}$

and same for D_Z.

Lemma 4.3.2. *Let* $\varepsilon_1 = \frac{200n}{\min\{m_X, m_Z\}}\varepsilon$. *Then*

$$D_Z(G_Z^{\varepsilon_1}), D_X(G_X^{\varepsilon_1}) \geq \frac{199}{200}.$$

Proof. Split up $\mathbf{H} = \mathbf{H}^Z + \mathbf{H}^X$, where $\mathbf{H}^Z$ is the part of the Hamiltonian corresponding to the Z stabilizers and $\mathbf{H}^X$ the part from the X stabilizers. Note that

$$\varepsilon n \geq \langle\psi|\mathbf{H}|\psi\rangle \geq \langle\psi|\mathbf{H}^Z|\psi\rangle = \mathop{\mathbf{E}}_{y\sim D_Z}|\mathbf{H}_Z y|.$$

This is because if $|\psi\rangle = \sum \alpha_y|y\rangle$, then

$$\begin{aligned}\langle\psi|\mathbf{H}_Z|\psi\rangle &= \sum_{a\in \mathrm{Row}(\mathbf{H}_Z)} \sum_{y,y'\in\mathbb{Z}_2^n} \alpha_y \alpha_{y'}^* \langle y'|\frac{1-Z^a}{2}|y\rangle \\ &= \sum_{a\in \mathrm{Row}(\mathbf{H}_Z)} \sum_{y\in\mathbb{Z}_2^n} |\alpha_y|^2 (a\cdot y) \\ &= \sum_{y\in\mathbb{Z}_2^n} |\alpha_y|^2 |\mathbf{H}_Z y|.\end{aligned}$$

Now let $q = D_Z(G_Z^{\varepsilon_1})$. Outside of $G_Z^{\varepsilon_1}$, $|\mathbf{H}_Z y|$ is $\geq \varepsilon_1 m_Z$. So

$$\varepsilon n \geq \mathop{\mathbf{E}}_{y\sim D_Z}|\mathbf{H}_Z y| \geq 0(q) + \varepsilon_1 m_Z(1-q),$$

and hence $q \geq 1 - \frac{\varepsilon n}{\varepsilon_1 m_X} \geq \frac{199}{200}$.

The same arguement works verbatim for the X measurements, giving the result. □

Now we will prove the theorem by showing

Lemma 4.3.3. *Take a* $[[n, k \geq 4, d]]$ *CSS code* $\mathrm{CSS}(\mathbf{H}_X, \mathbf{H}_Y)$ *satisfying the cluster property 3.5.5 with parameters* δ, c_1, c_2. *Then there are sets* $S_1, S_2 \in \mathbb{Z}_2^n$ *with* $d(S_1, S_2) \geq c_2 n$ *such that either*

$$D_X(S_1), D_X(S_2) \geq \frac{1}{400} \quad \text{or} \quad D_Z(S_1), D_Z(S_2) \geq \frac{1}{400}.$$

Proof. Let

$$\varepsilon_2 := \min\left\{\frac{(k-4)^2}{32n^2c_1}, \frac{1}{2c_1 n}, \varepsilon_1, \frac{c_2}{2c_1}, \frac{\delta_0}{2}\right\}.$$

Here ε_1 is from lemma 4.3.1.

Note that for $x, y \in G_Z^{\varepsilon_2}$, $x+y \in G_Z^{2\varepsilon_2}$. By the clustering property of Definition 3.5.5 we have

$d(x+y, \mathcal{C}_X^\perp) \leq 2c_1\varepsilon_2 n$ or $\geq c_2 n$. Define an equivalence relation

$$x, y \in G_Z^{\varepsilon_2}, \quad x \sim y \iff d(x+y, \mathcal{C}_X^\perp) \leq 2c_1\varepsilon_2 n.$$

This is indeed an equivalence relation, as if $x \sim y$ and $y \sim z$, then

$$d(x+z, \mathcal{C}_X^\perp) \leq d(x+y, \mathcal{C}_X^\perp) + d(y+z, \mathcal{C}_X^\perp) \leq 4c_1\varepsilon_2 n,$$

and by the dichotomy of the property, since we are assuming ε is small, it is not $\geq c_2 n$. Hence it is still $\leq 2c_1\varepsilon_2 n$.

So there is a partition of $G_Z^{\varepsilon_2}$ into the equivalence classes, call them $B_Z^1, B_Z^2, \ldots$. Due to the dichotomoy of the clustering property, each pair of equivalence classes has distance $\geq c_2 n$. Similarly, $G_X^{\varepsilon_1}$ splits up into $B_X^1, B_X^2, \ldots$.

We will bound the sizes of these clusters. Fix $z \in B_Z^i$. Then every other $z' \in B_Z^i$ has hamming distance atmost $2c_1\varepsilon_2 n$ from $z + w$ for some $w \in \mathcal{C}_X^\perp$. Note $|\mathcal{C}_X^\perp| = 2^{m_x}$. So

$$B_Z^i \subset \bigcup_{x \in \mathcal{C}_X^\perp} \overline{B_{2c_1\varepsilon_2 n}(x+z')}.$$

Now each closed ball of radius t with $t \leq n/2$ is the union of flipping i bits for $0 \leq i \leq t$. There are $\binom{n}{i} \leq \binom{n}{t}$ ways for each i, and so the ball has lenght atmost $(t+1)\binom{n}{t}$.

$$|B_Z^i| \leq 2^{k_x} \binom{n}{2c_1\varepsilon_2 n} (4c_1\varepsilon_2 n) \leq (2c_1\varepsilon_2 n + 1) 2^{m_x + \sqrt{2c_1\varepsilon_2 n}}.$$

We used stirling approximation on the binomial coefficient. We can similarly bound $|B_X^j|$.

Now we will show either $\forall i \, D_Z(B_Z^i) < 99/100$ or $\forall j \, D_X(B_X^j) < 99/100$. Indeed if for some i, $D_Z(B_Z^i) \geq 99/100$, then by the uncertainty principle lemma 4.3.1 we have

$$\begin{aligned} D_X(B_X^j) &\leq 2\sqrt{1 - D_Z(B_Z^i)} + \sqrt{\frac{|B_X^j| \cdot |B_Z^i|}{2^n}} \\ &\leq \frac{1}{5} + \sqrt{2^{m_x + m_y - n + 4\sqrt{2c_1\varepsilon_2 n}} (2c_1\varepsilon_2 n + 1)^2} \\ &= \frac{1}{5} + 4c_1\varepsilon_2 n 2^{-k/2 + 2\sqrt{2c_1\varepsilon_2 n}} \\ &< \frac{99}{100} \end{aligned}$$

Now suppose WLOG that $D_Z(B_Z^i) < 99/100$ for each i. Recall that $G_Z^{\varepsilon_2} = \cup_i B_Z^i$ and has

measure $\geq 199/200$ by lemma 4.3.2. we can find a q so that

$$D_Z\left(\bigcup_{i<q} B_Z^i\right) < \frac{1}{400} \quad \text{and } D_Z\left(\bigcup_{i<q+1} B_Z^i\right) \geq \frac{1}{400}.$$

Note that the second union cannot contain all the B_Zs, as the measure of B_Z^q is atmost 99/100, and so the union is

$$D_Z\left(\bigcup_{i<q+1} B_Z^i\right) = D_Z(B_Z^q) + D_Z\left(\bigcup_{i<q} B_Z^i\right) < \frac{99}{100} + \frac{1}{400} = \frac{397}{100}.$$

I.e, it is not the entire union which has bigger measure at 199/200. Hence

$$D_Z\left(\bigcup_{i\geq q+1} B_Z^i\right) = D_Z(G_Z^{\varepsilon_2}) - D_Z\left(\bigcup_{i<q+1} B_Z^i\right) > \frac{1}{400}.$$

Since each cluster has distance atleast $c_2 n$, so do the unions $\bigcup_{i\geq q+1} B_Z^i$ and $\bigcup_{i<q+1} B_Z^i$. This shows the result. □

We have proven the NLTS theorem, as combining this with the circuit lower bound of theorem 4.2.1, we get that either $|\psi\rangle$ or $H^{\otimes n}|\psi\rangle$ has a circuit of size atleast

$$\frac{1}{3}\log_2\left(\frac{(c_1 n)^2}{n\log(2/(1/400))}\right) = \Theta(\log(n)).$$

Note that we only looked at pure states. More or less the same proof would go through for mixed states pthat have low energy on $\mathbf{H}$, using the same circuit lower bound.

Bibliography

[Aar08] Scott Aaronson. *On Perfect Completeness for QMA*. arXiv:0806.0450 [quant-ph]. Aug. 2008. DOI: 10.48550/arXiv.0806.0450. URL: http://arxiv.org/abs/0806.0450 (visited on 08/14/2024).

[AB09] Sanjeev Arora and Boaz Barak. *Computational Complexity: A Modern Approach*. 1st. USA: Cambridge University Press, Mar. 2009. ISBN: 978-0-521-42426-4.

[ABN23] Anurag Anshu, Nikolas P. Breuckmann, and Chinmay Nirkhe. "NLTS Hamiltonians from good quantum codes". In: *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*. arXiv:2206.13228 [cond-mat, physics:quant-ph]. June 2023, pp. 1090–1096. DOI: 10.1145/3564246.3585114. URL: http://arxiv.org/abs/2206.13228 (visited on 07/07/2024).

[AE13] Dorit Aharonov and Lior Eldar. *Quantum Locally Testable Codes*. arXiv:1310.5664 [math-ph, physics:quant-ph]. Oct. 2013. DOI: 10.48550/arXiv.1310.5664. URL: http://arxiv.org/abs/1310.5664 (visited on 07/07/2024).

[AN22] Anurag Anshu and Chinmay Nirkhe. "Circuit lower bounds for low-energy states of quantum code Hamiltonians". In: *LIPIcs, Volume 215, ITCS 2022* 215 (2022). arXiv:2011.02044 [quant-ph], 6:1–6:22. ISSN: 1868-8969. DOI: 10.4230/LIPIcs.ITCS.2022.6. URL: http://arxiv.org/abs/2011.02044 (visited on 08/15/2024).

[Aro+98] Sanjeev Arora et al. "Proof verification and the hardness of approximation problems". In: *J. ACM* 45.3 (May 1998), pp. 501–555. ISSN: 0004-5411. DOI: 10.1145/278298.278306. URL: https://dl.acm.org/doi/10.1145/278298.278306 (visited on 08/14/2024).

[Bar+95] A. Barenco et al. "Elementary gates for quantum computation". In: *Physical Review A* 52.5 (Nov. 1995). arXiv:quant-ph/9503016, pp. 3457–3467. ISSN: 1050-2947, 1094-1622. DOI: 10.1103/PhysRevA.52.3457. URL: http://arxiv.org/abs/quant-ph/9503016 (visited on 08/13/2024).

[BLB04] Stéphane Boucheron, Gábor Lugosi, and Olivier Bousquet. "Concentration Inequalities". en. In: *Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, Australia, February 2 - 14, 2003, Tübingen, Germany, August 4 - 16, 2003, Revised Lectures*. Ed. by Olivier Bousquet, Ulrike von Luxburg, and Gunnar Rätsch. Berlin, Heidelberg: Springer, 2004, pp. 208–240. ISBN: 978-3-540-28650-9. DOI: 10.1007/978-3-540-28650-9_9. URL: https://doi.org/10.1007/978-3-540-28650-9_9 (visited on 08/13/2024).

[EH17] Lior Eldar and Aram W. Harrow. "Local Hamiltonians Whose Ground States are Hard to Approximate". In: *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*. arXiv:1510.02082 [quant-ph]. Oct. 2017, pp. 427–438. DOI: 10.1109/FOCS.2017.46. URL: http://arxiv.org/abs/1510.02082 (visited on 07/08/2024).

[FL16] Bill Fefferman and Cedric Lin. *Quantum Merlin Arthur with Exponentially Small Gap*. arXiv:1601.01975 [quant-ph]. Jan. 2016. DOI: 10.48550/arXiv.1601.01975. URL: http://arxiv.org/abs/1601.01975 (visited on 08/14/2024).

[KKR05] Julia Kempe, Alexei Kitaev, and Oded Regev. *The Complexity of the Local Hamiltonian Problem*. arXiv:quant-ph/0406180. Oct. 2005. DOI: 10.48550/arXiv.quant-ph/0406180. URL: http://arxiv.org/abs/quant-ph/0406180 (visited on 07/15/2024).

[KSV03] A. Yu. Kitaev, A. H. Shen, and M. N. Vyalyi. *Classical and Quantum Computation*. Vol. 110. ISSN: 00029890 Issue: 10 Journal Abbreviation: The American Mathematical Monthly. Dec. 2003. URL: https://www.jstor.org/stable/3647986?origin=crossref (visited on 08/14/2024).

[LZ22] Anthony Leverrier and Gilles Zémor. *Quantum Tanner codes*. arXiv:2202.13641 [quant-ph]. Sept. 2022. DOI: 10.48550/arXiv.2202.13641. URL: http://arxiv.org/abs/2202.13641 (visited on 07/07/2024).

[NN24] Anand Natarajan and Chinmay Nirkhe. "A distribution testing oracle separation between QMA and QCMA". In: *Quantum* 8 (June 2024). arXiv:2210.15380 [quant-ph], p. 1377. ISSN: 2521-327X. DOI: 10.22331/q-2024-06-17-1377. URL: http://arxiv.org/abs/2210.15380 (visited on 08/14/2024).

[PK22] Pavel Panteleev and Gleb Kalachev. *Asymptotically Good Quantum and Locally Testable Classical LDPC Codes*. arXiv:2111.03654 [quant-ph]. Jan. 2022. DOI: 10.48550/arXiv.2111.03654. URL: http://arxiv.org/abs/2111.03654 (visited on 07/07/2024).

[Sho95] Peter W. Shor. "Scheme for reducing decoherence in quantum computer memory". In: *Physical Review A* 52.4 (Oct. 1995). Publisher: American Physical Society, R2493–R2496. DOI: 10.1103/PhysRevA.52.R2493. URL: https://link.aps.org/doi/10.1103/PhysRevA.52.R2493 (visited on 08/14/2024).

[Wat08] John Watrous. *Quantum Computational Complexity*. arXiv:0804.3401 [quant-ph]. Apr. 2008. DOI: 10.48550/arXiv.0804.3401. URL: http://arxiv.org/abs/0804.3401 (visited on 08/13/2024).

[Wat18] John Watrous. *The Theory of Quantum Information*. Cambridge: Cambridge University Press, 2018. ISBN: 978-1-107-18056-7. DOI: 10.1017/9781316848142. URL: https://www.cambridge.org/core/books/theory-of-quantum-information/AE4AA5638F808D2CFEB070C55431D897 (visited on 08/13/2024).