Introduction to quantum computation: Quantum Complexity Classes

Lecture Notes for Quantum Computing

Lecturer: Dorit Aharonov   Scribe: Tomer Naveh

June 14, 2001

Abstract

This week's lecture discusses quantum complexity classes. It begins with a discussion of possible quantum analogs to the (classical) complexity class NP, then defines the complexity class QMA, and shows that the local hamiltonian problem (which can be seen as a quantum analog to 3SAT) is QMA-Complete.

1 Introduction

As we've seen before, $P \subseteq QP$ and $BPP \subseteq BQP$. It is also known that $BQP \subseteq P^{\sharp P}$. However, we do not know anything about the connection between these classes and the complexity class NP. Such a connection might be very interesting as NP is a very rich class with sound theory behind it.

2 Definition of QMA

We would like to define a complexity class which will be the quantum analog of NP. However, when trying to do so we encounter a major obstacle because errors and probabilistic computation are very common to quantum computation. Therefore we resort to defining QMA, the quantum analog of MA.

2.1 The complexity class MA

Informally, MA can be thought of as a probabilistic analog of NP, allowing for two-sided errors.

Definition 1 *$L \in MA$ if there exists a probabilistic polynomial time verifier V such that:*

- $\forall x \in L \;\; \exists y \;\; |y| = poly(|x|),\; Pr(V(x,y)=1) \geq \frac{2}{3}$
- $\forall x \notin L \;\; \forall y \;\; |y| = poly(|x|),\; Pr(V(x,y)=1) \leq \frac{1}{3}$

For $L \in MA$ the values $\frac{2}{3}$ and $\frac{1}{3}$ can be substituted by any $0 \leq p_0, p_1 \leq 1$ such that $|p_1 - p_0| = \frac{1}{poly(|x|)}$.

MA is naturally viewed as a game or interaction between 2 parties - Merlin, which has infinite computational power, and Arthur, which is limited to a polynomial time machine (the above V). Merlin should answer queries such as "is $x \in L$?", and accompany the answer with a polynomial witness y which Arthur can verify in polynomial time. Note that when showing that a problem is in MA, we should also show that Merlin cannot fool Arthur - i.e. that when $x \notin L$ there's no witness y that can persuade the verifier to believe that $x \in L$ with probability $> \frac{1}{3}$. We will define QMA analogously, where the verifier V is a quantum machine, and the witness y is a state of a polynomial number of qubits. Another possible definition (which we will not discuss) would be to take y as a classical witness. Clearly, $MA \subseteq QMA$.

3 The local hamiltonain problem

In this section we will define what can be thought of as the quantum analog of 3SAT, called "the local hamiltonian problem". We will also show that this problem is in QMA.

3.1 A different view of 3SAT

Let $\phi = C_1 \wedge C_2 \wedge \ldots \wedge C_k$ be a 3SAT formula on n variables. for every clause C_i we define a 8×8 matrix A_i, which is a projection on the unsatisfying assignment to this clause. For example, for the clause $X_1 \vee X_2 \vee X_3$ we get the matrix:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

If we look at vectors in ${Z_2}^8$, where each coordinate represents an assignment to the variables of C_i (for example, 10000000 represents the assignment (000)), then the standard basis of ${Z_2}^8$ is a basis of eigenvectors, where the vectors of the satisfying assignments correspond to the eigenvalue 0, and the vector of the unsatisfying assignment corresponds to the eigenvalue 1. Let $H_i = A_i \otimes I_{2^{n-3}}$ where A_i operates on the variables of C_i and $I_{2^{n-3}}$ operates on all the rest. Now we look at vectors in ${Z_2}^{2^n}$ which represent assignments to all the variables of ϕ. Let $v \in {Z_2}^{2^n}$ be a vector that represents an assignment A. v is a vector in the standard basis of ${Z_2}^{2^n}$. If A satisfies a clause C_i, then $H_i v = 0$. Otherwise, $H_i v = v$. We can view this as if the matrix H_i "penalizes" assignments that do not satisfy C_i. We denote $H = \sum_{i=1}^{k} H_i$. We observe that if v is a vector representing an assignment A, then $Hv = av$ where a is the number of unsatisfied clauses. Thus, 3SAT is equivalent to the following problem: "Is 0 an eigenvalue of H?"

3.2 The quantum analog

In this section we define the local hamiltonian problem, and show that it is in QMA. Observe that every A_i can also be written as $|i\rangle \langle i|, 0 \leq i \leq 7$. The quantum analog we discuss is a generalization of the above representation of 3SAT, where we allow A_i to be $|\alpha\rangle \langle\alpha|$ for every state $|\alpha\rangle$ of 3 qubits. H_i and H are defined accordingly. The problem is to decide whether H has an eigenvalue $< p_0$ or that all eigenvalues are $> p_1$, where $p_1 - p_0 > \frac{1}{poly(n)}$. We do not care about instances which do not fall into one of these two categories.

We would have liked the witness $|\xi\rangle$ to be an eigenvector of the minimal eigenvalue. A possible verification process would then be to compute the projections of $|\xi\rangle$ on each of the subspaces of the unsatisfying assignments to the clauses (with the H_i's). The problem with this suggested verification process is that we cannot compute the projections one after the other, because each projection changes the witness. Therefore the witness will be a tensor product of a polynomial number of copies of $|\xi\rangle$, i.e. $|\xi\rangle^{\otimes poly(n)}$. Now we will show how the verification process can be done with this definition of the witness.

If $|\xi\rangle$ is an eigenvector corresponding to an eigenvalue λ, then $\lambda = \langle\xi| H |\xi\rangle = \sum_i \langle\xi| H_i |\xi\rangle$. The verification process will estimate λ in the following way: for each copy of $|\xi\rangle$ the verifier randomly chooses $1 \leq i \leq k$ and projects $|\xi\rangle$ (the current copy) on $|\alpha_i\rangle$ (i.e. measures $|\xi\rangle$ with a basis that contains $|\alpha_i\rangle$) where $|\alpha_i\rangle$ is a vector such that $H_i = |\alpha_i\rangle\langle\alpha_i|$. By repeating it many times for the same i we get an estimation for $Pr(|\alpha_i\rangle\ is\ measured) = \langle\xi| |\alpha_i\rangle\langle\alpha_i| |\xi\rangle = \langle\xi| H_i |\xi\rangle$. After a polynomial number of iterations, each i will be chosen a polynomial number of times (w.h.p.) and so we get an estimation of λ which is better then the gap between p_0 and p_1 (again, w.h.p.).

Now we have to show soundness, i.e. that Merlin cannot fool Arthur. To show this, we need to show that if there's no valid witness, our verification process will return a false answer w.h.p. To see that, take the witness we get in such situation, denoted by $|\xi\rangle$. If we take a basis for each register of the witness, denoted $|\xi_{i_1}\rangle \ldots |\xi_{i_k}\rangle$ for some k, we can construct a basis to the witnesses space from all tensor products of choices of one vector from each basis. Then we can write the witness as a linear combination of this basis. The expected value of applying the verification operator on each register is then $\geq p_1$, and from linearity and because we apply it a polynomial number of times for each H_i, we get that w.h.p. our estimation will be $> p_0$, therefore we detect that it is a false witness.

4 QMA Completeness

In this section we will show that the local hamiltonian problem is QMA-Complete.

4.1 The Reduction

Let L be a problem in QMA. Then there exist a quantum circuit with gates $U_1, \ldots, U_l$ such that for an input $|x\rangle$ and a witness $|\xi\rangle$ the output qubit is $|1\rangle$ if $|x\rangle \in L$ and $|0\rangle$ otherwise. The reduction bears some resemblance to the one in the proof that 3SAT is NP-Complete. We look at the different stages of the computation of the machine of the problem L, operating on input x. We will define the witness $|\eta\rangle$ to the local hamiltonian problem to be a superposition of the stages of the computation i.e. $|\eta\rangle = \sum_{t=0}^{l} U_t U_{t-1} \ldots U_1 |\gamma_0\rangle \otimes |t\rangle$ where $|\gamma_0\rangle$ is the state at the beginning of the computation and $|t\rangle$ functions as a "clock" (we will see later that we need this clock). We will construct an input to the local hamiltonian problem, starting with a simple construction which illustrates the ideas of the reduction, and improve it each time we encounter a problem until we get a valid instance.

We start with a matrix $H = H_{in} + H_{out} + H_{prop}$ which is built as follows:

- H_{in} is a matrix that checks the input (verifies that it is indeed x).
- H_{out} is a matrix that checks that the output is 1.
- H_{prop} checks that the propagation of the computational process is done according to the given circuit.

Construction of H_{in}: For each bit of the input, if it is 0, we define a matrix $|1\rangle\langle 1| \otimes I$, and if it is 1, the matrix is $|0\rangle\langle 0| \otimes I$, where $|0\rangle\langle 0|$ or $|1\rangle\langle 1|$ operates of the relevant bit and I on the rest. H_{in} will be the sum of these matrices, thus it "penalizes" inputs that are different than x, and the penalty depends on the number of different bits.

Construction of H_{out}: $H_{out} = |0\rangle\langle 0| \otimes I$ where $|0\rangle\langle 0|$ operates on the first output qubit. This penalizes outputs that are not $\langle 1|$.

More precisely, H_{in} will be a tensor product of the previously defined matrix with $|0\rangle\langle 0|$ that operates on the clock qubits, and H_{out} will be tensored with $|l\rangle\langle l|$, where $|l\rangle\langle l|$ also operates on the clock. With this refinement of H_{in} and H_{out}, if we look at a superposition of all the states of a computation $\sum_{t=0}^{l} |\gamma_t\rangle \otimes |t\rangle$ (where $|\gamma_t\rangle = U_t U_{t-1} \dots U_1 |\gamma_0\rangle$) we get that:

$$H_{in}(\sum_{t=0}^{l} |\gamma_t\rangle \otimes |t\rangle) = H_{in}(|\gamma_0\rangle \otimes |0\rangle)$$

$$H_{out}(\sum_{t=0}^{l} |\gamma_t\rangle \otimes |t\rangle) = H_{out}(|\gamma_0\rangle \otimes |l\rangle)$$

Therefore H_{in} and H_{out} "donate" their part of the total penalty according to the first and last states, as we would expect.

Construction of H_{prop}: Now let us define H_{prop} and verify that it indeed checks that the propagation of the computation is correct. We start with a simplified definition of H_{prop}, which will help us understand the motivation but will be a bit incorrect.

$$H_{prop} = 2I - \sum_{k=1}^{l} (U_k \otimes |k\rangle\langle k-1| + U_k^{\dagger} \otimes |k-1\rangle\langle k|)$$

The first component in the summation advances a state in the computation one step forward, i.e. $|\gamma_t\rangle \otimes |t\rangle$ becomes $|\gamma_{t+1}\rangle \otimes |t+1\rangle$. The second component takes a state one step backwards, i.e. $|\gamma_t\rangle \otimes |t\rangle$ becomes $|\gamma_{t-1}\rangle \otimes |t-1\rangle$. Therefore, for a witness which represents a valid computation, the 2I and the sum are canceling each other and we get 0. This is true except for the first and the last steps of the computation, as we can see:

$$\begin{aligned}
H_{prop} |\eta\rangle &= (2I - \sum_{k=1}^{l} (U_k \otimes |k\rangle\langle k-1| + U_k^{\dagger} \otimes |k-1\rangle\langle k|))(\sum_{k=0}^{l} |\gamma_k\rangle |k\rangle) \\
&= 2\sum_{k=0}^{l} |\gamma_k\rangle |k\rangle - \sum_{k=1}^{l} (U_k |\gamma_{k-1}\rangle \otimes |k\rangle + U_k^{\dagger} |\gamma_k\rangle \otimes |k-1\rangle) \\
&= 2\sum_{k=0}^{l} |\gamma_k\rangle |k\rangle - \sum_{k=1}^{l} |\gamma_k\rangle |k\rangle - \sum_{k=1}^{l} |\gamma_{k-1}\rangle |k-1\rangle \\
&= 2\sum_{k=0}^{l} |\gamma_k\rangle |k\rangle - \sum_{k=1}^{l} |\gamma_k\rangle |k\rangle - \sum_{k=0}^{l-1} |\gamma_k\rangle |k\rangle \\
&= |\gamma_0\rangle |0\rangle + |\gamma_l\rangle |l\rangle
\end{aligned}$$

To fix this, we will change the previous definition of H_{prop} to be:

$$H_{prop} = 2I - \sum_{k=1}^{l} (U_k \otimes |k\rangle\langle k-1| + U_k^{\dagger} \otimes |k-1\rangle\langle k|) - I \otimes |0\rangle\langle 0| - I \otimes \langle l| |l\rangle .$$

One can easily check that for a witness $|\eta\rangle$ which represents a valid computation, we get $H_{prop}|\eta\rangle = 0$, so η is an eigenvector of eigenvalue 0.

We would like to show that we indeed get an instance of the local hamiltonian problem. However, when trying to do so we notice that all the parts of H dealing with the clock are not local - they are operations on $\log(l)$ bits. Therefore we change the clock to be unary (i.e. time $t = 4$ is represented as $111100000\ldots$). Now, as long as we can verify locally that the time is always a valid unary time, the operation $|k\rangle\langle k-1|$ becomes $|110\rangle\langle 100| \otimes I$ where $|110\rangle\langle 100|$ operates on bits $k-1, k, k+1$ of the clock and I on the rest. This is clearly a local operation. Similarly, $|k-1\rangle\langle k|$ becomes $|100\rangle\langle 110| \otimes I$. There are two exceptions to the above - when $k = 1$ and $k = l$, in order not to refer to bits 0 and $l+1$ of the clock which do not exist. For $k = 1$ we will drop the first bit of the 3-bit operator, so the operator $|1\rangle\langle 0|$ on the binary clock becomes $|10\rangle\langle 00|$, and similarly $|0\rangle\langle 1|$ becomes $|00\rangle\langle 10|$. For the bit $k = l$ we drop the 3^{rd} bit of the operators in the same manner. The only thing left to handle in order to have all clocks in unary form are the "correction terms" introduced, i.e. $I \otimes |0\rangle\langle 0|$ and $I \otimes |l\rangle\langle l|$. The $|0\rangle\langle 0|$ term will now become $|00\rangle\langle 00|$ on bits $1, 2$ of the clock and the $|1\rangle\langle 1|$ term becomes $|11\rangle\langle 11|$ on bits $l-1, l$. So we get

$$H_{prop} = 2I - \sum_{k=2}^{l-1} (U_k \otimes |110\rangle\langle 100| + U_k^\dagger \otimes |100\rangle\langle 110|)$$

$$+U_1 \otimes |10\rangle\langle 00| + U_1^\dagger \otimes |00\rangle\langle 10|$$

$$+U_l \otimes |11\rangle\langle 10| + U_l^\dagger \otimes |10\rangle\langle 11|$$

$$-I \otimes |00\rangle\langle 00| - I \otimes |11\rangle\langle 11| .$$

On the relevant bits as described above. Note that for a witness $|\eta\rangle$ which represents a valid computation, we still get $H_{prop}|\eta\rangle = 0$.

Now we define a matrix H_{clock} that locally checks that the clock bits are a valid unary representation of a clock. All we need to check is that two consequent bits cannot be "01″. This can be done by a sum of local projections, as follows:

$$H_{clock} = \sum_{k=1}^{l} |01\rangle\langle 01|_{k-1,k} \otimes I$$

Clearly, $|\eta\rangle$ (with a unary clock) is an eigenvector of eigenvalue 0 of both H_{prop} and H_{clock}.

Now we show that $H = H_{in} + H_{out} + H_{prop} + H_{clock}$ is a sum of local projections. It is obvious that all matrices except H_{prop} are such. H_{prop} is a hermitian matrix, therefore it can be diagonalized, and thus represented as $\sum \lambda_i |\gamma_i\rangle\langle\gamma_i|$ which is a sum of projections.

To complete the reduction, we need to show that if $x \notin L$, the minimal eigenvalue is large enough, i.e. the gap between this eigenvalue and 0 is at least $\frac{1}{poly(|x|)}$

Proposition 1 *If $x \notin L$ then the minimal eigenvalue is greater than $\frac{1}{l^5}$.*

Proof: We denote $H_1 = H_{in} + H_{out}$, $H_2 = H_{prop} + H_{clock}$.

Lemma 1 *Let H_1 and H_2 be two matrices with non-negative eigenvalues, and let N_1 and N_2 be the eigenspaces of the eigenvalue 0, respectively. If the angle between N_1 and N_2 is some $\alpha > 0$, and the second eigenvalue of both H_1 and H_2 is $\geq \lambda$ then the minimal eigenvalue of $H_1 + H_2 \geq \lambda \sin\alpha$.*

Proof: We denote by $|\delta\rangle$ an eigenvector of the minimal eigenvalue of $H_1 + H_2$, such that $|||\delta\rangle|| = 1$. The minimal value of $||(H_1 + H_2)|\delta\rangle||$ is achieved when the angle between $|\delta\rangle$ and N_1 and the angle between $|\delta\rangle$ and N_2 are both $\frac{\alpha}{2}$. In this case we get $||H_1|\delta\rangle|| \geq \lambda \sin\frac{\alpha}{2}$ and $||H_2|\delta\rangle|| \geq \lambda \sin\frac{\alpha}{2}$ because $|\delta\rangle$ can be decomposed to $|\delta_1\rangle \in N_1$ and $|\delta_2\rangle \perp N_1$ such that $|\delta\rangle = |\delta_1\rangle + |\delta_2\rangle$. Then, $H_1|\delta_1\rangle = 0$ and $|||\delta_2\rangle|| = \sin\frac{\alpha}{2}$ and since the second eigenvalue of H_1 is $\geq \lambda$ we get $||H_1|\delta\rangle|| \geq \lambda \sin\frac{\alpha}{2}$ and similarly for H_2. Now,

$$||(H_1 + H_2)|\delta\rangle||^2 = ||H_1|\delta\rangle||^2 + ||H_2|\delta\rangle||^2 + 2 < H_1|\delta\rangle, H_2|\delta\rangle >$$

$$\geq 2[\lambda^2 \sin^2\frac{\alpha}{2} + ||H_2|\delta\rangle|| \cdot ||H_2|\delta\rangle|| \cdot \cos\alpha] = 2\lambda^2 \sin^2\frac{\alpha}{2}(1 + \cos\alpha)$$

$$= 4\lambda^2 \sin^2\frac{\alpha}{2}\cos^2\frac{\alpha}{2} = (\lambda \sin\alpha)^2$$

■

Now we get back to the proof of the proposition. We will bound α, the angle between N_1 and N_2 (the null spaces of H_1 and H_2). Let us assume that $x \notin L$. N_1 is a superposition of the following:

$$|x\rangle\langle x| \otimes |0\rangle\langle 0|, |1\rangle\langle 1| \otimes |l\rangle\langle l|, \quad \forall A, \quad A \otimes |t\rangle\langle t| \quad 1 \leq t \leq l-1$$

N_2 is all valid computations, i.e. $|\eta\rangle$ such that

$$|\eta\rangle = \sum_{k=0}^{l} U_k \ldots U_1 |\alpha_0\rangle |k\rangle\langle k| \quad \forall |\alpha_0\rangle$$

We will bound $\cos^2\alpha = \langle\eta| N_1 |\eta\rangle$ as follows: Times $t = 1, ..., l-1$ in the uniform superposition $|\eta\rangle$ give us at most $\frac{1}{l+1}$ each. Since $x \notin L$, either the projection of η on $|x\rangle\langle x| |0\rangle\langle 0|$ is 0, or the projection on $|1\rangle\langle 1| |l\rangle\langle l|$ is $\leq \frac{1}{3}$ (it is not 0 since in QMA we allow for an error of $\frac{1}{3}$). Therefore times $t = 0$ and $t = l$ together give at most $\frac{4}{3(l+1)}$. We get that $cos^2\alpha \leq 1 - \frac{2}{3(l+1)}$.

Now we show that the second eigenvalues of H_1 and H_2 are large enough. The second eigenvalue of H_1 is ≥ 1 since H_{in} and H_{out} are projections and hence their eigenvalues are 0 and 1. Since the eigenspaces of the eigenvalue 1 of H_{in} and H_{out} are orthogonal (because they operate on different times), the claim holds.

The second eigenvalue of H_2 will be bounded as follows:

TO BE CONTINUED ...