CMSC 858L: Quantum Complexity

Instructor: Daniel Gottesman

Spring 2023

18 Witness Amplification and Quantum PCP

18.1 References

You can read more about the Quantum PCP conjecture in Aharonov, Arad, and Vidick, "The Quantum PCP Conjecture," arXiv:1309.7495, although there has been some progress since then, most notably the recent resolution of the NLTS conjecture. Witness amplification was introduced in Marriott and Watrous, "Quantum Arthur-Merlin Games," cs/0506068. Vidick and Watrous, "Quantum Proofs," arXiv:1610.01664 is a good reference for QMA and quantum interactive proofs (which we will discuss later).

18.2 Witness Amplification

We have been using the fact that we can amplify the acceptance probability of a QMA witness. The analogous property for classical MA is quite straightfoward: If we have a randomized algorithm to test the witness w with probability p, we can simply run the algorithm multiple times and take the majority result. The probability of getting the right answer for a single test is independent over multiple runs of the algorithm, so the chance of a majority of the trials giving the wrong answer is small.

For QMA, the roadblock to this strategy is that the witness is used up the first time we run a checking circuit. If we want to amplify the success probability by repeating, the most obvious way to do so is to require Merlin to give us multiple copies of the witness, and then we can run the checking circuit multiple times. This works, but does require some additional proof to verify that Merlin cannot cheat by giving us a large entangled state instead of the tensor product of witnesses.

However, surprisingly, there is a way to magnify the acceptance probability of a witness *without* having to increase the size of the witness. The main idea is to run the checking circuit, measure or otherwise record the output, and then reverse the checking circuit and see if the ancilla qubits return to their initial $|00\cdots 0\rangle$ state.

In particular, imagine that we start with the state $|\psi\rangle \otimes |00\cdots 0\rangle$ (the second register being ancilla qubits needed for the checking circuit), perform a unitary checking circuit C, concluding with a measurement, which we can assume is a projective measurement of a single qubit. Let $M_i = |i\rangle\langle i| \otimes I$ be a projection onto the measured final qubit having value i and let $P_0 = I \otimes |00\cdots 0\rangle\langle 00\cdots 0|$, $P_1 = I - P_0$ be projections onto the subspace with ancillas all 0 and the orthogonal subspace.

Consider the operators

$$A_{ij} = P_j C^\dagger M_i C. \tag{1}$$

(Note that these are generally not projectors.) If the checking circuit is accepted with 100% probability, then

$$A_{10}|\psi\rangle \otimes |00\cdots 0\rangle = |\psi\rangle \otimes |00\cdots 0\rangle \tag{2}$$

$$A_{00}|\psi\rangle \otimes |00\cdots 0\rangle = 0 \tag{3}$$

since the projection M_1 is trivial and M_0 annihilates the state. Conversely, if the checking is always rejected,

$$A_{00}|\psi\rangle \otimes |00\cdots 0\rangle = |\psi\rangle \otimes |00\cdots 0\rangle \tag{4}$$
$$A_{10}|\psi\rangle \otimes |00\cdots 0\rangle = 0. \tag{5}$$

Note that in both cases, for the outcome that can actually occur, we get the original witness back by reversing C.

In the case where the circuit is accepted with high probability, but not quite 1, we don't always succeed in getting back to $|\psi\rangle \otimes |00\cdots 0\rangle$, but we can get close. Instead, however, let us first focus on eigenstates of the operators A_{i0}, albeit ones with eigenvalues less than 1. Thus, for instance, we have

$$A_{i0}|\psi\rangle \otimes |00\cdots 0\rangle = \lambda_i|\psi\rangle \otimes |00\cdots 0\rangle. \tag{6}$$

Note that any eigenstate with non-zero eigenvalue of A_{i0} must be of the form $|\psi\rangle \otimes |00\cdots 0\rangle$ since A_{i0} ends with a projector onto such states and that since $A_{00} + A_{10} = P_0$, any eigenvector of A_{10} will also be an eigenvector of A_{00}. Furthermore,

$$(\langle\psi| \otimes \langle 00\cdots 0|)A_{i0}(|\psi\rangle \otimes |00\cdots 0\rangle) = (\langle\psi| \otimes \langle 00\cdots 0|)C^\dagger M_i C(|\psi\rangle \otimes |00\cdots 0\rangle), \tag{7}$$

which is p_i, the probability that the circuit C gives outcome i on input $|\psi\rangle$. Thus, $\lambda_i = p_i$.

Suppose we actually run the procedure implied by A_{i0}. That is, run C, measure the output qubit to have value i (which occurs with probability p_i) and then run $C^\dagger$ and measure to see if the ancilla qubits have been reset to all 0. What is the overall probability that this happens, including the output qubit having value i? It is

$$\begin{aligned}(\langle\psi| \otimes \langle 00\cdots 0|)A_{i0}^\dagger A_{i0}(|\psi\rangle \otimes |00\cdots 0\rangle) &= (\langle\psi| \otimes \langle 00\cdots 0|)C^\dagger M_i C P_0 C^\dagger M_i C(|\psi\rangle \otimes |00\cdots 0\rangle) && (8)\\ &= \sum_\ell |(\langle\psi_\ell| \otimes \langle 00\cdots 0|)C^\dagger M_i C(|\psi\rangle \otimes |00\cdots 0\rangle)|^2 && (9)\\ &= |(\langle\psi| \otimes \langle 00\cdots 0|)C^\dagger M_i C(|\psi\rangle \otimes |00\cdots 0\rangle)|^2 && (10)\\ &= p_i^2. && (11)\end{aligned}$$

Here, $|\psi_\ell\rangle$ runs over a basis of possible witness states including $|\psi\rangle$ and the third line follows because $|\psi\rangle|00\cdots 0\rangle$ is an eigenstate of A_{i0}.

Thus, the conditional probability of the ancillas being 0 conditioned on having outcome bit i is p_i. Also, note that for eigenstates $|\psi\rangle$, if we do happen to get the ancillas being in the state all 0, then we have restored the ancilla and can just repeat the process on the same state to get more data as to the values of p_0 and p_1.

But what about if we find the ancillas are *not* all 0? In that case, we are instead performing A_{i1}. We know that the conditional probability of ancillas not being 0 conditioned on output qubit being i is $p_{1\oplus i}$, but what is the final state? It is not generally going to be $|\psi\rangle$, which might pose a problem for repeating the measurement.

Let $|\phi_i\rangle = A_{i1}|\psi\rangle \otimes |00\cdots 0\rangle$. Note that $\langle\phi_i|\phi_i\rangle = p_0 p_1$. Suppose we go ahead and perform C on $|\phi_i\rangle$ and measure the output qubit. What state do we get if we get outcome k?

$$\begin{aligned}M_k C A_{i1}|\psi\rangle \otimes |00\cdots 0\rangle &= M_k C(I - P_0)C^\dagger M_i C|\psi\rangle \otimes |00\cdots 0\rangle && (12)\\ &= M_k M_i C|\psi\rangle \otimes |00\cdots 0\rangle - M_k C A_{i0}|\psi\rangle \otimes |00\cdots 0\rangle && (13)\\ &= \delta_{ik} M_k C|\psi\rangle \otimes |00\cdots 0\rangle - p_i M_k C|\psi\rangle \otimes |00\cdots 0\rangle. && (14)\end{aligned}$$

That is, whether $i = k$ or not, we get $M_k C|\psi\rangle \otimes |00\cdots 0\rangle$, the same state we get from the correct initial state on outcome probability k. If $i \neq k$, the normalization is p_i and the overall probability of this sequence of results is $p_i^2 p_k$. If $i = k$, the normalization is $1 - p_i = p_{i\oplus 1}$ and the overall probability of this sequence of results is $p_i p_{i\oplus 1}^2$. Since the probability of getting i and then P_1 is $p_0 p_1$, in either case, the conditional probability of getting outcome k is $p_{k\oplus i}$. That is, the probabilities are reversed from the original circuit: outcome 1 happens with probability p_0 and outcome 0 happens with probability p_1! And this is true regardless of what the measurement outcome was the first time we ran C.

One consequence is that after running C and measuring the output qubit in value k, we always get the same state, which depends on k but not on whether we started with $|\psi\rangle \otimes |00\cdots 0\rangle$ or $|\phi_i\rangle$ (for either i). And the probability of getting k didn't depend on whether we started with $|\phi_0\rangle$ or $|\phi_1\rangle$, only on whether the previous measurement result was P_0 or P_1. If we continue alternating C followed by a measurement of M_i and then $C^\dagger$ followed by measurement of P_j, the probability of getting outcome i or j only depends on the previous measurement outcome and not on any earlier ones.

Moreover, the pattern is very simple: After measuring P_j, we get the following conditional probabilities:

$$\begin{aligned}
\text{Prob}(i = 0|j = 0) &= p_0\\
\text{Prob}(i = 1|j = 0) &= p_1\\
\text{Prob}(i = 0|j = 1) &= p_1\\
\text{Prob}(i = 0|j = 1) &= p_0.
\end{aligned}$$

That is, $\text{Prob}(i = j) = p_0$ and $\text{Prob}(i \neq j) = p_1$. Similarly, after measuring M_i, we get conditional probabilities:

$$\begin{aligned}
\text{Prob}(j = 0|i = 0) &= p_0\\
\text{Prob}(j = 0|i = 1) &= p_1\\
\text{Prob}(j = 1|i = 0) &= p_1\\
\text{Prob}(j = 1|i = 1) &= p_0.
\end{aligned}$$

Once again, $\text{Prob}(j = i) = p_0$ and $\text{Prob}(j \neq i) = p_1$. Each outcome ($i$ or j) has the same value as the previous one with probability p_0 and has the opposite value with probability p_1. When we repeat this a number of times, we get a good estimate of p_0 and p_1 by counting the number of changes of outcome. This will enable us to determine if $p_0 \geq 2/3$ or $\leq 1/3$ with high probability.

This analysis was done under the assumption that $|\psi\rangle \otimes |00\cdots 0\rangle$ was an eigenvector of A_{i0}. Since $A_{00} + A_{10} = P_0$, the projector on the space $|\psi\rangle \otimes |00\cdots 0\rangle$, we can choose a basis of eigenvectors $|\psi_\alpha\rangle$ of A_{i0} in the subspace $|\psi\rangle \otimes |00\cdots 0\rangle$ and write any witness state as a superposition of eigenvectors. Since $A_{i0}|\psi_\alpha\rangle = p_i|\psi_\alpha\rangle$ and $A_{k0}A_{i1}|\psi_\alpha\rangle \propto |\psi_\alpha\rangle$ as well, the different eigenvectors of A_{i0} don't get mixed up under repeatedly applying A_{ij}. Thus, if the initial state is a superposition of different eigenvectors, each of them acts independently, meaning we see a sequence of measurement outcomes as if we had just a single eigenvector chosen randomly from the superposition.

18.3 Quantum PCP

A major result of classical complexity theory is the PCP theorem (for "probabilistically checkable proofs"), which states that every language in NP has a verification procedure (with appropriate witness) that only checks the witness in a constant number of locations. One application of the PCP theorem is to show that many NP-complete problems are hard to even approximate: For instance, suppose you have a k-SAT instance. Recognizing that finding a satisfying instance is going to be hard, you might instead wish to find a solution that satisfies a fraction f of all the clauses in the problem. However, PCP shows that this is not possible if $f > 7/8$, or at least not unless P = NP.

The analogous statement for QMA is open, and is one of the biggest open problems in the field. The Quantum PCP conjecture is most often phrased in terms of a local-Hamiltonian problem:

Conjecture 1. *The following language is QMA-complete: Instances of the form (H, E, Δ), for an n-qubit system, with quantities specified to polynomial accuracy, H a k-local Hamiltonian (for constant k), $\Delta = O(1)$ with the promise that either:*

1. *The ground state energy of H is at most En ("yes" instance)*

2. The ground state energy of H is at least $(E + \Delta)n$ ("no" instance).

Note that this is just a standard k-LOCAL HAMILTONIAN problem, but we are requiring that the promise gap is *linear in n.* (As opposed to $O(1/\text{poly}(n))$ as before.)

Again, this is an open problem; the above is a conjecture, and it is really not clear if it is true or not. There's no real evidence one way or the other. Note, though, that the hard-to-approximate instances of SAT are special cases of the language above with a Hamiltonian diagonal in the standard basis, so the language is at least NP-hard.

If the quantum PCP conjecture is true, that means that the ground state energy *per particle* is hard to approximate for some systems to even constant accuracy, let alone accuracy increasing with any parameter. It also implies that it is hard to find the thermal state for some systems at temperature greater than 0. To understand this statement, note that thermodynamics tells us that at thermal equilibrium, we have a mixture of states with different energies. The relative population of a particular eigenstate with energy E is

$$\text{Prob}(E) \propto e^{-\beta E}, \tag{15}$$

where $\beta = 1/kT$. T is the temperature in Kelvin and k is known as Boltzmann's constant. To get the actual population for this particular eigenstate, divide the above by the *partition function* $Z = \sum e^{-\beta E}$.

In particular, as $T \to 0$, the population is dominated by the ground state, whereas as $T \to \infty$, all states are equally populated. A thermal state is one that satisfies the correct distribution (**??**). Therefore at infinite temperature, finding the thermal state is easy: It is simply the maximally mixed state. However, at 0 temperature, for at least some systems, it is hard since the thermal state is just the ground state.

What about non-zero but low temperature? One might think that the thermal state will have a large admixture of the ground state, and that therefore this case will be hard as well: If you could create the thermal state, then with some high probability you could also create the ground state by simply measuring the energy of a thermal state, repeating polynomially many times until you find a solution.

However, there is one additional factor to take into account, which is the degneneracy of low-energy states. (**??**) only gives the weighting for a *single* state. But if we have n_E states of energy E, *each* of them has relative population $e^{-\beta E}$, meaning the total population of energy E states is

$$\sum_E n_E e^{-\beta E}/Z. \tag{16}$$

While the $e^{-\beta E}$ term decreases exponentially with E, it is possible (and often will be true) that the n_E term increases exponentially with E. If this is the case, then at any non-zero temperature T, the ground state is overwhelmed by the much larger numbers of higher-energy states. This means that measuring the energy in a thermal state of this sort is unlikely to produce ground states. That route to trying to show hardness of the thermal state hits a dead end.

However, if the Hamiltonian satisfies the conditions and results of the quantum PCP conjecture, then finding the thermal state is QMA-hard. We will show that for a Hamiltonian satisfying the promise of the quantum PCP conjecture, the LOCAL HAMILTONIAN problem reduces to finding the energy of a thermal state for that Hamiltonian. Under the quantum PCP conjecture, the k-LOCAL HAMILTONIAN with this promise would be QMA-complete, which means that finding the energy of a thermal state would also be QMA-complete.

Suppose you have an algorithm to measure the energy of the thermal state. Under the quantum PCP promise, this tells you about the ground state energy: If it is a "no" instance of the LOCAL HAMILTONIAN problem, then all eigenstates have energy at least $(E+\Delta)n$, and certainly we will get at least this value when we measure the energy of the thermal state. On the other hand, if we have a "yes" instance of the LOCAL HAMILTONIAN problem, then the ground state has energy En, which we can shift to 0, so its relative probability is 1. We don't know the distribution of higher-energy states, and there may be some or many with energies below $(E + \Delta)n$. There will also be some or many with energies above $(E + \Delta)n$. However, we know that the probability of such states in the thermal state is at most $n_E e^{-\beta(E+\Delta)n}$. Even if $n_E \propto 2^n$ (the total number of states), the overall population with energy at least $(E + \Delta)n$ is at most $e^{(2-\beta(E-\Delta))n}$, which for low enough temperature (large enough β), will always be negligible. Thus, the thermal state will be dominated by states with energy below $(E + \Delta)n$. This means that the energy of the thermal state will certainly be lower than $(E + \Delta)n$. Consequently, if we accept whenever the measured energy for the witness is less than $(E + \Delta)n$ and reject otherwise, we will have answered the instance the local Hamiltonian problem.