Quantum Multi Prover Interactive Proofs with Communicating Provers *

Michael Ben Or † Avinatan Hassidim ‡ Haran Pilpel §

Abstract

We introduce another variant of Quantum MIP, where the provers do not share entanglement, the communication between the verifier and the provers is quantum, but the provers are unlimited in the *classical* communication between them. At first, this model may seem very weak, as provers who exchange information seem to be equivalent in power to a simple prover. This in fact is not the case—we show that any language in NEXP can be recognized in this model efficiently, with just two provers and two rounds of communication, with a constant completeness-soundness gap. Similar ideas and techniques may help help with other models of Quantum MIP, including the question of noncommunicating provers with unlimited entanglement.

1 Introduction

Multi Prover Interactive Proofs (MIPs) have been studied extensively in the classical setting, and provide an exact characterization of NEXP [2]. Extending MIPs to the quantum setting poses many important open problems, and may give us more intuition regarding the power of entanglement. There are several possible generalizations for quantum multi-prover schemes, which differ in the power of the verifier (which can be quantum or classical), and in the relation between the provers—for example, how much entanglement they have. For a classical verifier, limited entanglement between the provers can only weaken the expressive power of the model (or not change it). When the verifier is quantum the situation can be more complicated, and entanglement can, in theory, increase or decrease the expressive power.

There are many interesting results in the model where the verifier is classical and the provers share (limited or unlimited) entanglement. The first results were obtained by Kobayashi and Matsumoto [17]. They proved that as long as the provers share a bounded (polynomial) amount of entanglement, the set of languages which can be

*Research supported by the Israel Science Foundation.

†Incumbent of the Jean and Helena Alfassa Chair in Computer Science. benor@cs.huji.ac.il, The Hebrew University, Jerusalem, Israel

‡avinatan@mit.edu MIT, Cambridge USA 02139

§haranp@math.huji.ac.il, The Hebrew University, Jerusalem, Israel. Research supported by the Giora Yoel Yashinsky memorial prize.

recognized is contained in NEXP, even if the verifier is quantum. Cleve et al. [8] provide examples where the proof is valid if the provers share no entanglement, but is no longer sound when they do. Preda [22] showed that if the provers are not limited to quantum entanglement, but instead have an unlimited amount of nonlocal boxes [21], then the set of recognizable languages is contained in EXP.

A limited family of quantum games is XOR-Games. In this type of games, the verifier is only allowed to look at the XOR of the answers sent by the provers. Cleve et al. [10] showed a parallel repetition lemma for these type of games, even if the provers share entanglement. However, Wehner [25] showed that if the provers share entanglement then these games are in EXP.

There are also some positive results when the provers are quantum and share entanglement. Cleve et al. [9] provide a proof system for NP when the verifier is classical and the provers share an unlimited amount of entanglement. The proof scheme provides a constant gap, but the communication is linear. Kempe et al. [13] give a quantum protocol for recognizing languages in NP by a quantum verifier with logarithmic communication, when the provers share unlimited entanglement. However, when $x \notin L$ the probability that the verifier will discover this is $1 - O(1/\text{poly}(n))$, which means that it is necessary to repeat the protocol a polynomial number of times to get constant soundness. Ito et al. [12] use this result, and give a two prover proof system for NEXP with a classical verifier which is resistant to entanglement with soundness of just $1 - 2^{-\text{poly}}$.

1.1 Our Results

An important assumption underlying the work on multi prover schemes is that the provers are not allowed to pass information between them. The results of Kitaev and Watrous [16] and Preda [22] could lead us to believe that a proof system with a quantum verifier and two provers who can pass classical information between them is limited to EXP. Surprisingly, this is not the case (assuming EXP $\neq$ NEXP). We show that:

Theorem 1.1. *Let V be a polynomial time verifier that can exchange quantum messages with two computationally unbounded provers. The provers share no entanglement, but can freely communicate classically between them. Then for any $L \in$ NEXP there is a two round protocol for the verifier and provers such that for any string x*

- *(completeness) If $x \in L$ then there are two prover strategies such that V will accept x with probability 1.*
- *(soundness) If $x \notin L$ then for any two prover strategies the probability that V will accept x is at most c for some constant $0 < c < 1$.*

The communication between the verifier and the provers is is polynomial in the length of the input[1].

We note that augmenting the provers in our model with unlimited entanglement gives something which is contained in EXP [16], as this is equivalent to quantum communication and thus to a single quantum prover. Bounding the verifier to be classical

[1] Equivalently we can state our result for NP, bounding the communication to be logarithmic.

would limit us to languages in PSPACE [24] (as this scenario is equivalent to a single prover and a classical verifier), so both conditions are necessary.

An important reason to study quantum MIPs is to better understand quantum theory, especially the power of entanglement. Surprisingly, our result, which is stated in a model with no entanglement between the provers, is based on following the entanglement between the provers and the verifier. Each message the verifier sends is a superposition of two classical queries. Measuring the message would ruin the superposition, and will be caught by the verifier. However, a strategy which does not measure the message "enough" does not extract enough useful classical information, and prevents the provers from coordinating answers via the classical channel. Most of the paper follows the amount of entanglement between the verifier and the provers during the protocol, making sure that either the provers do not extract enough information to answer with very high probability, or they have some chance of getting caught.

Another important task in quantum theory is to study the power of Local Operations and Classical Communication. In this model two entities are allowed to perform quantum computation locally, but they are limited to classical communication. They are usually cooperating to achieve some task, such as transforming one shared entangled state to another one. Characterizing the set of actions that can be performed by them is a very hard problem (in fact the large number of entanglement monotones comes from the fact that we do not have a complete characterization of what can be done in the LOCC model). One can view the results of this work as trying to look at the LOCC model from a complexity-theoretic perspective where the two entities are provers who are trying to fool the verifier.

1.2 Related Work

It is interesting to view the results of this paper in light of the complexity class QMA(2), defined by Kobayashi, Mastumoto and Yamakami [18]. Intuitively, this is the class of languages which can be recognized by a polynomial time quantum verifier with two unentangled quantum witnesses (the verifier is promised that the witnesses are unentangled). While there is no classical analog for this problem (having two classical witnesses is still NP), Liu, Christandl and Verstraete give evidence that QMA(2) strictly contains QMA [19]. Blier and Tapp [5] showed that a verifier can recognize an NP-complete language with soundness $1 - O(1/n^6)$. A constant soundness completeness gap in their results would imply our own. We note, however, that Aaronson et al. [1] give evidence towards QMA(2)$\subseteq$PSPACE, and therefore we do not expect that this is the case.

Private Information Retrieval schemes (PIRs) were studied by Chor et al. and by Kerenedis and DeWolf, among others [7, 15, 14]. The idea of using them for Multi Prover Protocols has been suggested by Cleve et al. [9]. Our protocol is in a sense a cheat sensitive PIR where the verifier can check whether the prover has tried to learn information. A similar quantum PIR scheme has been independently presented by Giovannetti, Lloyd and Maccone [11] in a different context. Cheat sensitive PIR's are usually characterized by their information disturbance tradeoff, which is defined as the probability of getting caught when the provers extract one bit of information on the query. However, in all existing PIR's this value is a lot less than a constant, even in the

simpler case of one server who wants to learn information on the query. As it is easy to find two assignments such that each clause is satisfied by one of them, even one bit of information can allow the provers to cheat the verifier. Moreover, in our protocol it is possible for the provers to extract a constant amount of information without getting caught. Therefore, instead of using information theoretic inequalities we tailor new bounds.

2 Preliminaries

We assume the reader is familiar with quantum computation (see [20] for basic notation).

Let $L \in$ NEXP. By standard PCP machinery, we can assume that given x the verifier has implicit efficient access to an exponentially long 3-SAT formula Φ, such that if $x \in L$ then Φ is satisfiable, and otherwise any assignment can satisfy at most a fraction of $1 - \gamma$ of the clauses for some constant $\gamma > 0$. We can also assume that each variable appears exactly 5 times, and each clause contains three different variables. Let C denote the set of clauses and V the set of variables. If a variable $v \in V$ appears in a clause $c \in C$ we write $v \in c$. Let $M = |C|$ denote the number of clauses and $N = |V|$ the number of variables. Let T be a truth assignment for Φ. For a variable $v \in V$, let $T(v)$ denote the value T assigns x. For a clause $c \in C$, if y contains the variables v_1^y, v_2^y, v_3^y, let $T(c) = T(v_1^c), T(v_2^c), T(v_3^c)$.

Alice (Bob) has a private Hilbert space H_A^p (H_B^p), with some finite arbitrarily large dimension t (we assume without loss of generality that the dimensions are identical). The messages between Alice (Bob) and the verifier will be sent by passing a state which is in a Hilbert space H_A^m (H_B^m). For convenience, we partition the private Hilbert space of the verifier into three parts, $H_v = H_v^{\text{aux}} \otimes H_A^v \otimes H_B^v$. The Hilbert spaces H_A^v, H_B^v will be used with messages sent to different provers, but they are private spaces that belong to the verifier. We let the verifier send and receive classical messages from Alice[2]. For the protocol we present, the dimensions of the Hilbert spaces used are $\dim(H_A^m) = 8M$, which would fit a clause y and the value an assignment T gives it, $T(y)$, $\dim(H_A^v) = M$, $\dim(H_B^m) = 2N$ which would fit a variable and the value it is assigned, and $\dim(H_B^v) = N$.

3 Algorithm

Let π be a probability distribution which chooses two clauses c, d uniformly at random from C, and two variables v, w such that v is chosen uniformly at random from the variables of c, and w is chosen uniformly from V. Figure 1 presents the protocol the verifier follows, a well as the answers expected from the provers.

Note that the verifier does not generate any entanglement between the provers - the states sent to Alice are unentangled with the ones sent to Bob, and they are measured separately. This means that it is possible to repeat the protocol in order to reduce the

[2]This can be done by using a larger space H_A^m, with the verifier measuring the part of the space which should be used for the classical message. Thus, this does not change the model, and is only done for clarity.

Protocol for VERIFIER:

1. Sample π to get c, d, v, w. Generate the states on $O(\log(N))$ qubits

$$\begin{aligned} \frac{1}{\sqrt{2}}(|cc\rangle + |dd\rangle) \otimes |000\rangle \quad &\in \quad H_A^v \otimes H_A^m \\ \frac{1}{\sqrt{2}}(|vv\rangle + |ww\rangle) \otimes |0\rangle \quad &\in \quad H_B^v \otimes H_B^m \end{aligned}$$

   Send Alice (Bob) the message space H_A^m (H_B^m), which consists of the last $m+3$ (respectively $n+1$) qubits.

2. Let T be a satisfying assignment for Φ (if one exists). Alice should apply the unitary which takes $|c\rangle \otimes |000\rangle \rightarrow |c\rangle \otimes |T(c)\rangle$ for any clause $c \in C$, and Bob should apply the unitary which takes $|v\rangle \otimes |0\rangle \rightarrow |v\rangle|T(v)\rangle$ for $v \in V$. Cheating provers may apply any local operations and classical communication (LOCC) protocol they want between them. Finally, Alice (Bob) returns the verifier the message space H_A^m (H_B^m).

3. The verifier sends Alice the classical values c, d, v, w. Alice returns 8 bits: $T(c), T(d), T(v), T(w)$. If Alice returned quantum values, the verifier measures them according to the standard basis.

4. The verifier does one of the following tests, each with probability $1/2$

   (a) **SWAP Test:** The verifier checks that the clause c is satisfied, and that $T(v)$ matches $T(c)$. He performs the SWAP test [6] between the state in $H_A^v \otimes H_A^m$ and $\frac{1}{\sqrt{2}}(|cc\rangle \otimes |T(c)\rangle + |dd\rangle \otimes |T(d)\rangle)$ and between the state in $H_B^v \otimes H_B^m$ and $\frac{1}{\sqrt{2}}(|vv\rangle \otimes |T(v)\rangle + |ww\rangle \otimes |T(w)\rangle)$, and accepts if all tests passed.

   (b) **Quantum Consistency Test:** The verifier measures H_B^v by projecting it on all the variables. If he sees w, the provers win. If he sees v, he projects H_B^m on $|v0\rangle, |v1\rangle$ and everything else. If the result is not either $|v0\rangle$ or $|v1\rangle$ Bob is caught. He does a similar check on $H_A^v \otimes H_A^m$, and rejects if $T(v)$ does not match $T(c)$.

Figure 1: The protocol applied by the verifier.

error probability. Also, note that the second check (the measurement) does not depend on the answers of the provers, and we can assume that it was done instead of sending Alice the classical values.

Completeness: With a common satisfying assignment the provers can apply the required quantum transformation, and all the tests will pass with probability 1.

4 Soundness of the Protocol

Most of this section will deal with the first test. The second test is only needed to make sure that the provers do not keep too much entanglement to the answers they send the verifier. This is handled in Subsection 4.6, and until then we assume that the SWAP test is applied. As this section is a little technical, we begin with an informal sketch of the main ideas.

4.1 Intuition

Consider the following two extreme strategies for the first round:

1. If Alice measures according to the standard basis, she can know one of the clauses which were sent by the verifier. However, such a measurement would destroy the entanglement between her and the verifier. Thus, she has a constant probability of getting caught in the second round of the protocol, regardless of the truth assignment she sends.

2. If both Alice and Bob apply unitary operations, then they are not utilizing the classical channel. Therefore, their success probability is related to the success probability of non-communicating provers, which is bounded.

There are two main obstacles in turning this observation into a proof. The first obstacle is that these are only two of the possible strategies available to the provers. For example, the provers might instead use a complicated protocol, which consists of many communication rounds between themselves, as well as exponentially weak measurements.

The second obstacle is that after the first round of the protocol, the verifier can not check whether the provers measured or not, as she does not know the truth assignment. To deal with the latter problem, the protocol has a second round, in which the provers give the verifier a classical description of the state she holds (and thus the verifier can verify that the provers did not measure too much). The two round protocol now resembles a PIR scheme (or a bit commitment scheme) in which the provers first commit to the assignment without knowing what questions the verifier asked, and then reveal the committed assignment.

We now go over the argument, as it appears in the rest of the paper. The first step is to purify the verifier. This enables us to consider the verifier's measurement after the provers have acted, and simplifies the analysis, without changing anything in the protocol. The second step is to strengthen the provers, and allow them to perform any separable outcome measurement (instead of an LOCC protocol). This modification can

only help cheating provers. The first round of the modified protocol now consists of the following stages.

1. The verifier generates a uniform superposition over the questions she can ask, namely she creates the state

$$\frac{1}{3M\sqrt{N}} \sum_{c,d,v,w \ , \ v \in c} |cdvw\rangle \otimes \frac{1}{\sqrt{2}} \left(|cc000\rangle + |dd000\rangle\right) \otimes \frac{1}{\sqrt{2}} \left(|vv0\rangle + |ww0\rangle\right)$$

   where the first register is called the auxiliary register, and is used to purify the protocol.

2. The provers perform a single separable measurement. Let k denote the result of this measurement.

3. The verifier measures the auxiliary register. This undoes the purification, and chooses what part of the assignment the verifier checks.

As the provers' measurement can have an arbitrarily large number of outcomes (where each outcome occurs with very small probability), we cannot afford to argue that the provers will be caught only on the probable outcomes. Instead, we argue that for every outcome of the provers' measurement, the provers have a constant probability of being caught. For the rest of this subsection, we fix a specific outcome k.

We begin by showing the following lemma, which is stated informally

Lemma 4.1 (Informal Lemma)**.** *Given measurement result k by the provers, either they are caught with some constant probability, or the distribution on the verifier's measurement results on the auxiliary register is close to uniform.*

Where the auxiliary register is the register which is used to purify the protocol. The main idea is to show that if the probability distribution is ϵ far from uniform between clauses, then there is a constant probability that after measuring the auxiliary register the verifier will be left in the state

$$\sqrt{p}|cc\psi_c\rangle + \sqrt{1-p}|dd\psi_d\rangle$$

for some $p \neq 1/2$ and states ψ_c, ψ_d (a similar lemma holds for variables). In this case, there is a probability that the provers will fail the SWAP test, regardless of what Alice sends in the second round of the protocol. Proving the global lemma when the provers can take correlated actions, and with the required parameters is a little involved, and takes most of Subsections 4.4, 4.5.

It is intuitive that if the provers perform a unitary transformation, their success probability is bounded. One way to prove this is to show a reduction from any strategy of the communicating provers (in which a unitary is applied in the first round), to a strategy for the provers in the classical game, where the verifier and the communication channels are classical, and the provers are not communicating. An obstacle in showing such a reduction is that we need to transform a query made by the classical verifier in the one round protocol into messages sent in the two round protocol. In particular,

in the second round of the quantum protocol, the verifier tells Alice classically what part of the assignment she will verify, and this cannot be simulated by the classical provers. Therefore, our reduction can only rely on the questions sent by the verifier in the first round of the quantum protocol (since they can be generated by the classical provers). Such a reduction can succeed if the quantum provers are already committed to an assignment after the first round, and thus Alice's answers in the second round is determined before getting the extra information from the verifier. Perhaps surprisingly, this amounts to showing that after the first round of the protocol, the provers are unentangled with the verifier.

Subsection 4.6 is devoted to analyzing the entanglement between the provers and the verifier after the first round of the protocol is over, that is, after Alice and Bob send back their answers, and the verifiers measures the register which is used to purify the protocol. We show that if the entanglement between Alice and the verifier in this stage is greater than some constant, then the provers have a constant probability of failing the quantum consistency test. This happens because if the state sent by Alice is entangled with Alice's private qubits, then there is a probability that the truth assignment measured by the verifier on the qubits sent by Alice will not match the truth assignment measured by the verifier on the qubits sent by Bob (here we rely on the fact that Alice and Bob are unentangled). This argument does not rely on the classical information sent by Alice in the second round. A standard argument enables us to assume that the provers are not entangled with the verifier at all after the first round.

To finish the proof, we consider the classical multi-prover game, in which Charlie and Diane are two noncommunicating provers who are trying to fool a classical verifier. We begin by presenting a reduction from any measurement result k (which doesn't generate entanglement after the first round) to a strategy for Charlie and Diane. The reduction we use for Charlie (respectively Diane) is to consider the residual state of Alice (respectively Bob) and the verifier after the first round of the quantum protocol, and to send the truth assignment which would match its classical description.

The success probability of Charlie and Diane in the classical strategy is related to the measurements done by the verifier in the quantum protocol. In particular, it is related to the *number* of different tuples that the verifier can measure in the auxiliary register (the register which is used to purify the protocol) on which the provers have high success probability.

If the strategy employed by the provers is a unitary transformation, all the measurement results on the auxiliary register are equally likely. Thus, the number of measurement results of that register on which the provers succeed is related to the success probability of the provers. This gives a bound on the success probability of the provers in this case. This reduction is robust; even if almost all measurement results of the auxiliary register are equally likely, there is still some relation between the number of measurement results of this register on which the quantum provers are likely to succeed, and the overall success probability of the classical provers. As we already showed that if (given k) the verifier has a far from uniform distribution on the measurement of the auxiliary register then the provers have constant probability of getting caught, this finishes the soundness proof.

4.2 The Modified Protocol

As stated above, the first modification is to purify the sampling of π, postponing it until after the provers act on the information. It uses H_v^{aux} with $\dim(H_v^{\text{aux}}) = M^2N^2$. The verifier generates

$$\begin{aligned}\psi_\pi &= \sum_{c,d\in C}\sum_{v\in c}\sum_{w\in V}(|cdvw\rangle \\ &\quad \otimes\frac{1}{\sqrt{2}}(|cc\rangle + |dd\rangle)\otimes|000\rangle \\ &\quad \otimes\frac{1}{\sqrt{2}}(|vv\rangle + |ww\rangle)\otimes|0\rangle) \\ &\in H_v^{\text{aux}}\otimes H_v^A\otimes H_m^A\otimes H_v^B\otimes H_m^B\end{aligned}$$

ignoring normalization factors.

As before, the verifier sends Alice (Bob) the Hilbert space H_m^A (H_m^B). Alice and Bob act on the message spaces they get and return H_m^A, H_m^B to the verifier. The verifier measures H_v^{aux} to get c, d, v, w and sends them to Alice as in Protocol 1. This modification does not change the cheating power of the provers (they cannot tell what protocol is being used).

The second modification is to replace the LOCC done by the provers in the first stage with a single joint separable measurement. Bennet et al. and Barnum [4, 3] proved that this is strictly stronger than LOCC. In particular they showed how to transform any LOCC protocol into such a measurement. As the provers are not entangled, we can assume that their private spaces are initialized with the state $|0\dots0\rangle$. Letting $\rho = |\psi_\pi\rangle\langle\psi_\pi|$, the provers' operation now becomes applying a measurement with operators

$$(I_{M^2N^2}\otimes I_M\otimes A_k\otimes I_N\otimes B_k)^\dagger(I_{M^2N^2}\otimes I_M\otimes A_k\otimes I_N\otimes B_k)$$

where I_p is the $p\times p$ identity matrix, A_k is an $8Md\times 8Md$ matrix, B_k is a $2Nd\times 2Nd$ matrix and

$$\sum_k (A_k\otimes B_k)^\dagger(A_k\otimes B_k) = I_{16NMd^2}$$

The Hilbert spaces H_m^A, H_m^B are then returned to the verifier.

We now calculate the probability that the verifier measured a tuple $r = (c, d, v, w)$, conditioned on the fact that the provers' measurement result was k. Denote

$$A_k(c) = \text{tr}(A_k(|c\rangle\langle c|\otimes I)A_k^\dagger)$$

where we are tracing over the private qubits of the prover and the qubits which define the assignment, and similarly $B_k(v) = \text{tr}(B_k(|v\rangle\langle v|\otimes I)B_k^\dagger)$. We shall omit the subindex k when it is clear from context, using $A(c)$ and $B(v)$. In Appendix A, we prove that for $c\neq d, v\neq\tilde{w}$

$$\Pr(c,d,v,w|k) = \frac{(A_k(c)+A_k(d))(B_k(v)+B_k(w))}{\sum_{\tilde{c},\tilde{d}\in C,\tilde{v}\in\tilde{c},\tilde{w}\in V}\Pr(\tilde{c},\tilde{d},\tilde{v},\tilde{w}|k)} \tag{1}$$

where if $c = d$ the numerator changes to $4A_k(c)(B_k(v) + B_k(w))$, and similarly for v, w.

We give some intuition for Equation (1). The numerator is the product of two factors, because when the verifier measures before the provers (which is physically equivalent) the provers are unentangled and operate on different spaces. Therefore the probability of k is just the $\text{tr}(A_k \rho A_k^\dagger) \cdot \text{tr}(B_k \rho B_k^\dagger)$. Alice's factor is composed of two terms, because tracing out the verifier Alice just gets a mixed state of $\frac{1}{2}|c\rangle\langle c| + \frac{1}{2}|d\rangle\langle d|$. Denote

$$W_A = W_{A_k} = \sum_i A_k(i) = \text{tr}(A_k)$$

$$W_B = W_{B_k} = \sum_i B_k(i) = \text{tr}(B_k)$$

$$\tilde{W} = \tilde{W}_k = \sum_{c \in C, v \in c} A_k(c) B_k(v)$$

In Appendix A we bound the denominator of (1) to get the following:

$$\Pr(c, d, v, w|k) \geq \frac{A_k(c)B_k(v) + A_k(c)B_k(w) + A_k(d)B_k(v) + A_k(d)B_k(w)}{2NM\tilde{W}_k + 22MW_{A_k}W_{B_k}} \quad (2)$$

4.3 Auxiliary Lemmas

In subsections 4.3-4.5, we assume in all the probability computations that the verifier performed the SWAP Test. This changes all the probabilities by a factor of two, and is done to simplify the notation.

We show that if $A_k(c)$ is not almost uniform in c, then for certain values of c, d, Alice has a good chance of failing the SWAP test. Formally:

Lemma 4.2. *Assume $A(c) \geq \alpha A(d)$, $\alpha > 1$. Then for any assignment T, the probability that the verifier will catch Alice cheating in the SWAP test is at least $\frac{1}{2} - \frac{\sqrt{\alpha}}{1+\alpha}$.*

The intuition is that the super-operator which acts on the state diminishes the entanglement between H_A^v and H_A^m. Therefore, there is a failure probability regardless of the assignment Alice will send in the second round of the protocol.

Proof. Let

$$\sigma \quad = \quad \text{tr}_{H_A^p}(I \otimes A_k)\rho(I \otimes A_k)^\dagger$$

where we ignored normalization factors. For a given assignment $T(c) \in \{0,1\}^3$ and $T(d) \in \{0,1\}^3$, let

$$|\psi\rangle \quad = \quad \frac{1}{\sqrt{2}}\left(|cc\rangle|T(c)\rangle + |dd\rangle|T(d)\rangle\right)$$

Taking $\delta = \sqrt{\langle\psi|\sigma|\psi\rangle}$, the fidelity between $|\psi\rangle$ and ρ, the SWAP test has probability at least $\frac{1-\delta^2}{2}$ to distinguish between them [6].

To calculate σ, we utilize the result in Appendix A. Write $M' = 8M + 8$. Since ρ consists of four elements in a rectangle

$$(M'c, M'c), (M'c, M'd), (M'd, M'c), (M'd, M'd)$$

differentiated by a distance of at least M', the nondiagonal elements do not contribute to the trace.

$|\psi\rangle$ is an equal superposition of two base vectors, one corresponding to the base state $|ccT(c)\rangle$ and the other to $|ddT(d)\rangle$. Thus the multiplication is effectively the sum of four elements arranged in a rectangle (multiplied by $1/2$). To calculate each of these four elements, we turn to Appendix A. Since, in the tensor product $I \otimes A_k$, any cell whose two coordinates differ by at least $M' = 8M + 8$ is zero, we can simplify and get:

$$\sigma[ccT(c), ccT(c)] = \text{tr}(A|cT(c)\rangle\langle cT(c)|A)$$
$$\sigma[ccT(c), ddT(d)] = \text{tr}(A|cT(c)\rangle\langle dT(d)|A)$$

where we abuse notation a little, and treat c as its numerical index in the list of clauses. We can now write the elements, sticking to the convention that the first m qubits describe the verifier's private space, the next m fit the clause in the message space and the last three fit the value of the assignment:

$$\begin{aligned}
\sigma(M'c + a, M'c + a) &= \sum_{i=1}^{t}\sum_{j=1}^{t} |A[8pc + ta + i, 8tc + ta + j]|^2 \\
\sigma(M'd + b, M'd + b) &= \sum_{i}\sum_{j} |A[8td + tb + i, 8td + tb + j]|^2 \\
\sigma(M'c + a, M'd + b) &= \sum_{i}\sum_{j} \overline{A[8tc + ta + i, 8tc + ta + j]} \cdot A[8tdy + tb + i, 8td + tb + j] \\
\sigma(M'd + b, M'c + a) &= \sum_{i}\sum_{j} \overline{A[8td + tb + i, 8td + tb + j]} \cdot A[8tc + ta + i, 8ty + ta + j]
\end{aligned}$$

Note that as $AA^\dagger$ is a measurement operator, we have that $A(y) \leq 1$, so $A(\tilde{y}) \leq 1/p$. Now calculating, reindexing by $s = M'y + T(y)$ and $r = M'\tilde{y} + T(\tilde{y})$, and folding the sum into the expression, we get:

$$\begin{aligned}
&\frac{\sigma[s,s] + |\sigma[r,r]| + \sigma[s,r]\sigma[r,s]}{2(A(y) + A(\tilde{y}))} \leq \\
&\frac{A(y) + A(\tilde{y}) + 2\sqrt{A(y)A(\tilde{y})}}{2(A(y) + A(\tilde{y}))} = \\
&\frac{1}{2} + \frac{\sqrt{A(y)A(\tilde{y})}}{A(y) + A(\tilde{y})} \leq \frac{1}{2} + \frac{\sqrt{\alpha}}{\alpha + 1} < 1
\end{aligned}$$

The first inequality follows from the Cauchy-Schwarz inequality, which states that $\sigma[s,r] \leq \sqrt{A(y)A(\tilde{y})}$, and similarly for $\sigma[r,s]$. Further, $A(y)$ is the sum of $\sigma[s,s]$ over all possible assignments for y, and thus dominates $\sigma[s,s]$.

The last inequality follows from the AM-GM inequality. More precisely, since the ratio is at least α, the extreme value is achieved when it is exactly α, which (when substituting) gives what we need. When $\alpha \geq \sqrt{2}$, this gives that the success probability of the provers is at most $1/2 + 2^{1/4}/(1+\sqrt{2}) \leq 0.993$. When $\alpha \geq 2$, we get $1/2 + \sqrt{2}/3 \leq 0.975$. We use these numbers later to calculate the constants, but they are otherwise not important. □

If the condition of lemma 4.2 holds, we say that the measurement has α-*damaged* the state. An analogous lemma holds for Bob.

Lemma 4.3. *If there exists a set $D \subset Y \times \tilde{Y} \times X \times \tilde{X}$ such that*

1. *For any $r = (c, d, v, w) \in D$ we have $v \in c$, and either Alice or Bob α-damage r for some constant α.*
2. *$\sum_{r \in D} \Pr(r|k) > \epsilon_D$ for some constant ϵ_D*

Then at least one of the provers gets caught in the SWAP test with probability $\epsilon_D \left(\frac{1}{2} - \frac{\sqrt{\alpha}}{1+\alpha}\right)$.

In this case we say that D is an $(\epsilon_{\mathrm{D}}, \alpha)$ *bad set*.

Proof. Given k, the probability that the verifier measures an element of D (in the second stage) is greater than ϵ_{D}. For any such element, at least one of the provers has probability greater than $\frac{1}{2} - \frac{\sqrt{\alpha}}{1+\alpha}$ to be caught. □

The proof of Theorem 1.1 now splits into two cases, according to the dominant term in the denominator of expression (2). If $NM\tilde{W}$ is the dominant term, then bad sets exist; this is shown in Section 4.4. If $NM\tilde{W}$ is small and MW_AW_B is the dominant term, then either a bad set exists or there is a "nice" set of clauses and variables which together give a cheating strategy for the classical protocol. This is shown in Sections 4.5–4.7.

4.4 Large $NM\tilde{W}_k$

Theorem 4.4. *If $NM\tilde{W}_k \geq 100MW_{A_k}W_{B_k}$ then at least one of the provers fails the SWAP test with probability greater than some constant, given by $\epsilon_{NMW_k} = \min\left\{\frac{1}{6.96 \cdot 10^9}, \frac{1}{4.2 \cdot 10^7}\right\}$.*

To prove this, we show that in order to pass the SWAP test, the weights of the clauses (Lemma 4.5) and the variables (Lemma 4.6) must be relatively uniformly distributed. On the other hand, the factor $NM\tilde{W}_k$ can be dominant only when they are not uniformly distributed. For the rest of this subsection, we assume the premise of Theorem 4.4.

We begin the proof by looking at the probability of each clause in the modified protocol given measurement result k (from now on we omit the subindex). If there are two sets of clauses such that

1. One set contains "heavy" clauses and the other contains "light" clauses

2. The "heavy" set contains a fraction of the weight

3. The "light" set contains a fraction of the clauses

then we can build a bad set. If this is not the case, then most clauses have almost the same probability. We begin generating those sets by dividing the clauses according to their weight, according to factors of two.

For $\tilde{c} \in C$, let $u(\tilde{c}) = \Sigma_{\tilde{v} \in \tilde{c}} A(\tilde{c}) B(\tilde{v})$, and for $S \subset C$, let $U(S) = \sum_{\tilde{c} \in S} u(\tilde{c})$. Let

$$S_i = \left\{ \tilde{c} : \frac{\tilde{W}}{2^{i+1}} < u(\tilde{c}) \leq \frac{\tilde{W}}{2^i} \right\}$$

Lemma 4.5. *If there exists an index j such that*

$$\begin{aligned} \sum_{i=0}^{j-1} U(S_i) &> \tilde{W}/100 \\ \sum_{i=j+1}^{\infty} U(S_i) &> \tilde{W}/100 \end{aligned}$$

then the provers get caught with constant probability $\frac{1}{6.96 \cdot 10^9}$, generated from a $(\frac{1}{4.8 \cdot 10^7}, \sqrt{2})$ bad set. We call the index j a separating *index.*

Proof. We construct such a bad set D. For any clause $\tilde{c}$ in variables v_1, v_2, v_3, let $\text{vmax}(\tilde{c})$ denote the variable $v_i \in \tilde{c}$ such that $B(v_i) = \max\{B(v_k) : k = 1, 2, 3\}$, and define $\text{vmin}(\tilde{c})$ analogously. Let $S_{\text{up}} = \cup_{i=0}^{j-1} S_i$, and $S_{\text{down}} = \cup_{i=j+1}^{\infty} S_i$. As $U(S_{\text{down}}) > \tilde{W}/100$, and S_{down} consists of "light" clauses, we must have $|S_{\text{down}}| > M/100$. Partition S_{down} arbitrarily into two sets S_l and S_r, such that $|S_l|, |S_r| \geq M/200$. The idea is that each clause in S_{up} will contribute $|S_l| \cdot |S_r|$ elements to D.

For each $c_{\text{up}} \in S_{\text{up}}, c_l \in S_l, c_r \in S_r$, we have $u(c_{\text{up}}) > 2u(c_l)$ and $u(c_{\text{up}}) > 2u(c_r)$. Taking the maximal variable in the sum for c_{up}, and the minimal variable for c_l, c_r, we get:

$$\begin{aligned} A(c_{\text{up}}) B(\text{vmax}(c_{\text{up}})) &> 2A(c_l) B(\text{vmin}(c_l)) \\ A(c_{\text{up}}) B(\text{vmax}(c_{\text{up}})) &> 2A(c_r) B(\text{vmin}(c_r)) \end{aligned}$$

Assume WLOG that $A(c_l) < A(c_r)$. Then

$$A(c_{\text{up}}) B(\text{vmax}(c_{\text{up}})) > 2A(c_l) B(\text{vmin}(c_r))$$

So in the tuple $(c_{\text{up}}, c_l, \text{vmax}(c_{\text{up}}), \text{vmin}(c_r))$ at least one of the provers damages the state by at least $\sqrt{2}$. We add this tuple to D. Note that we have added a (distinct) element to D for each of the $|S_l| \cdot |S_r|$ choices of c_l, c_r, as desired. Let $D(c_{\text{up}})$ denote the elements contributed to D by c_{up}.

The next step is to prove that D has constant probability. According to (1) the probability of D is bounded below by

$$\Pr(D) \geq \frac{\sum_{(c,d,v,w) \in D} (A(\tilde{c}) + A(\tilde{d}))(B(v) + B(w))}{\sum_{\tilde{c}, \tilde{d} \in C, \tilde{v} \in \tilde{c}, \tilde{w} \in V} \Pr(\tilde{c}, \tilde{d}, \tilde{v}, \tilde{w} | k)}$$

Under the conditions of Theorem 4.4, the denominator is bounded by

$$\sum_{\tilde{c},\tilde{d}\in C,\tilde{v}\in\tilde{c},\tilde{w}\in V} \Pr(\tilde{c},\tilde{d},\tilde{v},\tilde{w}|k) \quad < \quad 22MW_AW_B + 2NM\tilde{W} < 4NM\tilde{W}$$

and summing the probabilities of the elements in D, we get:

$$\begin{aligned}
\sum_{(c,d,v,w)\in D} \Pr(c,d,v,w|k) &= \sum_{c_{\text{up}}} \sum_{(d,w)\in D(c_{\text{up}})} \Pr((c_{\text{up}},d,\max(c_{\text{up}}),w)|k) \\
&\geq \frac{1}{4NM\tilde{W}} \sum_{c_{\text{up}}} \sum_{(d,w)\in D(c_{\text{up}})} A(c_{\text{up}})B(\max(c_{\text{up}})) \\
&= \frac{1}{4NM\tilde{W}} |S_l| \cdot |S_r| \sum_{c_{\text{up}}} A(c_{\text{up}})B(\max(c_{\text{up}})) \\
&\geq \frac{1}{4NM\tilde{W}} |S_l| \cdot |S_r| \sum_{c_{\text{up}}} u(c_{\text{up}})/3 \\
&\geq \frac{1}{4NM\tilde{W}} |S_l| \cdot |S_r| \tilde{W}/300 \\
&\geq \frac{1}{4NM\tilde{W}} \cdot \frac{M}{200} \cdot \frac{M}{200} \cdot \frac{\tilde{W}}{300} \\
\text{(because } M > N\text{)} \qquad &> \frac{NM\tilde{W}}{4NM\tilde{W} \cdot 200 \cdot 200 \cdot 300} = \frac{1}{4.8 \cdot 10^7}
\end{aligned}$$

where the first inequality comes from taking one out of the four terms in (2). □

If the condition of Lemma 4.5 does not hold, then there must be an index j such that $U(S_j) + U(S_{j+1}) > 0.98\tilde{W}$. Define $F = S_j \cup S_{j+1}$. We want to prove that most of the clauses in F have similar $A(\cdot)$ values. This is true because they are almost equiprobable, so if there are two large sets with different values then these sets together will generate a bad set. Remembering that $W_A = \sum_{\tilde{c}\in C} A(\tilde{c})$, we partition the clauses in F, in a way similar to that employed in the previous lemma:

$$T_i = \left\{ \tilde{c} \in F : \frac{W_A}{2^{i+1}} < A(\tilde{c}) \leq \frac{W_A}{2^i} \right\}$$

Lemma 4.6. *If there exists an index j such that $\sum_{i=0}^{j-1} |T_i| > |F|/100$, and $\sum_{i=j+1}^{\infty} |T_i| > |F|/100$, then the first prover gets caught with constant probability $\frac{1}{4.2\cdot 10^7}$, generated from a $(\frac{1}{1.2\cdot 10^6}, 2)$ bad set.*

Proof. Let $T_{\text{up}} = \cup_{i=0}^{j-1} T_i$, $T_{\text{down}} = \cup_{i=j+1}^{\infty} T_i$. Note that any clause from T_{up} at least 2-damages any clause in T_{down}. Take

$$D = \cup_{c_{\text{up}}\in T_{\text{up}}} \{\{c_{\text{up}}\} \times T_{\text{down}} \times \{\text{vmax}(c_{\text{up}})\} \times V\}$$

Note that $|T_{\text{down}}| > 0.98M/100 > M/200$, and as $T_{\text{up}} \subset F$, we have $U(T_{\text{up}}) \geq \frac{0.98\tilde{W}}{400} \geq \frac{\tilde{W}}{500}$, and thus

$$\Pr(D) \geq \frac{1}{4NM\tilde{W}} |T_{\text{down}}| N \frac{\tilde{W}}{1500} \geq \frac{1}{1.2 \cdot 10^6}$$

□

If the conditions of Lemmas 4.5 and 4.6 do not hold, then

$$\exists i : |T_i| + |T_{i+1}| > 0.98|F| \geq 0.98^2 M > 0.96M$$

Let $G = T_i \cup T_{i+1}$. As $G \subset F$, and as $\forall c_1, c_2 \in F : u(c_1) < 4u(c_2)$ we have

$$U(G) > 0.25 \cdot 0.98 U(F) > 0.25 \cdot 0.98^2 \tilde{W} \tag{3}$$

Note $\sum_{\tilde{c} \in G} \sum_{\tilde{v} \in \tilde{c}} B(\tilde{v}) \leq 5W_B$, as each variable appears 5 times. Also, since $\forall \tilde{c} \in G : A(\tilde{c}) > W_A/2^{i+1}$

$$0.96M \frac{W_A}{2^{i+1}} < \frac{W_A}{2^{i+1}}|G| < \sum_{\tilde{c} \in G} A(\tilde{c}) < W_A \tag{4}$$

Putting this together, we get

$$\begin{aligned} 0.25 \cdot 0.98^2 \tilde{W} &\overset{(3)}{<} U(G) = \sum_{\tilde{c} \in G} u(\tilde{c}) = \sum_{\tilde{c} \in G} \sum_{\tilde{v} \in \tilde{c}} A(\tilde{c}) B(\tilde{v}) \\ &\leq \sum_{\tilde{c} \in G} \sum_{\tilde{v} \in \tilde{c}} \frac{W_A}{2^{i-1}} B(\tilde{v}) = \frac{4W_A}{2^{i+1}} \sum_{\tilde{c} \in G} \sum_{\tilde{v} \in \tilde{c}} B(\tilde{v}) \\ &\leq \frac{20 W_A W_B}{2^{i+1}} \overset{(4)}{\leq} \frac{20 W_A W_B}{0.96M} \leq \frac{20 W_A W_B}{0.96N} \end{aligned}$$

But $NM\tilde{W} \geq 100 M W_A W_B$, which is a contradiction. This proves Theorem 4.4.

4.5 Small $NM\tilde{W}$

In this subsection we handle those values of k for which the premise of Theorem 4.4 does not hold, namely, $NM\tilde{W} < 100 M W_A W_B$. Define

$$S_i = \left\{ \tilde{c} \in C : \frac{W_A}{2^{i+1}} \leq A(\tilde{c}) < \frac{W_A}{2^i} \right\}$$

For a set $S \subset C$, let $W(S) = \Sigma_{\tilde{c} \in S} A(\tilde{c})$. We want to define a separating index, as we did in the previous section. However, in this section, if a large set of clauses is roughly equiprobable we will need to show that it cannot be satisfied. To do this, a necessary condition is that the set is large with respect to $(1-\gamma)M$, where the PCP theorem gives us that either ψ is satisfiable or any truth assignment can satisfy at most $(1-\gamma)M$ clauses. This motivates the following definition:

Lemma 4.7. *If $NM\tilde{W} < 100 M W_A W_B$ and there exists an index i such that*

$$\sum_{j=0}^{i-1} W(S_j) > \gamma 10^{-4} W_A \bigwedge \sum_{j=i+1}^{\infty} |S_j| > \gamma 10^{-4} M \tag{5}$$

then Alice is caught cheating with probability $\frac{\gamma^2}{2.6 \cdot 10^{12}}$, generated from a $\left(\frac{\gamma^2}{7.4 \cdot 10^{10}}, 2\right)$ bad set.

Proof. Let $S_{\text{up}} = \cup_{j=0}^{i-1} S_j$, $S_{\text{down}} = \cup_{j=i+1}^{\infty} S_j$. Let

$$D = \cup_{c \in S_{\text{up}}} \cup_{v \in c} \{c\} \times S_{\text{down}} \times \{v\} \times V$$

Every $(c, d, v, w) \in D$ is 2-damaged by Alice. On the other hand, by inequality (2) we get:

$$\begin{aligned} \Pr(c, d, v, w|k) &\overset{(2)}{\geq} \frac{(A(c) + A(d)(B(v) + B(w))}{22MW_AW_B + 2NM\tilde{W}} \\ &\geq \frac{A(y)B(\tilde{x})}{22MW_AW_B + 2NM\tilde{W}} \\ &\geq \frac{A(y)B(\tilde{x})}{222MW_AW_B} \end{aligned}$$

Summing this over D gives

$$\begin{aligned} \Pr(D) &\geq \sum_{c \in S_{\text{up}}} \sum_{v \in c} \sum_{d \in S_{\text{down}}} \sum_{w \in V} \frac{A(c)B(w)}{222MW_AW_B} \\ &\geq \sum_{c \in S_{\text{up}}} \frac{3 \cdot 10^{-4} \gamma MW_B A(c)}{222MW_AW_B} \\ &\geq \frac{3\gamma^2 W_A}{222 \cdot 10^8 W_A} \geq \frac{\gamma^2}{7.4 \cdot 10^{10}} \end{aligned}$$

□

Lemma 4.8. *If $NM\tilde{W} < 100MW_AW_B$ and the second condition of Lemma 4.7 does not hold, then there exists an index i such that for $F = S_i \cup S_{i+1}$ we have*

$$\begin{aligned} |F| &\geq (1 - 0.0002\gamma)M \\ W(F) &\geq (1 - 0.0002\gamma)W_A \end{aligned}$$

and $A(c) \geq W_A/(5M)$ for all $c \in F$.

The proof of the lemma is based on the fact that if the set of heavy clauses is light then it has to be small, and if the set of light clauses is small then it must be light. All the rest of the clauses are in F, and therefore it is both heavy and large.

Proof. Choose r to be the smallest index for which the first half of the condition does hold, i.e., $\sum_{j=0}^{r-1} W(S_j) > \gamma 10^{-4} W_A$. Then the second half of the condition cannot hold, i.e.

$$\sum_{j=r+1}^{\infty} |S_j| \leq \gamma 10^{-4} M$$

Take $i = r - 1$ (note that $r \neq 0$ because otherwise the first half of the condition does not hold). So:

$$\begin{aligned} |S_i| + |S_{i+1}| &= M - \sum_{j=0}^{i-1} |S_j| - \sum_{j=i+2}^{\infty} |S_j| \\ &\geq M - \sum_{j=0}^{i-1} |S_j| - \gamma 10^{-4} M \\ &\geq M - \gamma 10^{-4} M - \gamma 10^{-4} M \end{aligned}$$

where the last inequality follows since the total weight $\sum_{j=0}^{i-1} W(S_j) < \gamma 10^{-4} W_A$, but each clause in the sets S_j contributes at least $2^{-i} W_A$ to $W(S_j)$ while each clause outside of the sets S_j contributes at most $2^{-i-1} W_A$. A similar argument now applies to the weight $W(S_i) + W(S_{i+1})$. Finally, for each $c \in F$ we have

$$A(c) \geq \frac{W(F)}{4|F|} \geq \frac{(1 - 0.0002\gamma) W_A}{4M} \geq \frac{W_A}{5M}$$

□

Lemma 4.9 provides the equivalent claims of lemmata 4.7, 4.8 for Bob. Define the sets T_i analogously to S_i:

$$T_i = \left\{ v \in V : \frac{W_B}{2^{i+1}} \leq B(v) < \frac{W_B}{2^i} \right\}$$

Lemma 4.9. *Either Bob gets caught cheating with probability $\frac{\gamma^2}{3.9 \cdot 10^{12}}$ which is generated from a $\left(\frac{\gamma^2}{1.1 \cdot 10^{11}}, 2 \right)$ bad set, or else there exists an index i such that for $G = T_i \cup T_{i+1}$ we have $|G| > (1 - 0.0002\gamma)N$, $\Sigma_{v \in G} B(v) \geq (1 - 0.0002\gamma) W_B$ and $\forall v \in G$, $B(v) \geq \frac{W_B}{5N}$.*

Proof. If no such index exists then there is a separating index i such that letting $T_{\text{up}} = \cup_{j=0}^{i-1} S_j$, $T_{\text{down}} = \cup_{j=i+1}^{\infty} S_j$, we have

$$\begin{aligned} \sum_{v \in T_{\text{up}}} B(v) &> 10^{-4} \gamma W_B \\ |T_{\text{down}}| &> 10^{-4} \gamma N \end{aligned}$$

Let

$$D = \cup_{v \in T_{\text{up}}} \cup_{c : v \in c} \{c\} \times C \times \{v\} \times T_{\text{down}}$$

Then

$$\begin{aligned}\Pr(D|k) &\geq \sum_{(c,d,v,w)} \frac{A(d)B(v)}{222MW_AW_B} \\ &\geq \sum_{v\in T_{\text{up}}} \frac{\gamma NW_AB(v)}{222MW_AW_B} \\ &\geq \frac{\gamma^2 NW_B}{2.22\cdot 10^{10}MW_A} \\ &\geq \frac{\gamma^2 M}{1.1\cdot 10^{11}M} = \frac{\gamma^2}{1.1\cdot 10^{11}}\end{aligned}$$

where we used the fact that each variable appears in the formula 5 times. As before, if a separating index does not exist most of the weight (and most of the variables) lie in two adjacent steps. As these steps are adjacent, the variables on them are almost equiprobable. □

We now define a set of clauses, which are all almost equiprobable, when considering the information the provers have about the clause, as well as information the provers have about variables. This combines the results of Lemmas 4.9,4.7 and 4.8.

$$H = \{c \in F : \forall v \in c, v \in G\} \tag{6}$$

As $|G| \geq (1 - 0.0002\gamma)N$, and each variable appears 5 times we have

$$|H| \geq (1 - 0.0002\gamma)M - 5\cdot 0.0002\gamma N \geq (1 - 0.002\gamma)M$$

So far we have proved that either the provers have a constant probability of getting caught, or the set H is very large, and all the legal tuples inside H have almost the same probability.

4.6 The Measurement Test

To finish the proof we later claim that if the provers can cheat very well when the measurement result is k, then A_k and B_k also define a good classical cheating strategy. This may not be true however, if, following the first round, the provers are strongly entangled to the verifier. The Measurement Test is designed to catch the provers if this is the case. In this subsection we assume that the verifier performed the Measurement Test. This changes the probabilities by a factor of 2, and saves the need to multiply all the computations by $1/2$. The goal of this subsection is to prove the following theorem.

Theorem 4.10. *If there is a protocol for QMIP with communicating provers where the provers fail with probability at most ϵ, then there is a protocol in which the provers fail with probability which is at most a constant times ϵ and after the first round of the communication they are unentangled with the verifier.*

To prove this theorem, we deal with two cases:

1. We show that if after the first round the entanglement between the verifier and the provers is greater than ϵ, the the provers are caught with constant probability.

2. If the entanglement after the first round between the verifier and the provers is small, then there exists a product state such that if we replace the state of the system after the first round with this product state the success probability of the provers will not change by much.

We begin by showing that making small changes to the state does not change the success probability by much, and thus it suffices to show that the state is close to a nonentangled state.

Lemma 4.11. *Denote by u the pure state of the system after the first round, and by $P(u)$ the success probability of the provers at that stage, if they are allowed to make any separable measurement. Then if $\langle v|u\rangle > 1-\epsilon$, then the success probability of the provers $P(v)$ if v we change the state of the system to v is at most $P(u)+\epsilon$.*

We omit the proof of this approximation lemma. It is based on the fact that the success probability of the protocol is derived from a probability distribution on measurement results, and these statistics cannot change by much if u, v are very close.

We look at B_k's effect on $|vv0\rangle$, although in the protocol it operates on $\frac{1}{\sqrt{2}}(|vv0\rangle + |ww0\rangle)$. Let

$$|\phi^B(v)\rangle = (I \otimes B_k)(|vv0\rangle \otimes |0_B\rangle) = \sum_{b,\tilde{v}} a^B_{\tilde{v}b}|v\tilde{v}b\rangle \otimes \psi^B_{\tilde{v}b}$$

where $|0_B\rangle$ is the initial state in Bob's Hilbert space H^p_B, and the $\tilde{v}$'s span a basis.

Lemma 4.12. *If $\sum_{b,\tilde{v}\neq v} |a^B_{\tilde{v}b}|^2 > \epsilon$, then Bob fails the Measurement Test with probability at least $\epsilon/2$.*

Proof. Regardless of $T(v)$ and the measurement done by the provers, If the verifier measures either $|v\tilde{v}0\rangle$ or $|v\tilde{v}1\rangle$ with $v \neq \tilde{v}$ the provers are caught cheating. This occurs with probability $\sum_{b,\tilde{v}\neq v} |a^B_{\tilde{v}b}|^2$. □

Using Lemma 4.11 we can assume that

$$|\phi^B(v)\rangle = \sum_b a^B_{vb}|vvb\rangle \otimes \psi^B_b$$

by changing the success probability by a tiny amount.

A similar lemma holds for the system held by Alice and the verifier. Formally, by a small change in the success probability, we can assume that the state in $H^v_A \otimes H^m_A \otimes H^p_A$ is of the form

$$|\phi^A(c)\rangle = \sum_{T(c)} a^A_{cT(c)}|ccT(c)\rangle \otimes \psi^A_{T(c)}$$

where the sum is over at most 8 elements (as there are three qubits for the truth assignment.)

Lemma 4.13. *If there are two different values $\alpha, \beta \in \{0,1\}^3$ such that $a^A_\alpha, a^A_\beta > \epsilon/8$, and c is sent in superposition with some d then the provers are caught with probability at least $\epsilon/372$.*

Proof. With probability at least $1/2$, the first measurement by the verifier returned c. Let v be a variable which is assigned different values by α and β. As we are in H (the variables are almost uniform), with probability at least $1/12$ we know that v was sent by the verifier to Bob. With probability $1/2$ the verifier will measure v and not w.

Note that the measurements made by the provers and the verifier commute (because they are on different spaces). Ignoring the answers of the provers, there is probability at least $\epsilon/8$ for the verifier to measure $|c\alpha\rangle$, and probability at least $\epsilon/8$ for $|c\beta\rangle$. As $\phi^B(v)$ and $\phi^A(c)$ are unentangled, there is probability at least $\epsilon/8$ that the measurement on $\phi^B(v)$ will not match the one on $\phi^A(c)$. In this case, the provers fail. □

Using the approximation of Lemma 4.11 we assume that

$$|\phi^A(c)\rangle = |ccT(c)\rangle \otimes \psi^A_{T(c)}$$

for a specific value $T(c)$. By a similar argument we can assume:

$$|\phi^B(v)\rangle = |vvT(v)\rangle \otimes \psi^B_{T(v)}$$

for some value $T(v)$. This finishes the proof of Theorem 4.10.

In the next subsection, we use Theorem 4.10 to generate a strategy for the classical game (namely using $T(x), T(y)$ as an assignment). This will show that a high success probability in our quantum variant implies a high success probability for the classical variant. As the classical success probability is bounded, this will give a bound for the quantum success probability. Before we begin, we go over the classical setting.

4.7 Classical Setting

Let Charlie and Diana be two classical provers who are faced with a classical verifier. The verifier sends Charlie a random clause c, and Diana a random variable v which appears in c. Charlie is expected to answer with the values that some satisfying assignment gives the variables in c, and Diana with the value that the same assignment gives v. Remember that according to the PCP theorem, either the formula is satisfiable, or there exists a global constant γ such that any assignment satisfies at most a fraction of $(1-\gamma)$ of the clauses. It can be shown that in the second case, the success probability of Charlie and Diana is bounded by $1 - \frac{\gamma}{3}$.

We assume that the measurement result k was such that H exists as in 6, and the provers are not entangled to the verifier (by Theorem 4.10) after the first round. This enables us to prove a reduction from the quantum case to the classical one. First, a simple lemma.

Lemma 4.14. *If $\langle u|v\rangle \leq 1/2$ and $|u| = |v| = |w| = 1$ then*

$$\langle u|w\rangle > 1-\epsilon \Longrightarrow \langle v|w\rangle < 1/2 + \frac{\sqrt{3\epsilon}}{2} - \frac{\epsilon}{2}$$

The proof follows from Taylor's approximation. A specific case: if $\epsilon < 0.01$ the bound is less than 0.99.

Let FP(c, d, v, w, k) denote the probability that the provers failed to convince the verifier, given the measurement results (c, d, v, w) and k.

Lemma 4.15. *If there is an index k, measurement operators A_k, B_k and a set of clauses $R \subset C$ such that*

1. $|R| \geq (1 - \epsilon_1)M$
2. $\forall c \in R : \forall v \in c : |\{(d, w) \in C \times V : FP(c, d, v, w, k) > \epsilon_3/2\}| < \epsilon_2 NM$
3. $\epsilon_3 < 1/200$
4. *After the first stage in the purified protocol there is no entanglement between the provers and the verifier (see Theorem 4.10)*

then there is a classical strategy for Charlie and Diana which gives them a success probability of at least $(1 - \epsilon_1)(1 - \epsilon_2)(1 - \epsilon_3)$.

Proof. Charlie and Diana can now simulate Alice and Bob. Charlie gets as an input a clause c from the verifier. He chooses a random clause d, and generates the quantum state

$$(I \otimes A_k)\frac{1}{\sqrt{2}}\left(|cc000\rangle + |dd000\rangle\right) \otimes |0\rangle$$

which is exactly the quantum state shared between Alice and the verifier in the quantum protocol. Since we assume that k passes the measurement test, this state can be written as

$$\frac{1}{\sqrt{2}}\left(|ccT(c)\rangle + |ddT(d)\rangle\right) \otimes |\text{garbage}\rangle$$

where the garbage qubits are in H_A^p, Alice's private space. Note that Charlie cannot simulate the second round of the protocol, in which Alice is being told c, v classically. Therefore, Charlie reports the standard basis state which is closest to $|T(c)\rangle$.

Diane simulates Bob in a similar manner. Note that if Charlie and Diane are computationally unbounded but are not quantum they can just simulate a quantum computer to compute what assignment they should send.

We analyze the success probability of Charlie and Diane. There are three events we need to consider:

1. The verifier in the classical game sends a clause and variable for which $A_k \otimes B_k$ succeeds with high probability.
2. Charlie and Diane succeed in simulating the classical answers Alice would return in the second round of the quantum game.
3. Alice and Bob succeed in the quantum game, conditioned on the fact that k is usually good for c, v.

If all three events happen, Charlie and Diane succeed.

As the verifier in the classical case is random, the probability that a $c \in R$ is at least $(1-\epsilon_1)$. Given that $c \in R$, a Markovian bound on the success probability of the provers gives the factor $(1-\epsilon_2)(1-\epsilon_3)$ (taking care of the first and third bad events). We concentrate on the second event, and show that it cannot happen if $\epsilon_3 < 1/200$. To analyze this, we expand the state after Charlie applied $I \otimes A_k$

$$\begin{aligned}(I \otimes A_k)\frac{1}{\sqrt{2}}\left(|cc000\rangle + |dd000\rangle\right)|0\rangle &= \sum_{i=0}^{7}\frac{1}{\sqrt{2}}\alpha_{i,c}|cc\rangle|i\rangle|g_{c,i}\rangle + \frac{1}{\sqrt{2}}\alpha_{i,d}|dd\rangle|i\rangle|g_{d,i}\rangle \\ &= \sum_{i=0}^{7}\frac{1}{\sqrt{2}}\alpha_{i,c}|cc\rangle|i\rangle \otimes |g_c\rangle + \frac{1}{\sqrt{2}}\alpha_{i,d}|dd\rangle|i\rangle \otimes |g_d\rangle\end{aligned}$$

where $g_{c,i}$, and $g_{d,i}$ are on the private space of the prover, and they are independent of i due to the assumption that Charlie is not entangled with the verifier after the first round. The state $|i\rangle$ is the truth assignment, and Charlie sends the i for which $\alpha_{i,c}$ is maximal. If in the quantum protocol Alice would choose a different truth assignment $T(c) \neq i$, it means that $\alpha_{c,T(c)} \leq \frac{1}{2}$. But in this case, Alice gets caught with probability greater than $1/200$: With probability $1/2$ the verifier performs the SWAP test, and with according to Lemma 4.14, the state Alice sent doesn't pass the test with probability at least $1/100$. A similar analysis holds for Diane and Bob, showing that if $\epsilon_3 \leq 200$ both classical provers correctly guess the answers Alice gives in the second round.

□

Lemma 4.16. *If the failure probability of Alice and Bob given result k is less than $\frac{\gamma^3}{5.55\cdot 10^{13}}$ then there exists a set R with the properties stated in Lemma 4.15, with $\epsilon_1 = 0.003\gamma$, $\epsilon_2 = \gamma 10^{-3}$ and $\epsilon_3 = \gamma 10^{-4}$.*

Proof. Since the failure probability is less than $\frac{1}{6.96\cdot 10^9}$, we must have, by Theorem 4.4, that $NM\tilde{W} < 100MW_AW_B$. By Lemmas 4.7, 4.8 and 4.9, we have a set H as in (6) such that $|H| \geq (1-0.002\gamma)M$, and

$$\begin{aligned}\forall y \in H &: A(y) > W_A/(5M) \\ \forall y \in H : \forall v \in c &: B(v) > W_B/(5N)\end{aligned}$$

Using inequality (2), this means that for any tuple (c,d,v,w) where $c \in H$ and $v \in c$

$$\begin{aligned}\Pr(c,d,v,w|k) &\geq \frac{A(c)B(v)}{222MW_AW_B} \\ &\geq \frac{W_AW_B}{25MN\cdot 222MW_AW_B} \\ &= \frac{1}{5.55\cdot 10^3 NM^2}\end{aligned}$$

Denote

$$\begin{aligned}L(c,v) &= \{(d,w) : \mathrm{FP}(c,d,v,w,k) > 10^{-4}\gamma\} \\ H_{\text{fail}} &= \{c \in H : \exists v \in c : |L(c,v)| > 10^{-3}\gamma NM\}\end{aligned}$$

For any clause $c \in H_{\text{fail}}$, let $\text{fail}(c) \in c$ denote the variable in v for which $L(c,v)$ is maximal. We bound Alice and Bob's failure probability from below, to get an upper bound on $|H_{\text{fail}}|$.

$$\begin{aligned}
&\Pr(\text{The provers fail}) \geq \\
&\quad \sum_{\substack{c \in H_{\text{fail}}, v \in c \\ (d,w) \in L(c,v)}} \text{FP}(c,d,v,w,k) \Pr(c,d,v,w|k) \geq \\
&\quad \sum_{c \in H_{\text{fail}}} \sum_{(d,w) \in L(c,\text{fail}(c))} \gamma 10^{-4} \Pr(c,d,\text{fail}(c),w:k) \geq \\
&\quad \sum_{c \in H_{\text{fail}}} \gamma 10^{-4} |L(c,\text{fail}(c))| \Pr(c,d,\text{fail}(c),w:k) \geq \\
&\quad \sum_{c \in H_{\text{fail}}} \frac{\gamma^2 NMW_AW_B}{25NM \cdot 10^7 \cdot 222MW_AW_B} = \frac{\gamma^2 |H_{\text{fail}}|}{M \cdot 5.55 \cdot 10^{10}}
\end{aligned}$$

where the last inequality follows from taking a tuple in H. As Pr(The provers fail) $< \frac{\gamma^3}{5.55 \cdot 10^{13}}$, we have

$$|H_{\text{fail}}| < \frac{\gamma^3}{5.55 \cdot 10^{13}} \cdot \frac{M \cdot 5.55 \cdot 10^{10}}{\gamma^2} = 10^{-3} \gamma M$$

Taking $R = H \backslash H_{\text{fail}}$, we get

$$|R| \geq (1 - 0.002\gamma)M - |H_{\text{fail}}| \geq (1 - 0.003\gamma)M$$

as required. □

This enables us to finally prove Theorem 1.1:

Proof of Theorem 1.1. Assume Φ is not satisfiable, and assume by contradiction that the provers had some strategy which would work with success probability $\geq 1 - \frac{\gamma^3}{5.55 \cdot 10^{13}}$. Then there has to be a measurement result k such that the success probability given k is at least $1 - \frac{\gamma^3}{5.55 \cdot 10^{13}}$. However, according to Lemma 4.16, either the provers are caught with probability greater than $\frac{\gamma^3}{5.55 \cdot 10^{13}}$ (which contradicts our assumption on the success probability), or there exists a set R as in the premises of that lemma. However, this would imply that there is a strategy in the classical protocol with success probability $> (1 - 0.003\gamma)(1 - \gamma 10^{-3})(1 - \gamma 10^{-4}) > 1 - \gamma/3$, which is a contradiction. □

5 Conclusions and Open Problems

We have shown that NEXP can be recognized in a quantum MIP protocol, even if the provers have unlimited classical communication between them, but are not allowed to share entanglement. Our protocol achieves perfect completeness and constant soundness. It only sends $O(\log(N))$ qubits, and thus can also be used for NP-complete languages with a polylogarithmic communication. Some interesting questions still remain open:

- What is the correct upper bound on the power of this proof system? Note that if the provers were allowed to make any joint separable measurement it would be exactly NEXP. Does adding provers or communication rounds help? What happens if there is just one quantum round?
- Is there a parallel repetition lemma for protocols when the provers are allowed to communicate with each other? The original proof of [23] does not apply here.
- What happens in the related scenario, when the provers are allowed to share entanglement but are not allowed to communicate? Can similar ideas work here?
- Does our result hold when the provers have a bounded amount of entanglement in addition to their communication channel?

6 Acknowledgments

We thank Scott Aaronson, Dorit Aharonov, Richard Cleve, Irit Dinur, Julia Kempe, Debbie Lueng, Oded Regev, Peter Shor and John Watrous for many helpful discussions.

References

[1] Scott Aaronson, Salman Beigi, Andrew Drucker, Bill Fefferman, and Peter Shor. The power of unentanglement. In *IEEE Conference on Computational Complexity*, pages 223–236. IEEE Computer Society, 2008.

[2] Babai, Fortnow, and Lund. Non-deterministic exponential time has two-prover interactive protocols (addendum). *CMPCMPL: Computational Complexity*, 2, 1992.

[3] H. Barnum, M. A. Nielsen, and B. W. Schumacher. Information transmission through a noisy quantum channel. *Phys. Rev. A*, 57(7):4153–4175, June 1998.

[4] Bennett, DiVincenzo, Fuchs, Mor, Rains, Shor, Smolin, and Wootters. Quantum nonlocality without entanglement, 1998.

[5] Hugue Blier and Alain Tapp. All languages in NP have very short quantum proofs, September 05 2007.

[6] Harry Buhrman, Richard Cleve, John Watrous, and Ronald de Wolf. Quantum fingerprinting, February 01 2001. Comment: 8 pages, LaTeX, one figure.

[7] Chor, Kushilevitz, Goldreich, and Sudan. Private information retrieval. *JACM: Journal of the ACM*, 45, 1998.

[8] Cleve, Hoyer, Toner, and Watrous. Consequences and limits of nonlocal strategies. In *Annual IEEE Conference on Computational Complexity (formerly Annual Conference on Structure in Complexity Theory)*, volume 19, 2004.

[9] Richard Cleve, Dmitry Gavinsky, and Rahul Jain. Entanglement-resistant two-prover interactive proof systems and non-adaptive private information retrieval systems, July 11 2007. Comment: 8 pages.

[10] Richard Cleve, William Slofstra, Falk Unger, and Sarvagya Upadhyay. Strong parallel repetition theorem for quantum XOR proof systems, April 11 2006. Comment: 17 pages, no figures.

[11] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Quantum private queries, October 11 2007. Comment: 4 pages, 1 figure.

[12] Tsuyoshi Ito, Hirotada Kobayashi, Daniel Preda, Xiaoming Sun, and Andrew Chi-Chih Yao. Generalized tsirelson inequalities, commuting-operator provers, and multi-prover interactive proof systems. In *IEEE Conference on Computational Complexity*, pages 187–198. IEEE Computer Society, 2008.

[13] Julia Kempe, Hirotada Kobayashi, Keiji Matsumoto, Ben Toner, and Thomas Vidick. Entangled games are hard to approximate. In *FOCS: Foundations of Computer Science*, October 2008.

[14] Kerenidis and de Wolf. Exponential lower bound for 2-query locally decodable codes via a quantum argument. *JCSS: Journal of Computer and System Sciences*, 69, 2004.

[15] Kerenidis and de Wolf. Quantum symmetrically-private information retrieval. *IPL: Information Processing Letters*, 90, 2004.

[16] Kitaev and Watrous. Parallelization, amplification, and exponential time simulation of quantum interactive proof systems. In *STOC: ACM Symposium on Theory of Computing (STOC)*, 2000.

[17] Kobayashi and Matsumoto. Quantum multi-prover interactive proof systems with limited prior entanglement. *JCSS: Journal of Computer and System Sciences*, 66, 2003.

[18] Hirotada Kobayashi, Keiji Matsumoto, and Tomoyuki Yamakami. Quantum certificate verification: Single versus multiple quantum certificates. *CoRR*, quant-ph/0110006, 2001. informal publication.

[19] Y. K. Liu, M. Christandl, and F. Verstraete. N-representability is QMA-complete, September 17 2006.

[20] Michael Nielsen and Isaac Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 2000.

[21] Sandu Popescu and Daniel Rohrlich. Causality and nonlocality as axioms for quantum mechanics. Technical Report HPL-BRIMS-97-23, Hewlett Packard Laboratories, October 29 1997.

[22] Preda. Manuscript, 2008.

[23] Raz. A parallel repetition theorem. *SICOMP: SIAM Journal on Computing*, 27, 1998.

[24] A. Shamir. IP = PSPACE. *J. ACM*, 39:869–877, 1992.

[25] Stephanie Wehner. Entanglement in interactive proof systems with binary answers. *CoRR*, abs/quant-ph/0508201, 2005. informal publication.

A Calculating Probabilities

Let $r = (c, d, v, w)$. We wish to estimate $\Pr(r|k)$. Bayes' rule gives

$$\Pr(r|k) = \frac{\Pr(k|r)\Pr(r)}{\Pr(k)} = \frac{\Pr(k|r)\Pr(r)}{\sum_s Pr(k|s)\Pr(s)}$$

where s denotes any legal tuple $s = (\tilde{c}, \tilde{d}, \tilde{v}, \tilde{w})$ with $\tilde{c}, \tilde{d} \in C$, $\tilde{v}, \tilde{w} \in V$ and $\tilde{v} \in \tilde{c}$. As the prior distribution for all legal tuples is identical, we are only interested in calculating $\Pr(k|s)$ for any legal tuple $s = (\tilde{c}, \tilde{d}, \tilde{v}, \tilde{w})$.

In the protocol we presented, the provers first apply their measurement and get k, and then the verifier measures to get s. However, it is physically equivalent to assume the verifier measured first. As the states sent to the provers are unentangled after tracing out the verifier, we have that

$$\Pr(k|s) = \text{tr}((I \otimes A_k)\rho_A(I \otimes A_k)^\dagger) \cdot \text{tr}((I \otimes B_k)\rho_B(I \otimes B_k)^\dagger)$$

where ρ_A is the state in $H_A^v \otimes H_A^M$, ρ_B is the state in $H_B^v \otimes H_B^M$, and the identity is applied on the verifier's side.

When considering states in $H_A^v \otimes H_A^m \otimes H_A^p$, we stick to the convention that the first m qubits define the verifier's private space, then next $m + 3$ describe the message qubits, and the last t define Alice's private space. We can now calculate

$$\begin{aligned} A_k(c) &= \text{tr}(A_k(|c\rangle\langle c| \otimes I)A_k) \\ &= \sum_{j=1}^{8Mt} \sum_{h=8t(c-1)+1}^{8tc} A_k[j,h]\overline{A_k[j,h]} \\ &= \sum_{j=1}^{8Mt} \sum_{h=8t(c-1)+1}^{8tc} |A_k[j,h]|^2 \end{aligned}$$

where we abuse notation and treat c as an index. Similarly,

$$\begin{aligned} B_k(v) &= \text{tr}(B_k(|v\rangle\langle v| \otimes I)B_k) \\ &= \sum_{j=1}^{2Nt} \sum_{h=2t(v-1)+1}^{2tv} B_k[j,h]\overline{B_k[j,h]} \\ &= \sum_{j=1}^{2Nt} \sum_{h=2t(v-1)+1}^{2tv} |B_k[j,h]|^2 \end{aligned}$$

where again v was treated as an index. We now assume that the v, w is being traced out, and only look at the probabilities for c, d, generated from $\text{tr}((I \otimes A_k)\rho_A(I \otimes A_k)^\dagger)$. As $A_k(c)$ is just the trace out of the private data of the prover and the qubits which fit the assignment, then $A_k(c) = \text{tr}((I \otimes A_k)\rho_A(I \otimes A_k)^\dagger)$. We are analyzing the following expression:

$$\text{tr}((I_M \otimes A_{8Mt})\rho_A(I_M \otimes A_{8Mt})^\dagger)$$

Up to normalization, ρ_A is a matrix which contains exactly four 1s, arranged: (a, a), (a, b), (b, a), (b, b). However, as we shall soon see, either $a = b$ (in which case we have a single cell with a 4 in it) or else $|a - b| \geq 8Mt$ and thus, by the previous paragraph, we can ignore the off-diagonal entries. To conclude, in both cases we can restrict our attention to the diagonal entries.

Thus the structure of the ρ_A matrix is:

$$\rho = \frac{1}{\sqrt{2}}(|cc\rangle + |dd\rangle) \otimes |000\rangle\langle 000|(\langle cc| + \langle dd|)\frac{1}{\sqrt{2}} \otimes |0_t\rangle\langle 0_t| \in H_v^A \otimes H_M^A \otimes H_p^A$$

Note that the term 0_t refers to element in a space of dimension t, as opposed to 000, an element in a space of dimension 2^3. If $c = d$ then obviously there is only one nonzero cell in the final matrix, on the diagonal. Otherwise, since $|cc\rangle$ is located in the cell $Mc + c = (M + 1)c$, and $d \neq c$, they are differentiated (after tensoring) by at least $(M + 1) \cdot 8 \cdot t > 8Mt$, as required. Let

$$A_k(i) = \sum_{j=1}^{8Mt} \sum_{h=8t(i-1)+1}^{8ti} A_k[j, h]\overline{A_k[j, h]} = \sum_{j=1}^{8Mt} \sum_{h=8t(i-1)+1}^{8ti} |A_k[j, h]|^2$$

The probability that the verifier measures c, d in the modified protocol given k is

$$\begin{aligned}
P(c, d|k) &= \frac{P(k|c, d)P(c, d)}{P(k)} \\
&= \frac{P(k|c, d)P(c, d)}{\sum_{\tilde{c},\tilde{d}} P(k|\tilde{c}, \tilde{d})P(\tilde{c}, \tilde{d})} \\
&= \frac{\text{tr}(A_k \rho_{c,d} A_k^\dagger)}{\sum_{\tilde{c},\tilde{d}} \text{tr}(A_k \rho_{\tilde{c},\tilde{d}} A_k^\dagger)} \\
\text{(equal unless } c = d\text{)} \quad &\geq \frac{A_k(c) + A_k(d)}{\sum_{\tilde{c}\neq\tilde{d}}(A_k(\tilde{c}) + A_k(\tilde{d})) + \sum_{\tilde{c}} 4A_k(\tilde{c})} \\
&= \frac{A_k(c) + A_k(d)}{\sum_{\tilde{c},\tilde{d}}(A_k(\tilde{c}) + A_k(\tilde{d})) + \sum_{\tilde{c}} 2A_k(\tilde{c})} \\
&= \frac{A_k(c) + A_k(d)}{2MW_{A_k} + 2W_{A_k}}
\end{aligned}$$

where W_{A_k} is the total weight: $W_{A_k} = \sum_{\tilde{c}} A_k(\tilde{c})$. Note that if $c = d$ we use $4A_k(c)$ instead of $A_k(c) + A_k(d)$.

A.1 Bounding the Denominator

Let

$$\begin{aligned} W_{A_k} &= \sum_i A_k(i) \\ W_{B_k} &= \sum_i B_k(i) \\ \tilde{W}_k &= \Sigma_{\tilde{c}\in C,\tilde{v}\in\tilde{c}} A_k(\tilde{c})B_k(\tilde{v}) \end{aligned}$$

We want to bound the denominator in

$$\Pr(c,d,v,w|k) = \frac{(A_k(c)+A_k(d))(B_k(v)+B_k(w))}{\sum_{\tilde{c},\tilde{d}\in C,\tilde{v}\in c,\tilde{w}\in V}\Pr(\tilde{c},\tilde{d},\tilde{v},\tilde{w}|k)}$$

Note that if $\tilde{c}=\tilde{d}$, then $\mathrm{tr}((I\otimes A_k)\rho_A(I\otimes A_k)^\dagger) = 4A_k(\tilde{c})$. However, when $\tilde{c}\neq\tilde{d}$, we account this twice (because any of them can be considered first in the sum). Thus, the denominator becomes

$$\sum_{\tilde{c},\tilde{d}}\sum_{\tilde{v}\in\tilde{c},\tilde{w}}(A_k(\tilde{c})+A_k(\tilde{d}))(B_k(\tilde{v})+B_k(\tilde{w}))+2\left(\sum_{\tilde{c}=\tilde{d},\tilde{v},\tilde{w}}+\sum_{\tilde{c},\tilde{d},\tilde{v}=\tilde{w}}\right)+4\sum_{\tilde{c}=\tilde{d},\tilde{v}=\tilde{w}} \quad (7)$$

where all the sums are on $(A_k(\tilde{c})+A_k(\tilde{d}))(B_k(\tilde{v})+B_k(\tilde{w}))$, and factors of two and four come from $\tilde{c}=\tilde{d}$, and $\tilde{v}=\tilde{w}$. We begin by bounding the first two sums (which will contribute most of the weight).

$$\sum_{\tilde{c},\tilde{d}}\sum_{\tilde{v}\in\tilde{c},\tilde{w}}(A_k(\tilde{c})+A_k(\tilde{d}))(B_k(\tilde{v})+B_k(\tilde{w})) = \sum_{\tilde{c},\tilde{d}}\sum_{\tilde{v}\in\tilde{c},\tilde{w}} A_k(\tilde{c})B_k(\tilde{v})+A_k(\tilde{c})B_k(\tilde{w})+A_k(\tilde{d})B_k(\tilde{v})+A_k(\tilde{d})B_k(\tilde{w})$$

We now look at each of the four terms separately:

$$\begin{aligned} \sum_{\tilde{c},\tilde{d}}\sum_{\tilde{v}\in\tilde{c},\tilde{w}} A_k(\tilde{c})B_k(\tilde{v}) &= \sum_{\tilde{d},\tilde{w}}\sum_{\tilde{c},\tilde{v}\in\tilde{c}} A_k(\tilde{c})B_k(\tilde{v}) = NM\sum_{\tilde{c},\tilde{v}\in\tilde{c}} A_k(\tilde{c})B_k(\tilde{v}) = NM\tilde{W}_k \\ \sum_{\tilde{c},\tilde{d}}\sum_{\tilde{v}\in\tilde{c},\tilde{w}} A_k(\tilde{c})B_k(\tilde{w}) &= 3M\sum_{\tilde{c},\tilde{w}\in V} A_k(\tilde{c})B_k(\tilde{w}) = 3MW_{A_k}W_{B_k} \\ \sum_{\tilde{c},\tilde{d}}\sum_{\tilde{v}\in\tilde{c},\tilde{w}} A_k(\tilde{d})B_k(\tilde{v}) &= 5NW_{A_k}W_{B_k} < 5MW_{A_k}W_{B_k} \\ \sum_{\tilde{c},\tilde{d}}\sum_{\tilde{v}\in\tilde{c},\tilde{w}} A_k(\tilde{d})B_k(\tilde{w}) &= 3MW_{A_k}W_{B_k} \end{aligned}$$

We used the fact that Φ is 3-SAT, and that each variable appears exactly five times.

We return to bounding the sums in (7). By fixing $\tilde{c}$, we get that if $\tilde{c}=\tilde{d}$ the second sum is bounded, relative to the first, by a factor of $2/M$. Fixing $\tilde{v}=\tilde{w}$, we can bound the third sum by a factor of $2/N$. Fixing both, the fourth sum is bounded by a factor of $4/(NM)$. We get an overall bound for the denominator of:

$$(NM\tilde{W}_k+3MW_{A_k}W_{B_k}+5NW_{A_k}W_{B_k}+3MW_{A_k}W_{B_k})(1+2/M+2/N+4/(NM))$$

Since M and N are arbitrarily large, and $M \geq N$, we deduce our bound:

$$2(NM\tilde{W}_k + 11MW_{A_k}W_{B_k})$$

which finally gives

$$\Pr(c, d, v, w|k) \geq \frac{A_k(c)B_k(v) + A_k(c)B_k(w) + A_k(d)B_k(v) + A_k(d)B_k(w)}{2NM\tilde{W}_k + 22MW_{A_k}W_{B_k}}$$