

# The quantum communication complexity of the pointer chasing problem: the bit version

Rahul Jain\*

Jaikumar Radhakrishnan\*

Pranab Sen†

## Abstract

We consider the two-party quantum communication complexity of the bit version of the pointer chasing problem, originally studied by Klauck, Nayak, Ta-Shma and Zuckerman [KNTZ01]. We show that in any quantum protocol for this problem, the two players must exchange  $\Omega(\frac{n}{k^4})$  qubits. This improves the previous best bound of  $\Omega(\frac{n}{2^{2^{O(k)}}})$  in [KNTZ01], and comes significantly closer to the best upper bounds known  $O(n+k \log n)$  (classical deterministic [PRV01]) and  $O(k \log n + \frac{n}{k}(\log^{\lceil k/2 \rceil}(n) + \log k))$  (classical randomized [KNTZ01]). Our proof uses a round elimination argument with correlated input generation, making better use of the information theoretic tools than in previous papers.

## 1 Introduction

We consider the following pointer chasing problem in the two-party communication model [Yao79, Yao93].

Let  $V_A$  and  $V_B$  be disjoint sets of size  $n$ . Alice is given a function  $F_A : V_A \rightarrow V_B$  and player Bob is given a function  $F_B : V_B \rightarrow V_A$ . Let  $F \triangleq F_A \cup F_B$ . There is a fixed vertex  $s$  in  $V_B$ . The players need to exchange messages and determine the least significant bit of  $F^{(k+1)}(s)$ , where  $k$  and  $s$  are known to both parties in advance.

If Bob starts the communication, there is a straightforward classical deterministic protocol where one of the players can determine the answer after  $k$  messages of  $\log n$  bits have been exchanged. It appears much harder, however, to solve the problem efficiently with  $k$  messages, when Alice is required to send the first message. We refer to this as the pointer chasing problem  $P_k$ .

**Background:** The pointer chasing problem has been studied in the past to show rounds versus communication tradeoffs in classical communication complexity. Nisan and Wigderson [NW93] showed (following some earlier results of Papadimitriou and Sipser [PS84], and Duris, Galil and Schnitger [DGS87]) that the players must exchange  $\Omega(\frac{n}{k} - k \log n)$  bits to solve  $P_k$ ; their bound was improved by Klauck [Kla00] to  $\Omega(\frac{n}{k} + k)$ . These lower bounds hold even if randomization is allowed. A deterministic protocol with  $O(n + k \log n)$  bits of communication was given by Ponzio, Radhakrishnan and Venkatesh [PRV01], and a classical randomized protocol with  $O(k \log n + \frac{n}{k}(\log^{\lceil k/2 \rceil}(n) + \log k))$  bits by Klauck, Nayak, Ta-Shma and Zuckerman [KNTZ01]. Thus, the lower and upper bounds are quite close in the classical setting.

In the quantum communication complexity model, this problem has been studied recently by Klauck, Nayak, Ta-Shma and Zuckerman [KNTZ01], who, using interesting information-theoretic techniques, showed

---

\*STCS, Tata Institute of Fundamental Research, Mumbai 400005, India. Email: {rahulj, jaikumar}@tcs.tifr.res.in.

†LRI, UMR 8623, Université de Paris–Sud, 91405 Orsay, France. Email: pranab@lri.fr.

a lower bound of  $\Omega(\frac{n}{2^{2^{O(k)}}})$ . This bound deteriorates rapidly with  $k$ , and becomes trivial for  $k \geq \log \log n$ . We improve this lower bound.

**Result:** In any bounded error quantum protocol for the pointer chasing problem  $P_k$ , Alice and Bob must exchange  $\Omega(\frac{n}{k^4})$  qubits.

**Our proof technique:** The underlying information theoretic tools we use are, in fact, mainly taken from the paper Klauck, Nayak, Ta-Shma and Zuckerman [KNTZ01]. Our proofs use the round elimination method, stated explicitly in the classical communication complexity setting by Miltersen, Nisan, Safra and Wigderson [MNSW98]. This technique was applied in the quantum setting by Klauck *et al.*, who developed several tools, notably the *average encoding theorem* and the *local transition theorem*. Their argument was refined further by Sen and Venkatesh [SV98]. Recently, Jain, Radhakrishnan and Sen [JRS02] showed an optimal  $\Omega(n \log^{(k)} n)$  lower bound for the *full version* of the pointer chasing problem, where the players must determine the full description of  $F^{(k+1)}(s)$ , and not just its least significant bit. This was also obtained by the round elimination argument. In this paper, we adapt this argument to the *bit version* of the problem. For this, we consider a slightly different pointer chasing problem, where the two players are allowed to generate their own inputs and then proceed to compute the answer. To keep this problem non-trivial we must impose some restrictions on the way the players behave. First, we insist that the inputs they generate must be sufficiently rich. Second, the amount of communication before the input is generated, is limited. In previous round elimination arguments, the inputs were supplied to the two players from ‘outside’. While this worked well for many problems, for the pointer chasing problem it made things difficult. However, letting the players generate their inputs gives rise to new technical difficulties, because the inputs they generate are not exactly what we want, but only close to it. So, we need to apply a correction step, that converts a protocol whose inputs have a distribution close to the one we desire into one where the inputs are exactly what we want. Overall, we believe, the main contribution of this work is in showing how existing information theoretic tools can be better exploited for round elimination in quantum communication protocols.

## 1.1 Organization of the rest of the paper

In the next section, we define the pointer chasing problem formally and derive our main result assuming a Round Elimination Lemma. In Section 3, we collect the probabilistic and information theoretic tools that are required for the proof. Finally, in Section 4, we describe the round elimination argument in detail.

## 2 Lower bound for the pointer chasing problem

In this section, we formally define the problem and our main result assuming a Round Elimination Lemma, which will be proved in later section.

**Quantum communication protocols:** We consider two party quantum communication protocols as defined by Yao [Yao93]. Let  $E, F, G$  be arbitrary finite sets and  $f : E \times F \rightarrow G$  be a function. There are two players Alice and Bob, who hold qubits. When the communication game starts, Alice holds  $|x\rangle$  where  $x \in E$  together with some ancilla qubits in the state  $|0\rangle$ , and Bob holds  $|y\rangle$  where  $y \in F$  together with some ancilla qubits in the state  $|0\rangle$ . Thus the qubits of Alice and Bob are initially in computational basis states, and the initial superposition is simply  $|x\rangle_A |0\rangle_A |y\rangle_B |0\rangle_B$ . Here the subscripts denote the ownership of the qubits by Alice and Bob. The players take turns to communicate to compute  $f(x, y)$ . Suppose it is Alice’s

turn. Alice can make an arbitrary unitary transformation on her qubits and then send one or more qubits to Bob. Sending qubits does not change the overall superposition, but rather changes the ownership of the qubits, allowing Bob to apply his next unitary transformation on his original qubits plus the newly received qubits. At the end of the protocol, the last recipient places the answer in a special register Ans.

**Definition 1 (Safe transformation, protocols)** Let  $\mathcal{H}$  and  $\mathcal{K}$  be a finite-dimensional Hilbert spaces, with bases  $(|h\rangle : h \in H)$  and  $(|k\rangle : k \in K)$ . We say that a unitary transformation  $U$  on  $\mathcal{H} \otimes \mathcal{K}$  acts safely on  $\mathcal{H}$  if there exist unitary transformations  $(U_h : h \in H)$  acting on  $\mathcal{K}$  such that for all  $h \in H$  and  $k \in K$ ,

$$U : |h\rangle|k\rangle \mapsto |h\rangle U_h |k\rangle.$$

We say that a protocol acts safely on a register  $R$ , if all unitary transformations in the protocol act safely on  $R$ , and  $R$  is never sent as part of a message. We say that a protocol is safe if Alice and Bob act safely on their input registers.

When the inputs are classical, we can always assume that the protocol is safe. This is possible since the inputs to Alice and Bob are in computational basis states. So, the players can make a secure copy of their inputs before beginning the protocol.

## 2.1 The pointer chasing problem $P_k$

**The input:** Alice's input is a function  $F_A : V_A \rightarrow V_B$ . Bob's input is a function  $F_B : V_B \rightarrow V_A$ .  $V_A$  and  $V_B$  are disjoint sets of size  $n$  each. We assume that  $n = 2^r$  for some  $r \geq 1$ .

**The golden path:** There is a fixed vertex  $s \in V_B$ . Let  $F \triangleq F_A \cup F_B$ ; let  $\text{ans} \triangleq \text{lsb}(F^{(k+1)}(s))$ . Here  $\text{lsb}(x)$  is the least significant bit of  $x$ ; we assume that vertices in  $V_A$  and  $V_B$  have binary encodings of length  $\log n$ .

**The communication:** Alice and Bob exchange messages  $M_1, \dots, M_k$ , having lengths  $c_1 n, \dots, c_k n$ , via a safe quantum protocol in order to determine ans. Alice starts the communication, that is, she sends  $M_1$ . The player receiving  $M_k$  places a guess for ans in the register Ans. We require that the bit obtained by measuring Ans in the computational basis<sup>1</sup> should be the correct answer (i.e. equal to  $\text{lsb}(F^{(k+1)}(s))$  with probability at least  $\frac{3}{4}$ , for all  $F_A, F_B$ .

## 2.2 The predicate $Q_k^A$

We will show our lower bound for  $P_k$  using an inductive argument. It will be convenient to state our induction hypothesis by means of a predicates  $Q_k^A$  and  $Q_k^B$ , defined below. Roughly, the induction proceeds as follows. We show that if there is an efficient protocol for  $P_k$ , then  $Q_k^A$  is true. We then show independently that  $Q_\ell^A$  implies  $Q_{\ell-1}^B$  and  $Q_\ell^B$  implies  $Q_{\ell-1}^A$ , and that  $Q_0^A$  and  $Q_0^B$  are false. Thus, there is no efficient protocol for  $P_k$ .

We now define  $Q_k^A(c_1, \dots, c_k, n_a, n_b, \epsilon)$  for  $k \geq 1$ . Then, separately, we define  $Q_0^A$ . For  $k \geq 0$ ,  $Q_k^B$  is the same as  $Q_k^A$ , with the roles of Alice and Bob reversed. Consequently, all our statements involving  $Q_k^A$  and  $Q_k^B$  have two forms, where one is obtained from the other by reversing the roles of Alice and Bob. We will typically state just one of them, and let the reader infer the other.

The predicate  $Q_k^A(c_1, \dots, c_k, n_a, n_b, \epsilon)$  holds if there is a quantum protocol of the following form.

<sup>1</sup>From now on, all measurements are to be performed using the computational basis.

**Input generation:** Alice and Bob ‘generate’ most of their inputs themselves. Alice has  $n$  input registers ( $F_A[u] : u \in V_A$ ) and Bob has  $n$  input registers ( $F_B[v] : v \in V_B$ ). There is a fixed vertex  $s \in V_B$ , that is known to both players. Each of Alice’s registers has  $\log n$  qubits so that it can hold a description of a vertex in  $V_B$ ; similarly, each of Bob’s registers can hold a description of a vertex in  $V_A$ . In addition, Alice and Bob have registers for their ‘work’ qubits  $W_A$  and  $W_B$ .

When the protocol starts, Alice’s registers are all initialized to 0. On Bob’s side, the register  $F_B[s]$  starts off with the uniform superposition  $|\mu\rangle \triangleq \frac{1}{\sqrt{n}} \sum_{a \in V_A} |a\rangle$ ; the other registers are all 0.

Alice starts by generating a pure state in  $\widetilde{M}_1 M_1$ , where  $\widetilde{M}_1, M_1$  are each  $c_1 n$  qubit registers. Then she applies a unitary transformation  $U_A$  on her registers other than  $M_1$  to generate a state in registers  $F_A$  and  $W_A$ . Alice then sends  $M_1$  to Bob.

Now, Bob generates his input using the message  $M_1$  as follows. He applies a unitary transformation  $U_B$  on the registers that he owns at this point:

- $M_1$ , the message registers just received from Alice;
- $F_B[s]$  the register holding the start pointer, which is in the state  $|\mu\rangle$  in tensor with the other register;
- ( $F_B[b] : b \in V_B - \{s\}$ ) and the registers  $W_B$  holding the work qubits of  $B$ , which contain 0.

$U_B$  must operate “safely” on  $F_B[s]$ .  $F_B$  holds the ‘generated input’ to Bob for the pointer chasing problem, and  $W_B$  Bob’s ‘work qubits’.

We will use  $F_A, F_B$  also to refer to the actual states of the respective registers;  $f_A, f_B$  will denote the states that would result, were we to measure  $F_A, F_B$ . Thus, typically  $F_A, F_B$  will be parts of a pure state (the global state of Alice’s and Bob’s qubits) whereas  $f_A, f_B$  will be mixtures of computational basis states.

For our predicate  $Q_k^A(c_1, \dots, c_k, n_a, n_b, \epsilon)$  to hold, this input generation process must satisfy some conditions.

**Requirement 1(a):** There is a subset  $X_A \subseteq V_A$  of size at most  $n_a$  such that the variables ( $f_A(u) : u \in V_A$ ) are independent, and for  $u \in V_A - X_A$ ,  $f_A(u)$  is uniformly distributed.

**Requirement 1(b):** There is a subset  $X_B \subseteq V_B - \{s\}$  of size at most  $n_b$  such that the random variables ( $f_B(v) : v \in V_B$ ) are independent, and  $f_B(v)$  for  $v \in V_B - X_B$  is uniformly distributed. Note that  $f_B[s]$  is automatically uniformly distributed, because initially  $F_B[s]$  contains the uniform superposition, and  $U_B$  acts safely on  $F_B[s]$ .

**Communication:** After  $U_A, U_B$  have been applied, Alice and Bob follow a quantum protocol exchanging further messages  $M_2, \dots, M_k$  of lengths  $c_2 n, \dots, c_k n$ . Bob sends the message  $M_2$ . The rest of the protocol is required to act safely on registers  $F_A, F_B$ . At the end of the protocol, the player who receives  $M_k$  places a qubit in a special register Ans. The protocol then terminates.

**The probability of error:** Once the protocol has terminated, all registers are measured. Let ans denote the value observed in Ans, and let  $f_A$  and  $f_B$  be the values observed in  $F_A$  and  $F_B$ ; we treat  $f_A$  and  $f_B$  as functions (from  $V_A$  to  $V_B$  and  $V_B$  to  $V_A$  respectively). Let  $f \triangleq f_A \cup f_B$ . Note that ans and  $f$  are random variables.

**Requirement 2:**  $\Pr[\text{ans} = \text{lsb}(f^{(k+1)}(s))] \geq 1 - \epsilon$ .

**Base case:** In  $Q_0^A(\epsilon)$ , there is no input generation phase or communication. Bob and Alice start as before, with  $|\mu\rangle$  in Bob's register  $F_B[s]$ . Alice produces a guess ans for  $\text{lsb}(f_B(s))$ , which must be correct with probability at least  $1 - \epsilon$ . Clearly, we have the following base case for our induction.

**Proposition 1** *If  $Q_0^A(\epsilon)$  is true then  $\epsilon \geq \frac{1}{2}$ .*

Our goal is to show that if  $Q_k^A$  holds, then  $c_1 + c_2 + \dots + c_k = \Omega(k^{-4})$ . By the following lemma, this implies a lower bound  $\frac{n}{k^4}$  for  $P_k^A$ .

**Lemma 1** *If there is a safe quantum protocol for  $P_k^A$  with  $v_0 = s \in V_B$ , messages of lengths  $c_1n, \dots, c_kn$ , and worst case error at most  $\frac{1}{4}$ , then  $Q_k^A(c_1, \dots, c_k, n_A = 0, n_B = 0, \frac{1}{4})$  is true.*

**Proof:** We are given a safe quantum protocol  $\mathcal{P}$  for  $P_k$ , where Alice sends the first message  $M_1$ . Consider the operation of  $\mathcal{P}$  when uniform superpositions are fed for  $F_A$  and  $F_B$ . Consider the state of Alice just before  $M_1$  is sent to  $B$ . This state has two parts.

1. The qubits that Alice keeps with herself,  $F_A W_A$ , where  $F_A$  is  $n \log n$  qubits long.
2. The  $c_1n$  qubits that constitute the message  $M_1$ .

Let  $\widetilde{M}_1 M_1$  contain a canonical purification of  $M_1$ , where  $\widetilde{M}_1$  is  $c_1n$  qubits long. Clearly, it is within Alice's powers to first generate the canonical purification in  $\widetilde{M}_1 M_1$ , and then apply a unitary transformation  $U_A$  on  $\widetilde{M}_1$  plus some initially zero ancilla qubits in order to generate the correct state of  $F_A W_A M_1$ . Alice then sends  $M_1$

In our protocol, on Bob's side,  $F_B[s]$  already has a uniform superposition in tensor with the rest of Alice's and Bob's qubits. Then, Bob generates the rest of his "input",  $F_B[v], v \neq s$  as a uniform superposition in tensor with everything else. The registers  $W_B$  are set to  $|0\rangle$ . At this point, the state of  $F_A W_A M_1 F_B W_B$  is exactly the same as it would be in  $\mathcal{P}$  after Bob receives the first message. From now on, Alice and Bob operate exactly as in  $\mathcal{P}$ , which is "safe" on  $F_A, F_B$ . The above parameters for  $Q_k^A$  can now be verified easily.  $\square$

The following lemma is the key to our inductive argument.

**Lemma 2 (Round elimination)** (a) *For  $k \geq 2$ , if  $Q_k^A(c_1, \dots, c_k, n_A, n_B, \epsilon)$  holds (with  $n_A < n$ ) then  $Q_{k-1}^B(c_1 + c_2, c_3, \dots, c_k, n_A, n_B + 1, \epsilon')$  holds with  $\epsilon' = \left(\frac{n}{n-n_A}\right) \left[\epsilon + 3((2 \ln 2)c_1)^{\frac{1}{4}}\right]$ .*

(b) *If  $Q_1^A(c_1, n_A, n_B, \epsilon)$  holds (with  $n_A < n$ ), then  $Q_0^A(\epsilon')$  holds, where  $\epsilon'$  is exactly as in part (a).*

The next section is devoted to the proof of this lemma. Now, let us assume this lemma and prove our main lower bound.

**Theorem 1** *Suppose  $k \leq n^{\frac{1}{4}}$  and  $Q_k^A(c_1, \dots, c_k, 0, 0, \frac{1}{4})$  holds. Then  $c_1 + c_2 + \dots + c_k = \Omega(k^{-4})$ .*

**Proof:** (Sketch) By  $k - 1$  applications of Part (a) of Lemma 2 (a) and one application of Part (b), we conclude that either  $Q_0^A(\epsilon')$  or  $Q_0^B(\epsilon')$  holds with  $\epsilon' \leq \left(\frac{n}{n-k}\right)^k \left[\frac{1}{4} + 3k((2 \ln 2)(c_1 + c_2 + \dots + c_k))^{\frac{1}{4}}\right]$ . Our theorem follows immediately from this and Proposition 1.  $\square$

Now, by using Lemma 1, we can derive from this our lower bound for  $P_k$ .

**Corollary 1 (Main result)** *In any protocol for  $P_k$ , Alice and Bob must exchange a total of  $\Omega(\frac{n}{k^4})$  qubits.*

### 3 Preliminaries

We now recall some basic definitions and facts from probability and information theory, which will be useful in proving our main result. For excellent introductions to classical and quantum information theory, see the books by Cover and Thomas [CT91] and Nielsen and Chuang [NC00] respectively.

If  $A$  is a quantum system with density matrix  $\rho$ , then  $S(A) \triangleq S(\rho) \triangleq -\text{Tr } \rho \log \rho$  is the *von Neumann entropy* of  $A$ . If  $A, B$  are two disjoint quantum systems, the *mutual information* of  $A$  and  $B$  is defined as  $I(A : B) \triangleq S(A) + S(B) - S(AB)$ .

**Fact 1 (see [KNTZ01])** *Let  $X_1, \dots, X_n$  be classical random variables and let  $M$  be a quantum encoding of  $X \triangleq X_1 \dots X_n$ . Then,  $I(X : M) \geq \sum_{i=1}^n I(X_i : M)$ . Also, if  $M$  is  $n$  qubits long, then  $I(X : M) \leq n$ .*

We will be working with various measures of distance between classical and quantum states. For distributions  $D$  and  $D'$  on a finite set  $X$ , their total variational distance is given by  $\|D - D'\|_1 = \sum_{x \in X} |D(x) - D'(x)|$ . We will use the following elementary fact, which we state without proof.

**Fact 2** *Suppose  $D, D'$  are two probability distributions on the same finite set  $X$ , whose total variation distance is  $\|D - D'\|_1 = \delta$ . Then, there exists a stochastic matrix  $P = (p_{xx'})_{xx' \in X}$ , such that  $D = PD'$  and  $\sum_{x' \in X} P(x', x')D(x') = 1 - \frac{1}{2}\delta$ . Let  $\mathcal{H}$  be a Hilbert space with basis  $(|x\rangle : x \in X)$ . Let  $C$  be a unitary transformation on  $\mathcal{H} \otimes \mathcal{H}$  that maps basis vectors of the form  $|x'\rangle|\mathbf{0}\rangle$  (where  $\mathbf{0}$  is a special element of  $X$ ) according to the rule*

$$|x'\rangle|\mathbf{0}\rangle \rightarrow |x'\rangle \otimes \sum_{x \in X} \sqrt{p_{xx'}}|x\rangle,$$

*and maps other standard basis vectors suitably. Suppose  $R'$  and  $R$  are registers that can hold states in  $\mathcal{H}$ , where  $R'$  contains a mixture of basis states with distribution  $D'$  and  $R$  is in the state  $|\mathbf{0}\rangle$ . Apply  $C$  to  $(R', R)$ , and then measure the registers in the computational basis. Let the resulting random variables (taking values in  $X$ ) be  $Z'$  and  $Z$ . Then,  $Z'$  has distribution  $D'$ ,  $Z$  has distribution  $D$  and  $\Pr[Z \neq Z'] \leq \frac{1}{2}\delta$ . Note, that  $C$  acts safely on  $R'$ .*

The trace norm of a linear operator  $A$  is defined as  $\|A\|_t \triangleq \text{Tr } \sqrt{A^\dagger A}$ . The following fundamental theorem (see [AKN01]) shows that the trace distance between two density matrices  $\rho_1, \rho_2$ ,  $\|\rho_1 - \rho_2\|_t$ , bounds how well one can distinguish between  $\rho_1, \rho_2$  by a measurement.

**Theorem 2 ([AKN01])** *Let  $\rho_1, \rho_2$  be two density matrices on the same Hilbert space. Let  $\mathcal{M}$  be a general measurement (i.e. a POVM), and  $\mathcal{M}\rho_i$  denote the probability distributions on the (classical) outcomes of  $\mathcal{M}$  got by performing measurement  $\mathcal{M}$  on  $\rho_i$ . Let the  $\ell_1$  distance between  $\mathcal{M}\rho_1$  and  $\mathcal{M}\rho_2$  be denoted by  $\|\mathcal{M}\rho_1 - \mathcal{M}\rho_2\|_1$ . Then*

$$\|\mathcal{M}\rho_1 - \mathcal{M}\rho_2\|_1 \leq \|\rho_1 - \rho_2\|_t$$

We will need the following ‘‘average encoding theorem’’ of Klauck *et al.* [KNTZ01]. Intuitively speaking, it says that if the mutual information between a classical random variable and its quantum encoding is small, then the various quantum ‘‘codewords’’ are close to the ‘‘average codeword’’.

**Theorem 3 (Average encoding theorem [KNTZ01])** *Suppose  $X, Q$  are two disjoint quantum systems, where  $X$  is a classical random variable, which takes value  $x$  with probability  $p_x$ , and  $Q$  is a quantum encoding  $x \mapsto \sigma_x$  of  $X$ . Let the density matrix of the average encoding be  $\sigma \triangleq \sum_x p_x \sigma_x$ . Then*

$$\sum_x p_x \|\sigma_x - \sigma\|_t \leq \sqrt{(2 \ln 2)I(X : Q)}$$

We will also need the following “local transition theorem” of Klauck *et al.* [KNTZ01].

**Theorem 4 (Local transition, [KNTZ01])** *Let  $\rho_1, \rho_2$  be two mixed states with support in a Hilbert space  $\mathcal{H}$ ,  $\mathcal{K}$  any Hilbert space of dimension at least the dimension of  $\mathcal{H}$ , and  $|\phi_i\rangle$  any purifications of  $\rho_i$  in  $\mathcal{H} \otimes \mathcal{K}$ . Then, there is a local unitary transformation  $U$  on  $\mathcal{K}$  that maps  $|\phi_2\rangle$  to  $|\phi'_2\rangle \triangleq (I \otimes U)|\phi_2\rangle$  ( $I$  is the identity operator on  $\mathcal{H}$ ) such that*

$$\| |\phi_1\rangle\langle\phi_1| - |\phi'_2\rangle\langle\phi'_2| \|_t \leq 2\sqrt{\|\rho_1 - \rho_2\|_t}$$

## 4 Round elimination: proof of Lemma 2

We consider Part (a) first. Part (b) follows using similar argument, and we do not describe them explicitly. Suppose  $Q_k^A(c_1, c_2, \dots, c_k, n_A, n_B, \epsilon)$  is true. That is, there is a protocol  $\mathcal{P}$  satisfying Requirements 1 and 2 in the definition of  $Q_k^A$ . We need to show that there is a protocol that satisfies the requirements for  $Q_{k-1}^B$  with parameters stated in Lemma 2 (a).

In what follows, subscripts of pure and mixed states will denote the registers which are in those states. For  $u \in V_A$ , we use the subscript  $u$  instead of  $F_A[u]$ . Similarly, for  $v \in V_B$ , we use the subscript  $v$  instead of  $F_B[v]$ . For example, we say that the register  $F_B[s]$  is initially in the state  $|\mu\rangle_s = \frac{1}{\sqrt{n}} \sum_{u \in V_A} |u\rangle_s$ .

Let  $|\psi^A\rangle$  be the (pure) state of Alice’s registers just before she sends  $M_1$  to Bob. At this point the state of all the registers taken together is the pure state

$$|\psi_{\text{in}}\rangle = |\psi^A\rangle \otimes \frac{1}{\sqrt{n}} \sum_{a \in V_A} |a\rangle_s |\mathbf{0}\rangle_R, \quad (1)$$

where  $R$  is the set of registers corresponding to the rest of  $B$ ’s input ( $F_B[v] : v \in V_B - \{s\}$ ), and work qubits  $W_B$ . For  $a \in V_A$ , we may expand  $|\psi^A\rangle$  as

$$|\psi^A\rangle = \frac{1}{\sqrt{\ell_a}} \sum_{b \in V_B} |b\rangle_a |\psi_{a \rightarrow b}^A\rangle, \quad (2)$$

where  $\ell_a = 1$  if  $a \in X_A$  and  $\ell_a = n$  otherwise. Here,  $|\psi_{a \rightarrow b}^A\rangle$  is a pure state of Alice’s registers ( $F_A(v) : v \in V_A - \{a\}$ ) and  $W_A$ . Note that  $|\psi_{a \rightarrow b}^A\rangle$  is precisely the state of these registers when  $F_A[a]$  is measured and found to be in state  $|b\rangle$ . (If  $\Pr[f_A[a] = b] = 0$ , then  $|\psi_{a \rightarrow b}^A\rangle \triangleq 0$ .) From (1) and (2), we have

$$|\psi_{\text{in}}\rangle = \frac{1}{\sqrt{n}} \sum_{a \in V_A} \frac{1}{\sqrt{\ell_a}} \sum_{b \in V_B} |b\rangle_a |\psi_{a \rightarrow b}^A\rangle |a\rangle_s |\mathbf{0}\rangle_R \quad (3)$$

At this point the first message  $M_1$  is sent to Bob. Let the rest of the protocol starting from this point be  $\mathcal{P}'$ ; that is, in  $\mathcal{P}'$  Bob starts by generating his input from  $M_1$  and  $F_B[s]$ , sends the message  $M_2$  to  $A$ , to which Alice responds with  $M_3$ , and so on. At the end of  $\mathcal{P}'$  we have a register containing the answer which we measure to find  $\text{ans}$ , and the input registers of Alice and Bob, which when measured yield  $f_A$  and  $f_B$ .

Let  $\epsilon_{a \rightarrow b}$  be the probability of error when  $\mathcal{P}'$  is run starting from the state  $|b\rangle_a |\psi_{a \rightarrow b}^A\rangle |a\rangle_s |\mathbf{0}\rangle_R$ . Thus, we have

$$\epsilon_{a \rightarrow b} = \Pr[\text{ans} \neq \text{lsb}(f^{(k+1)}(s)) \mid f_B[s] = a \text{ and } f_A[a] = b],$$

in the original protocol  $\mathcal{P}$  (or in  $\mathcal{P}'$ , when it is run starting from  $|\psi_{\text{in}}\rangle$ ). In particular, we have

$$\epsilon = \mathbb{E}_{a,b}[\epsilon_{a \rightarrow b}] \geq \frac{n - n_a}{n} \mathbb{E}_{a \in {}_u V_A - X_A, b \in {}_u V_B}[\epsilon_{a \rightarrow b}]. \quad (4)$$

In the first expectation,  $(a, b)$  are chosen with the same distribution as  $(f_B[s], f_A[f_B[s]])$  of the given protocol  $\mathcal{P}$ ; in the second, they are chosen uniformly from the sets specified.

**Overview:** We want to eliminate the first message sent by Alice, at the cost of increasing the probability of error slightly, but preserving the total length of the communication. This is based on the following idea (taken from [KNTZ01]). Let  $M_{1,a \rightarrow b}$  be the state of the registers holding the first message when the entire state of Alice's registers is  $\psi_{a \rightarrow b}^A$ ; that is,  $M_{1,a \rightarrow b}$  is the state of the message registers corresponding to message  $M_1$ , when we measure  $F_A[a]$  and observe  $|b\rangle$  there. Note, that  $\psi_{a \rightarrow b}^A$  is a purification of  $M_{1,a \rightarrow b}$ . Also, the state of the first message in  $\mathcal{P}$ ,  $M_1$  is the average, taken over the choices of  $b$ , of  $M_{1,a \rightarrow b}$ .

Suppose there is an  $a \in V_A - X_A$  such that for all  $b$ , the message  $M_{1,a \rightarrow b}$  is independent of  $b$ , that is, it is always the fixed state  $M^*$ . Then, we can eliminate the first message. Informally stated, this amounts to restricting ourselves to the subcase of the protocol when Bob's first pointer  $F_B[s]$  is fixed at  $|a\rangle$ , and Bob generates  $M^*$  himself, and sends some small advice along with his message  $M_2$ , to enable Alice to reproduce the right entanglement between her registers and Bob's. Unfortunately, we will not be able to show that there is an  $a$  and an  $M^*$  such that  $M_{1,a \rightarrow b} = M^*$ , for all  $b$ . Instead, we will show that there is an  $M^*$  that will be close to  $M_{1,a \rightarrow b}$  for typical  $b$ . In fact, the message  $M_1$  (which is the average of  $M_{1,a \rightarrow b}$  as  $b$  varies) will be our  $M^*$ .

Let  $(M_1, \widetilde{M}_1)$  be the canonical purification of the first message of the protocol  $\mathcal{P}$ . Our first goal is to show that if  $M_1$  is close to  $M_{1,a \rightarrow b}$ , then Alice can create a state close to  $|\psi_{a \rightarrow b}^A\rangle$  from  $(M_1, \widetilde{M}_1)$  by applying a unitary transformation on  $\widetilde{M}_1$ . More precisely, suppose  $\|M_{1,a \rightarrow b} - M_1\|_t \triangleq \delta_{a \rightarrow b}$ . Then, by the Local Transition Theorem, there is a unitary transformation  $U_{a \rightarrow b}$  that when applied to  $\widetilde{M}_1$  (together with ancilla qubits initialized to zero) takes the pure state  $(M_1, \widetilde{M}_1)$  to a state  $\tilde{\psi}_{a \rightarrow b}^A$  such that

$$\left\| |\psi_{a \rightarrow b}^A\rangle\langle\psi_{a \rightarrow b}^A| - |\tilde{\psi}_{a \rightarrow b}^A\rangle\langle\tilde{\psi}_{a \rightarrow b}^A| \right\|_t \leq 2\sqrt{\delta_{a \rightarrow b}}. \quad (5)$$

In particular, if the protocol  $\mathcal{P}'$  is run starting from the state  $|\tilde{\psi}_{a \rightarrow b}^A\rangle|\mu\rangle_s|\mathbf{0}\rangle_R$  (instead of  $|\psi_{a \rightarrow b}^A\rangle|\mu\rangle_s|\mathbf{0}\rangle_R$ ), the probability of error is at most  $\epsilon_{a \rightarrow b} + 2\sqrt{\delta_{a \rightarrow b}}$ .

#### 4.1 The protocol $\mathcal{P}_{a \rightarrow b}$

Now, we fix  $a \in V_A$  and  $b \in V_B$  and consider the case when  $f_B(s) = a$  and  $f_A(a) = b$ . We now describe a protocol that functions for this situation (see Figure 1). This is just an intermediate protocol. Later we will describe how we obtain our final protocol (satisfying the requirements of  $Q_{k-1}^B$ ) from this. It will be helpful, meanwhile, to keep in mind that in our final protocol, the roles of  $A$  and  $B$  will be reversed,  $F_B[s]$  will be fixed at  $|a\rangle$  (we will add  $s$  to  $X_B$ ),  $a$  will be our new  $s$ , and the state of  $F_A[a]$  will not be fixed at  $|b\rangle$  but will be the uniform superposition  $|\mu\rangle$ .

**Step 1:** Alice generates the canonical purification  $(M_1, \widetilde{M}_1)$ . Alice applies  $U_{a \rightarrow b}$  to  $\widetilde{M}_1$  (plus some ancilla) to produce the state  $|\tilde{\psi}_{a \rightarrow b}^A\rangle$  in the registers  $(M_1, F_A, W_A)$ .

**Step 2:** Alice and Bob proceed according to the protocol  $\mathcal{P}'$  starting from the state  $|\tilde{\psi}_{a \rightarrow b}^A\rangle = |\tilde{\psi}_{a \rightarrow b}^A\rangle|a\rangle_s|\mathbf{0}\rangle_R$ , where, as before,  $R$  is the set of registers of Bob corresponding to  $(F_B[v] : v \in V_B - \{s\})$  and work qubits  $W_B$ .

Figure 1: The intermediate protocol  $\mathcal{P}_{a \rightarrow b}$

**Revised Step 1:**

- Alice generates the canonical purification  $(M_1, \widetilde{M}_1)$ . Alice applies  $U_{a \rightarrow b}$  to  $\widetilde{M}_1$  (plus some ancilla) to produce the state  $|\widetilde{\psi}_{a \rightarrow b}^A\rangle$  in the registers  $(M_1, F_A, W_A)$ . Alice sends  $M_1$  to Bob.
- Next, to produce input registers satisfying Requirement 1(a), Alice uses a fresh set of registers  $\hat{F}_A$  and sets  $\hat{F}_A[a] = |b\rangle$ . Next, Alice applies a unitary transformation to registers  $(\hat{F}_A[a], F_A, \tilde{F}_A)$  defined by

$$|b\rangle_{\hat{F}[a]}|\psi\rangle_{F_A, \tilde{F}_A} \rightarrow |b\rangle_{\hat{F}[a]}C_{a \rightarrow b}|\psi\rangle_{F_A, \tilde{F}_A}.$$

Before the application of this the registers  $\tilde{F}_A$  are initialized to  $|\mathbf{0}\rangle$  (as in the statement of Fact 2). Alice then copies  $(\tilde{F}_A[u] : u \in V_A - \{a\})$  into  $(\hat{F}_A[u] : u \in V_A - \{a\})$ . The input generation for Alice is now complete.

*Note that at this point if we measure  $(F_A, \hat{F}_A)$ , the resulting random variables  $(f'_{A,a \rightarrow b}, \hat{f}_{A,a \rightarrow b})$  have distribution precisely  $D'_{a \rightarrow b}$  and  $D_{a \rightarrow b}$ . Furthermore, (see Fact 2),*

$$\Pr[f'_{A,a \rightarrow b} \neq \hat{f}_{A,a \rightarrow b}] \leq \frac{1}{2} \cdot 2\sqrt{\delta_{a \rightarrow b}}. \quad (6)$$

**Step 2:** From this point on, Alice and Bob just follow  $\mathcal{P}'$  described above. On receiving  $M_1$ , Bob generates his input and work qubits by appropriately applying the unitary transformation  $U_B$ . He then generates message  $M_2$  and sends it to Alice.

*Let  $|\phi_{a \rightarrow b}\rangle$  denote the state of the entire system just after  $M_2$  is sent to Alice.*

After this, Alice and Bob continue as before. In particular, the Alice continues to use her old input register  $F_A$  (safely) as before. The registers  $\hat{F}$  are not used until the end, when they are measured in order to decide if the answer returned by the protocol is correct.

Figure 2: The revised protocol  $\mathcal{P}_{a \rightarrow b}$

**Remark on the inputs generated:** Suppose we measure registers  $F_A$  just after  $U_{a \rightarrow b}$  has been applied in the above protocol. Let  $f'_{A,a \rightarrow b}$  be the resulting random variable with distribution  $D'_{a \rightarrow b}$ . On the other hand, if we were to measure the same registers in the state  $|\psi_{a \rightarrow b}^A\rangle$ , then the resulting random variable is  $f_{A,a \rightarrow b}$  whose distribution is  $D$ ; that is,  $D$  is the distribution of  $f_A$  conditioned on the event  $f_A[a] = b$ . Then, it follows from (5) and Theorem 2 that

$$\|D_{a \rightarrow b} - D'_{a \rightarrow b}\|_1 \leq 2\sqrt{\delta_{a \rightarrow b}}. \quad (7)$$

We will want Alice's input registers to satisfy Requirement 1(b). Unfortunately, the distribution  $D'$  may not satisfy this requirement automatically, but (7) will help us 'correct' this.

Next consider Bob's input registers. In  $\mathcal{P}$ , Bob's register  $F_B[s]$  contained the uniform superposition  $\mu$  and he generated the input in the rest of the registers himself from  $M_1$  using the unitary transformation  $U_B$ . The input he generated satisfied Requirement 1(b). In  $\mathcal{P}_{a \rightarrow b}$ , Bob applies the same transformation  $U_B$  on  $M_1$ , but  $F_B[s]$  is now  $|a\rangle$  and not  $|\mu\rangle$ . Suppose  $F_B$  is measured at this stage resulting in the random variable  $f_{B,a \rightarrow b} : V_B \rightarrow V_A$ . Note that  $f_{B,a \rightarrow b}$  has the same distribution as  $f_B$  conditioned on the event  $f_B(s) = a$ .

Thus,

- B1.  $f_{B,a \rightarrow b}$  is constant on  $X_A \cup \{s\}$  (in fact,  $f_B[s] = a$ ), and
- B2. the set of random variables  $(f_{B,a \rightarrow b}[v] : v \in V_B - X_B - \{s\})$  are independent and uniformly distributed over  $V_A$ .

**Probability of error in  $\mathcal{P}_{a \rightarrow b}$ :** By (5) and Theorem 2, the probability of error of  $\mathcal{P}_{a \rightarrow b}$ , which we denote by  $\tilde{\epsilon}_{a \rightarrow b}$ , is at most  $\epsilon_{a \rightarrow b} + 2\sqrt{\delta_{a \rightarrow b}}$ .

**Correcting Alice's input registers:** The random variable  $f'_{A,a \rightarrow b}$  that results from measuring  $F_A$  has a distribution  $D'_{a \rightarrow b}$  which is close to the desired distribution  $D_{a \rightarrow b}$  of  $f_{A,a \rightarrow b}$  (by (7) above). It will be easier to satisfy Requirement 1(b), however, if we could arrange that the distribution of Alice's inputs is exactly  $D_{a \rightarrow b}$ . To do this, we use Fact 2; let  $C_{a \rightarrow b}$  be the unitary transformation corresponding to  $D'_{a \rightarrow b}$  and  $D_{a \rightarrow b}$ . We revise the protocol  $\mathcal{P}_{a \rightarrow b}$  by including this operation (see Figure 2).

**Error probability of the revised protocol:** At that end of the protocol, we measure all registers and obtain the answer ans, and the inputs  $\hat{f}_{A,a \rightarrow b}$  and  $f_{B,a \rightarrow b}$ . We also have  $f'_{A,a \rightarrow b}$  corresponding to Alice's old input registers  $F_A$ . Let  $\hat{f}_{a \rightarrow b} = \hat{f}_{A,a \rightarrow b} \cup f_{B,a \rightarrow b}$  and  $f'_{a \rightarrow b} = f'_{A,a \rightarrow b} \cup f_{B,a \rightarrow b}$ . This revised protocol makes an error whenever  $\text{ans} \neq \text{lsb} \hat{f}_{a \rightarrow b}^{\hat{(k+1)}}(s)$ . We then have

$$\begin{aligned}
\hat{\epsilon}_{a \rightarrow b} &\stackrel{\Delta}{=} \Pr[\text{ans} \neq \text{lsb} \hat{f}_{a \rightarrow b}^{\hat{(k+1)}}(s)] \\
&\leq \Pr[\hat{f}_{a \rightarrow b} \neq f'_{a \rightarrow b}] + \Pr[\text{ans} \neq \text{lsb} f'_{a \rightarrow b}^{\hat{(k+1)}}(s)] \\
&\leq \frac{1}{2} \cdot 2\sqrt{\delta_{a \rightarrow b}} + \epsilon_{a \rightarrow b} + 2\sqrt{\delta_{a \rightarrow b}} \\
&= \epsilon_{a \rightarrow b} + 3\sqrt{\delta_{a \rightarrow b}}. \tag{8}
\end{aligned}$$

## 4.2 The final protocol: $\mathcal{P}_a$

A small modification now gives us our final protocol, which will satisfy the requirements for  $Q^{k-1}$ . We make two changes to the revised version of  $\mathcal{P}_{a \rightarrow b}$ . First, instead of Alice sending  $M_1$  and retaining  $\widetilde{M}_1$ , now Bob creates the canonical purification  $(M_1, \widetilde{M}_1)$  and sends Alice  $\widetilde{M}_1$ , while retaining  $M_1$ . Second, in  $\mathcal{P}_{a \rightarrow b}$ , the register  $\hat{F}_A[a]$  is fixed to the value  $|b\rangle$ . Now, however, Alice starts with  $|\mu\rangle$  in  $\hat{F}[a]$ . With these modifications, Alice's role in the input generation phase of the new protocol is similar to Bob's role in the protocol we started with. The resulting protocol  $\mathcal{P}_a$  (see Figure 3) depends on the choice of  $a$ . Using an averaging argument we will conclude that there is a choice for  $a \in V_A$  so that  $\mathcal{P}_a$  satisfies the requirements for  $Q_{k-1}^B$  as needed in Lemma 2(a).

**The probability of error of  $\mathcal{P}_a$ :** For  $a \in V_A - X_A$ , let  $\hat{\epsilon}_a$  be the probability of error of  $\mathcal{P}_a$ . Then, by (8), we have

$$\hat{\epsilon}_a = \mathbb{E}_{b \in {}_a V_B} [\hat{\epsilon}_{a \rightarrow b}] \tag{9}$$

$$\leq \mathbb{E}_{b \in {}_a V_B} [\epsilon_{a \rightarrow b} + 3\sqrt{\delta_{a \rightarrow b}}]. \tag{10}$$

The new input registers for Alice will be denoted by  $\hat{F}_A$ . The old input registers will continue to exist, but they will count as work qubits of Alice. Initially, in the register  $\hat{F}_A[a]$  we place a uniform superposition  $|\mu\rangle$ . All other registers are initialized to 0.

**Step 1:** Bob generates the canonical purification  $(M_1, \widetilde{M}_1)$  of the first message of  $\mathcal{P}$ . He sets his register  $F_B[s]$  to the state  $|a\rangle$ , and using the transformation  $U_B$  generates his inputs  $F_B$  and work qubits  $W_B$ . Then, he generates the first message of protocol  $\mathcal{P}'$  (this corresponds message  $M_2$  of the  $\mathcal{P}$ ), and sends this message along with  $\widetilde{M}_1$  to Alice.

**Step 2:** (a) One receiving  $\widetilde{M}_1$ , Alice applies a unitary transform on registers  $(\hat{F}_A[a], \widetilde{M}_1, A)$  to generate a state in registers  $F_A$  (the old input registers) and  $W_A$  (the work qubits of the original protocol). Here,  $A$  is a set of ancilla qubits initialized to 0. This unitary transformation acts according to the rule

$$|b\rangle_{\hat{F}[a]}|\theta\rangle_{\widetilde{M}_1, A} \mapsto |b\rangle_{\hat{F}[a]}U_{a \rightarrow b}|\theta\rangle_{\widetilde{M}_1, A}.$$

Note that this transformation is safe on  $\hat{F}[a]$ .

(b) Since  $F_A$  is not in the desired state, Alice applies the correction used in the revised Step 1 of  $\mathcal{P}_{a \rightarrow b}$ . That is, she applies a unitary transformation to registers  $(\hat{F}_A[a], F_A, \tilde{F}_A)$  defined by

$$|b\rangle_{\hat{F}[a]}|\psi\rangle_{F_A, \tilde{F}_A} \mapsto |b\rangle_{\hat{F}[a]}C_{a \rightarrow b}|\psi\rangle_{F_A, \tilde{F}_A}.$$

Before the application of this the registers  $\tilde{F}_A$  are initialized to 0. Alice then copies  $(\tilde{F}_A[u] : u \in V_A - \{a\})$  into  $(\hat{F}_A[u] : u \in V_A - \{a\})$ . For the purpose of satisfying Requirement 1(b),  $\tilde{F}_A$  are to be treated as  $A$ 's input register.

The state of entire system at this point is precisely  $\frac{1}{\sqrt{n}} \sum_{b \in V_B} |\phi_{a \rightarrow b}\rangle$ , where  $|\phi_{a \rightarrow b}\rangle$  is the state at the corresponding point in the revised protocol  $\mathcal{P}_{a \rightarrow b}$  (see Figure 2). The rest of the protocol operates safely on  $F_A, \hat{F}_A$  and  $F_B$ . In fact, no unitary transform will now be applied to registers  $\hat{F}_A$ .

**Step 3:** Alice resumes the protocol  $\mathcal{P}'$ . Note that Bob has already executed the first step of  $\mathcal{P}'$  and sent the first message (which corresponds to message  $M_2$  of the original protocol). Alice responds to this message as before.

While executing  $\mathcal{P}'$ , the old input registers  $F_A$  are used. The new registers  $\hat{F}_A$  are not touched by any unitary transformation from now on. At the end, however, when we try to decide if an error has been made, we will measure all registers, and check if the answer  $\text{ans}'$  agrees with the answer  $\text{ans}(\hat{f}_A, f_B)$ , where  $\hat{f}_A$  is the random variable obtained by measuring the new input registers  $\hat{F}_A$ .

Figure 3: The protocol  $\mathcal{P}_a$

We need to show that there exists an  $a$  such that  $\hat{\epsilon}_a$  is small. For this we consider the average of  $\hat{\epsilon}_a$  as  $a$  is chosen uniformly from  $V_A - X_A$ :

$$\mathbb{E}_{a \in V_A - X_A} [\hat{\epsilon}_a] \leq \mathbb{E}_{a,b} [\epsilon_{a \rightarrow b} + 3\sqrt{\delta_{a \rightarrow b}}], \quad (11)$$

where on the right  $a$  is chosen uniformly from  $V_A - X_A$  and  $b$  is chosen independently and uniformly from  $V_B$ . (From now on, when we average over  $a$  and  $b$ , we will assume that they are chosen in this manner.) By (4), we have

$$\mathbb{E}_{a,b} [\epsilon_{a \rightarrow b}] \leq \left( \frac{n}{n - n_a} \right) \epsilon. \quad (12)$$

It remains to bound  $\mathbb{E}_{a,b} [\sqrt{\delta_{a \rightarrow b}}]$ . Consider the state obtained by measuring Alice's input registers  $F_A$  just before  $M_1$  is sent to Bob in the original protocol. As stated earlier, if the value  $b$  is observed for  $F_A[a]$ , then the state of the message registers will be  $M_{1,a \rightarrow b}$ ; also,  $M_1$  is the average of these states, that is,  $M_1 = \frac{1}{n} \sum_{b \in V_B} M_{1,a \rightarrow b}$ .

**Claim 1** For  $a \in V_A - X_A$ ,  $\mathbb{E}_b [\delta_{a \rightarrow b}] \leq \sqrt{(2 \ln 2) I(f_A[a] : M_1)}$ .

**Proof:** Consider the encoding of elements of  $V_B$  given by  $b \mapsto M_{1,a \rightarrow b}$  by restricting attention the registers  $F_A[a]$  and  $M_1$ . Our claim now follows from the Average Encoding Theorem (Theorem 3) and the definition of  $\delta_{a \rightarrow b}$ .  $\square$

**Claim 2**  $\mathbb{E}_{a \in V_A - X_A} [I(f_A[a] : M_1)] \leq \left( \frac{n}{n - n_a} \right) c_1$ .

**Proof:** Using Fact 1 and (7), we have  $c_1 n \geq I(f_A : M_1) \geq \sum_{a \in V_A} I(f_A[a] : M_1) \geq \sum_{a \in V_A - X_A} I(f_A[a] : M_1)$ .  $\square$

By combining these two claims, and noting that the square-root function is concave, we obtain

$$\mathbb{E}_{a,b} [\delta_{a \rightarrow b}] \leq \mathbb{E}_a [\sqrt{(2 \ln 2) I(f_A[a] : M_1)}] \leq \sqrt{(2 \ln 2) \mathbb{E}_a [I(f_A[a] : M_1)]} \leq \sqrt{\left( \frac{n}{n - n_a} \right) (2 \ln 2) c_1}. \quad (13)$$

This implies, again because the square root is concave, that

$$\mathbb{E}_{a,b} [\sqrt{\delta_{a \rightarrow b}}] \leq \left[ \left( \frac{n}{n - n_a} \right) (2 \ln 2) c_1 \right]^{\frac{1}{4}}. \quad (14)$$

Now we return to (11), and use (12) and (14) to obtain

$$\mathbb{E}_a [\hat{\epsilon}_a] \leq \left( \frac{n}{n - n_a} \right) \left[ \epsilon + 3((2 \ln 2) c_1)^{\frac{1}{4}} \right].$$

Thus, there exists an  $a \in V_A - X_A$  such that

$$\hat{\epsilon}_a \leq \left( \frac{n}{n - n_a} \right) \left[ \epsilon + 3((2 \ln 2) c_1)^{\frac{1}{4}} \right].$$

Now, it can be verified, the protocol  $\mathcal{P}_a$  satisfies the requirements for  $Q_{k-1}^B(c_1 + c_2, c_3, \dots, c_k, n_A, n_B + 1, \hat{\epsilon}_a)$ . This shows Part (a) of Lemma 2. Part (b) can be established similarly.

## References

- [AKN01] D. Aharonov, A. Kitaev, and N. Nisan. Quantum circuits with mixed states. In *Proceedings of International Colloquium on Automata, Languages and Programming (ICALP)*, pages 358–369, 2001.
- [CT91] T. Cover and J. Thomas. *Elements of Information Theory*. Wiley Series in Telecommunications. John Wiley and Sons, 1991.
- [DGS87] P. Duris, Z. Galil, and G. Schnitger. Lower bounds on communication complexity. *Information and Computation*, 73:1–22, 1987.
- [JRS02] R. Jain, J. Radhakrishnan, and P. Sen. Privacy and interaction in quantum communication complexity and a theorem about the relative entropy of quantum states. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, 2002.
- [Kla00] H. Klauck. On quantum and probabilistic communication: Las Vegas and one-way protocols. In *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing*, pages 644–651, 2000.
- [KNTZ01] H. Klauck, A. Nayak, A. Ta-Shma, and D. Zuckerman. Interaction in quantum communication and the complexity of set disjointness. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing*, pages 124–133, 2001.
- [MNSW98] P. B. Miltersen, N. Nisan, S. Safra, and A. Wigderson. On data structures and asymmetric communication complexity. *Journal of Computer and System Sciences*, 57(1):37–49, 1998.
- [NC00] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
- [NW93] N. Nisan and A. Wigderson. Rounds in communication complexity revisited. *SIAM Journal of Computing*, 22:211–219, 1993.
- [PRV01] S. Ponzio, J. Radhakrishnan, and S. Venkatesh. The communication complexity of pointer chasing. *Journal of Computer and System Sciences*, 62(2):323–355, 2001.
- [PS84] C. Papadimitriou and M. Sipser. Communication complexity. *Journal of Computer and System Sciences*, 28:260–269, 1984.
- [SV98] P. Sen and S. Venkatesh. Lower bounds in quantum cell probe model. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing*, pages 20–30, 1998.
- [Yao79] A. C-C. Yao. Some complexity questions related to distributed computing. In *Proceedings of the 11th Annual ACM Symposium on Theory of Computing*, pages 209–213, 1979.
- [Yao93] A. C-C. Yao. Quantum circuit complexity. In *Proceedings of the 34th Annual IEEE Symposium on Foundations of Computer Science*, pages 352–361, 1993.