Privacy and interaction in quantum communication complexity and a theorem about the relative entropy of quantum states

Rahul Jain[*] Jaikumar Radhakrishnan[*] Pranab Sen[†]

Abstract

We prove a theorem about the relative entropy of quantum states, which roughly states that if the relative entropy, $S(\rho\|\sigma) \stackrel{\Delta}{=} \mathrm{Tr}\ \rho(\log\rho - \log\sigma)$, of two quantum states ρ and σ is at most c,

then $\rho/2^{O(c)}$ 'sits inside' σ. Using this 'substate' theorem, we give tight lower bounds for the privacy loss of bounded error quantum communication protocols for the index function problem. We also give tight lower bounds for the k-round bounded error quantum communication complexity of the pointer chasing chasing problem, when the wrong player starts, and all the $\log n$ bits of the kth pointer are desired.

1 Introduction

The main contribution of this paper is a theorem, called *Substate Theorem*, about relative entropy; it states, roughly, that if the relative entropy, $S(\rho\|\sigma) \stackrel{\Delta}{=} \mathrm{Tr}\ \rho(\log\rho - \log\sigma)$, of two quantum states ρ and σ is at most c, then $\rho/2^{O(c)}$ *sits inside σ*. This implies, for example, that if some event occurs in ρ with probability p, then it occurs in σ with probability at least $p/2^{O(c/p)}$. We shall present below two natural problems in whose solution this result plays a crucial part. First, let us motivate the substate theorem by considering its classical analog. Let P and Q be probability distributions on the set $[n]$ with relative entropy bounded by c, that is

$$S(P\|Q) \stackrel{\Delta}{=} \sum_{i\in[n]} P(i)\log_2\frac{P(i)}{Q(i)} \ \leq\ c. \qquad (1)$$

When $c \ll 1$, this implies that P and Q are close to each other; indeed, one can show that (see [CT91, Lemma 12.6.1])

$$\|P-Q\|_t \stackrel{\Delta}{=} \sum_{i\in[n]} |P(i)-Q(i)| \ \leq\ \sqrt{(2\ln 2)c}. \qquad (2)$$

That is, the probability of an event $\mathcal{E} \subseteq [n]$ in P is close to its probability in Q: $|P(\mathcal{E})-Q(\mathcal{E})| \leq \sqrt{(c\ln 2)/2}$. We are, however, concerned with the situation when $c \gg 1$. In that case, (2) becomes weak: we cannot even infer from it that an event $\mathcal{E}$ with probability $3/4$ in P has positive probability in Q. But is it true that when $S(P\|Q) < +\infty$ $P(\mathcal{E}) > 0$, then $Q(\mathcal{E}) > 0$? Yes! To see this, let us reinterpret the expression in (1) as the expectation of $\log P(i)/Q(i)$ as i is chosen according to P. Thus, one is lead to believe that if $S(P\|Q) \leq c < +\infty$, then $\log P(i)/Q(i)$ is typically bounded by c, that is, $P(i)/Q(i)$ is typically bounded by 2^c. One can formalise this intuition and show, for all $r > 1$,

$$\Pr_{i\in P}\left[\frac{P(i)}{Q(i)} \geq 2^{r(c+1)}\right] \leq \frac{1}{r}.$$

Let $\mathsf{Good} \stackrel{\Delta}{=} \{i : P(i)/2^{rc} \leq Q(i)\}$, $P'(i) \stackrel{\Delta}{=} P(i \mid i \in \mathsf{Good})$. That is in P' we just discard the bad values of i, and normalise. Now, $\frac{r-1}{r2^{r(c+1)}}P'$ is dominated by Q everywhere. We have thus proved the following.

Proposition 1 *If $S(P\|Q) \leq c$, then for all $r > 1$, there exists a distribution P' such that $|P-P'|_1 \leq \frac{2}{r}$ and $Q = \alpha P' + (1-\alpha)P''$, where P'' is some other distribution and $\alpha = 2^{-O(rc)}$.*

Let us return to our event $\mathcal{E}$ that occurred with some small probability p in P. Now, if we take r to be $2/p$, then $\mathcal{E}$ occurs with probability at least $p/2$ in P', and hence appears with probability $p/2^{O(rc)}$ in Q . Thus, we have shown that even though P and Q are far apart as distributions, events that have positive probability (no matter how small) in P, continue to have positive probability in Q.

The main contribution of this paper is a quantum analog of Proposition 1.

Result 1 (Substate Theorem) Suppose ρ and σ are quantum states with $S(\rho\|\sigma) \leq c$. Then, for all $r > 1$, there

[*] 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. Supported by the EU 5th framework program QAIP IST-1999-11234, and by CNRS/STIC 01N80/0502 and 01N80/0607 grants.

are states ρ' and ρ'' such that $\|\rho - \rho'\|_t \leq 4/\sqrt{r}$ and $\sigma = \alpha\rho' + (1-\alpha)\rho''$, where $\alpha = 2^{O(rc)}$.
(This has been stated here in a form that brings out the analogy with the classical statement above. In Section 3, we have a more nuanced statement which is better suited for our applications.)

1.1 The pointer chasing problem

Our first application of the Substate Theorem concerns the pointer chasing problem in two-party communication complexity.

> Let V_A and V_B be disjoint subsets of size n. Player A is given a function $F_A : V_A \to V_B$ and player B is given a function $F_B : V_B \to V_A$. Let $F = F_A \cup F_B$. There is a fixed vertex s in V_B. A and B need to communicate to determine $t = F^{(k+1)}(s)$, where k and s are known to both parties in advance.

If B starts the communication, then there is a straightforward classical deterministic protocol where one of the players can determine t after k messages of $\log n$ bits have been exchanged. It appears much harder, however, to solve the problem efficiently with k messages, when A 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 a lot 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 A and B must exchange $\Omega(n/k^2 - k\log n)$ bits to solve P_k; their bound was improved by Klauck [Kla00] to $\Omega(\frac{n}{k}+k)$. These lower bounds hold even when A and B are allowed to toss coins and err with some small probability. Furthermore, they hold for the bit version of the problem, where one only wants to determine (say) the least significant bit of t, and not all of t. For this bit version of the problem, a deterministic protocol with $O(n + k\log n)$ bits of communication was shown by Ponzio, Radhakrishnan and Venkatesh [PRV01]. Thus, the lower and upper bounds are quite close in the the classical setting for this version of the problem. For the full version of the problem, where one needs to determine all of t, the best upper bound, $O(n\log^{(k)} n)$, comes from a classical deterministic protocol due to Damm, Jukna and Sgall [DJS98]. Note that for constant k, this is superlinear. Ponzio, Radhakrishnan and Venkatesh [PRV01] showed a matching lower bound in the classical setting.

The pointer chasing problem was studied recently in the quantum communication complexity model by Klauck, Nayak, Ta-Shma and Zuckerman [KNTZ01], who, using interesting information-theoretic techniques, showed a lower bound of $\Omega(n/2^{2^{O(k)}})$ for the *bit version* of this problem. They *did not consider* the full version of the problem. (Note that the classical application of the lower bound for P_k to monotone circuit depth in the paper of Nisan and Wigderson [NW93, Theorem 2.7] is valid for the full version of the problem, not just for the bit version.) We fill this gap.

Result 2: For any constant k, the bounded error quantum communication complexity of the pointer jumping problem P_k (full pointer version) is $\Omega(n\log^{(k)} n)$.

1.2 Privacy and communication complexity

Our second application of the substate theorem concerns the index function problem [MNSW98, ANTV99, Nay99]:

> INDEX$_n$: There are two players A and B. A is given an input $x \in \{0,1\}^n$ and B is given an index $i \in [n]$. They must exchange messages so that in the end B knows x_i.

Background: Miltersen, Nisan, Safra and Wigderson [MNSW98] considered this problem (under the name *set membership problem*) in the classical setting, and showed that if B sends a total of at most b bits, then A must send $n/2^{O(b)}$ bits. Note that this is optimal as there is a trivial protocol where B sends the first b bits of his index to A, and A replies by sending the corresponding part of her bit string.

In the quantum setting, Nayak [Nay99] (see also Cleve et al. [CvDNT98]), showed that if B sends no messages at all, then A must send at least $\Omega(n)$ bits. This bound holds even if the players share EPR pairs in advance, or if A and B interact but B's messages do not depend on his input i. However, the case where B is allowed to send a few qubits based on his input in order to reduce the communication from A, does not seem to have been considered before.

In this paper, we generalise the Nayak's result to a statement of the following form: if B 'leaks' only a small number of bits of information about his input, then A must send a large number of bits. Before we present our result, let us explain what we mean when we say that B 'leaks' only a small number of bits of information about his input. Fix a protocol for the index function problem. Assume that B's input J is a random index $i \in [n]$. Suppose B operates faithfully according to the protocol, but A deviates from it and manages to get her registers R entangled with J: we say that B leaks only b bits of information about his input if the mutual information between J and R, $I(J : R)$, is at most b. This must hold for all strategies adopted by A, which have the property that the reduced density matrix of

Bob's qubits is at all times the same as in the original protocol. In other words, A wants to cheat and gather a lot of information about B's input, but B should not be able to figure out that A is cheating. Note that we do not assume that B's messages contain only b qubits, they can be arbitrarily long. In the quantum setting, A has a big bag of tricks she can use in order to extract information from B; for example, she can place a superposition of states in the registers corresponding to he input and extract information about B's input (see [CvDNT98, Kla02] for details).

Klauck [Kla02] recently studied privacy in quantum communication protocols. In Klauck's setting, two players collaborate to compute a function, but at any point, one of the players might decide to terminate the protocol and try to infer something about the input of the other player using the bits in his possession. The players are *honest but curious*: in a sense, they don't deviate from the protocol in any way other than, perhaps, by stopping early. In this model, Klauck shows that there is a protocol for the *set disjointness* function where neither player reveals more than $O((\log n)^2)$ bits of information about his input, whereas in every classical protocol, at least, one of the players leaks $\Omega(\sqrt{n}/\log n)$ bits of information about his input. Klauck, however, proves no *lower bounds* for privacy loss in the quantum setting. Our model of privacy is more stringent. We allow malicious players who can deviate arbitrarily from the protocol, but with the restriction that the honest player does not realise the difference. Note that this precludes the malicious player from prematurely aborting the protocol.

Result 3 (informal statement) If there is a protocol for the index function problem where B leaks only b bits of information about his input i, then A must *leak* $\Omega(n/2^{O(b)})$ bits of information about her input x.

Corollary (informal statement) For the index function problem, one of the players must leak $\Omega(\log n)$ bits of information about his input.

General result and other problems: The index function problem is just one of several problems where a statement like Result 2 can be proved using our technique. In fact, it follows easily that if the communication matrix of the function has VC-dimension at least k, then one of the players must leak at least $\Omega(\log k)$ bits of information about his input. In particular, this implies an $\Omega(\log n)$ loss in privacy for the set disjointness and inner product modulo 2 problems.

1.3 Organisation of the rest of the paper

In the next section, we give some information theoretic preliminaries and formally define our model of privacy loss in quantum communication protocols. In Section **??**, we give a complete proof of Result 2, assuming the Substate Theorem. In Section 3, we give proof of the substate thoerem. In Section 4, we give the proof of the lower bound for the Pointer Chasing problem.

2 Preliminaries

2.1 Notation and definitions

We use $H(X)$ to denote the Shannon entropy of a classical random variable X. If A is a quantum system with density matrix ρ, then $S(A) \stackrel{\Delta}{=} S(\rho) \stackrel{\Delta}{=} -\mathrm{Tr}\,\rho \log \rho$ 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 $I(A : B) \stackrel{\Delta}{=} S(A) + S(B) - S(AB)$. If ρ, σ are density matrices in the same Hilbert space, their *relative entropy* is defined as $S(\rho \| \sigma) \stackrel{\Delta}{=} \mathrm{Tr}\,(\rho(\log \rho - \log \sigma))$. The trace norm of a linear operator A is defined as $\|A\|_t \stackrel{\Delta}{=} \mathrm{Tr}\,\sqrt{A^\dagger A}$. The trace distance between two linear operators A, B is defined as $\|A - B\|_t$. For distributions D and D' on a finite set X, their total variational distance is given by $\|D - D'\|_1 \stackrel{\Delta}{=} \sum_{x \in X} |D(x) - D'(x)|$. We will use the notation $A \geq B$ for Hermitian operators A, B in the same finite dimensional Hilbert space $\mathcal{H}$ as a shorthand for the statement '$A - B$ is positive semidefinite'. Thus, $A \geq 0$ denotes that A is positive semidefinite. We use $B(\rho, \sigma) \stackrel{\Delta}{=} \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_t$ to denote the *Bhattacharya distuinguishability* [**?**] of density matrices ρ, σ. Note that this is the square root of the *fidelity* of Jozsa [Joz94].

The above notations and definitions are standard. For excellent introductions to classical and quantum information theory, see the books by Cover and Thomas [CT91] and Nielsen and Chuang [NC00] respectively.

In this paper, we consider two party quantum communication protocols as defined by Yao [Yao93].

Definition 1 (Safe transformation, protocols) *Let $\mathcal{H}$ and $\mathcal{K}$ be finite-dimensional Hilbert spaces, with computational orthonormal 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 $k \in K$,*

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

We say that a quantum communication 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 to a quantum communication protocol are classical, we can always assume that the protocol is safe, since the players can make a secure copy of their inputs before beginning the protocol. From now on we assume that all our protocols are safe.

2.2 Some basic facts

We will use the following elementary facts, which we state without proof.

Fact 1 *Suppose X, Q are two disjoint finite dimensional 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*

$$I(X : Q) = \sum_x p_x S(\sigma_x \| \sigma).$$

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

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

and maps other computational basis vectors suitably, preserving orthonormality. 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 $\Pr[Z \neq Z'] \leq \frac{\delta}{2}$. Note that C acts safely on R'.

We will require the following minimax theorem from game theory. It follows by combining Proposition 20.3 (which shows the existence of Nash equilibrium a^* in strategic games) and Proposition 22.2 (which connects Nash equilibrium and the min-max theorem for games defined using a pay-off function such as u) of Osborne and Rubinstein's [OR94] book on game theory.

Fact 3 *Let A_1, A_2 be non-empty, convex and compact subsets of $\mathbb{R}^n$ for some n. Let $u : A_1 \times A_2 \to \mathbb{R}$ be a continuous function, such that*

- *$\forall a_2 \in A_2$, the set $\{a_1 \in A_1 : u(a_1, a_2) \geq u(a_1', a_2) \, \forall a_1' \in A_1\}$ is convex; and*
- *$\forall a_1 \in A_1$, the set $\{a_2 \in A_2 : u(a_1, a_2) \leq u(a_1, a_2') \, \forall a_2' \in A_2\}$ is convex.*

Then, there is an $a^* \in A_1 \times A_2$ such that

$$\max_{a_1 \in A_1} \min_{a_2 \in A_2} u(a_1, a_2) = u(a^*) = \min_{a_2 \in A_2} \max_{a_1 \in A_1} u(a_1, a_2).$$

We will also be using several information theoretic facts that can be found in the books by Cover and Thomas [CT91] and Nielsen and Chuang [NC00] without explicitly stating them here.

3 Proof of the substate theorem

To prove the Substate Theorem, it will be useful to define a new notion of distinguishability between density matrices. We shall call this notion *observational divergence*.

Definition 2 (Observational divergence) *Let ρ, σ be density matrices in the same finite dimensional Hilbert space $\mathcal{H}$. Their observational divergence is defined as*

$$D(\rho \| \sigma) \triangleq \sup_F \left(\mathrm{Tr}\,(F\rho) \log \frac{\mathrm{Tr}\,(F\rho)}{\mathrm{Tr}\,(F\sigma)} \right),$$

where F above ranges over POVM elements on $\mathcal{H}$ such that $\mathrm{Tr}\,(F\sigma) \neq 0$.

We note that the relative entropy is an upper bound on the divergence to within an additive constant.

Lemma 1 *Let ρ, σ be density matrices in the same finite dimensional Hilbert space $\mathcal{H}$. Then, $D(\rho \| \sigma) < S(\rho \| \sigma) + 1$.*

Proof: **(Sketch)** Follows from the Lindblad-Uhlmann monotonicity of relative entropy. ■

We now prove the following lemma, which can be thought of as a Substate Theorem when the first density matrix is in fact a pure state.

Lemma 2 *Let $|\psi\rangle$ be a pure state and σ be a density matrix in the same finite dimensional Hilbert space $\mathcal{H}$. Let $k \triangleq D\left((|\psi\rangle\langle\psi|) \| \sigma\right)$. If $k > 0$, then for all $r > 1$, there exists a pure state $|\phi\rangle$ (depending on r) such that*

$$\| |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \|_t \leq \frac{4}{\sqrt{r}} \quad \text{and} \quad \left(\frac{r-1}{r 2^{rk}} \right) |\phi\rangle\langle\phi| \leq \sigma.$$

Proof: **(Sketch)** We assume without loss of generality that $0 < k < +\infty$. Consider $M \triangleq \sigma - (|\psi\rangle\langle\psi|/2^{rk})$. Since $-(|\psi\rangle\langle\psi|/2^{rk})$ has exactly one non-zero eigenvalue and this eigenvalue is negative viz. $-1/2^{rk}$, and σ is positive semidefinite, M is a hermitian matrix with at most one negative eigenvalue.

If $M \geq 0$ we take $|\phi\rangle$ to be $|\psi\rangle$. The lemma trivially holds in this case.

Otherwise, let $|w\rangle$ be the eigenvector corresponding to the unique negative eigenvalue $-\alpha$ of M. One can check that $|\langle\psi|w\rangle|^2 < \frac{1}{r} < 1$. In particular, this shows that $|\psi\rangle, |w\rangle$ are linearly independent.

Let $n \triangleq \dim(\mathcal{H})$. Let $\{|v\rangle, |w\rangle\}$ be an orthonormal basis for the two dimensional subspace of $\mathcal{H}$ spanned by $\{|\psi\rangle, |w\rangle\}$. Extend it to $\{|v_1\rangle, \ldots, |v_{n-2}\rangle, |v\rangle, |w\rangle\}$, an orthonormal basis for the entire space $\mathcal{H}$. In this basis we have the following matrix equation,

$$\begin{bmatrix} F & e & d \\ e^\dagger & a & b \\ d^\dagger & \overline{b} & c \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0^\dagger & x & y \\ 0^\dagger & \overline{y} & z \end{bmatrix} = \begin{bmatrix} P & l \\ l^\dagger & -\alpha \end{bmatrix}, \tag{3}$$

where the first, second and third matrices are σ, $|\psi\rangle\langle\psi|/2^{rk}$ and M respectively. F is an $(n-2) \times (n-2)$ matrix, P is an $(n-1) \times (n-1)$ matrix, d, e are $(n-2) \times 1$ matrices and l is an $(n-1) \times 1$ matrix. a, c, x, z, α are non-negative real numbers and b, y are complex numbers. The zeroes above denote all zero matrices of appropriate dimensions. The $\dagger$ denotes conjugate transpose of matrices and $\overline{b}$ denote the complex conjugate of scalar b.

By inspection, one can show that $b = y \neq 0$, $c > 0$ and $ac \geq |b|^2$.

We can now write $\sigma = \sigma_1 + \sigma_2$, where

$$\sigma_1 \triangleq \begin{bmatrix} F & e & 0 \\ e^\dagger & a - \frac{|b|^2}{c} & 0 \\ 0^\dagger & 0 & 0 \end{bmatrix}$$

and

$$\sigma_2 \triangleq \begin{bmatrix} 0 & 0 & 0 \\ 0^\dagger & \frac{|b|^2}{c} & b \\ 0^\dagger & \overline{b} & c \end{bmatrix}.$$

Note that $\sigma_2 \geq 0$, (in fact, σ_2 has one dimensional support). It can be checked that $\sigma_1 \geq 0$. Hence, $\sigma \geq \sigma_2$. Let $|\phi\rangle\langle\phi| \triangleq \frac{\sigma_2}{\text{Tr } \sigma_2}$. By a direct computation, one can check that Tr $\sigma_2 > \frac{r-1}{r2^{rk}}$ and $\| |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \|_t < \frac{4}{\sqrt{r}}$. This establishes the first assertion of the lemma and completes the proof. ■

We next prove the following lemma, which can be thought of as an 'observational substate' lemma.

Lemma 3 *Consider two finite dimensional Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, $\dim(\mathcal{K}) \geq \dim(\mathcal{H})$. Let ρ, σ be density matrices in $\mathcal{H}$. Let $|\psi\rangle$ be a purification of ρ in $\mathcal{H} \otimes \mathcal{K}$. Let F be a POVM element on $\mathcal{H} \otimes \mathcal{K}$. Then there exists a purification $|\phi\rangle$ of σ in $\mathcal{H} \otimes \mathcal{K}$ such that $q \geq \frac{p}{2^{k'/p}}$, where $p \triangleq$ Tr $(F|\psi\rangle\langle\psi|)$, $q \triangleq$ Tr $(F|\phi\rangle\langle\phi|)$ and $k' \triangleq 4D(\rho\|\sigma) + 2$.*

Proof: **(Sketch)** We assume without loss of generality that $0 < D(\rho\|\sigma) < +\infty$ and that $p > 0$. Let $n \triangleq \dim(\mathcal{H})$ and $\{|\alpha_i\rangle\}_{i=1}^n$ be the orthonormal eigenvectors of F with corresponding eigenvalues $\{\lambda_i\}_{i=1}^n$. We have,

$$p = \sum_{i=1}^n \lambda_i |\langle\alpha_i|\psi\rangle|^2 \quad \text{and} \quad q = \sum_{i=1}^n \lambda_i |\langle\alpha_i|\phi\rangle|^2.$$

Define,

$$|\theta'\rangle \triangleq \frac{\sum_{i=1}^n \lambda_i \langle\alpha_i|\psi\rangle |\alpha_i\rangle}{\sqrt{p}} \quad \text{and} \quad |\theta\rangle \triangleq \frac{|\theta'\rangle}{\||\theta'\rangle\|}.$$

Note that $p = |\langle\psi|\theta\rangle|^2 \||\theta'\rangle\|^2$ and $0 < \||\theta'\rangle\|^2 \leq 1$. Using the Cauchy-Schwarz inequality, one can check that $|\langle\phi|\theta\rangle|^2 \||\theta'\rangle\|^2 \leq q$ and

$$\frac{p}{2^{k'/p}} = \frac{|\langle\psi|\theta\rangle|^2 \||\theta'\rangle\|^2}{2^{k'/(|\langle\psi|\theta\rangle|^2 \||\theta'\rangle\|^2)}} \leq \frac{|\langle\psi|\theta\rangle|^2 \||\theta'\rangle\|^2}{2^{k'/|\langle\psi|\theta\rangle|^2}}.$$

Hence, it will suffice to show that there exists a purification $|\phi\rangle$ of σ in $\mathcal{H} \otimes \mathcal{K}$ such that

$$|\langle\phi|\theta\rangle|^2 \geq \frac{|\langle\psi|\theta\rangle|^2}{2^{k'/|\langle\psi|\theta\rangle|^2}}.$$

Define the density matrix τ in $\mathcal{H}$ as $\tau \triangleq \text{Tr}_{\mathcal{K}} |\theta\rangle\langle\theta|$. There is a purification $|\phi\rangle$ of σ in $\mathcal{H} \otimes \mathcal{K}$ and a POVM $\{F_1, \ldots, F_l\}$ in $\mathcal{H}$ such that,

$$|\langle\phi|\theta\rangle| = B(\tau, \sigma) = \sum_{i=1}^l \sqrt{c_i b_i}, \tag{4}$$

where $c_i \triangleq$ Tr $(F_i\tau)$ and $b_i \triangleq$ Tr $(F_i\sigma)$. Let $a_i \triangleq$ Tr $(F_i\rho)$. Then,

$$0 < \sqrt{p} \leq |\langle\psi|\theta\rangle| \leq B(\tau, \rho) \leq \sum_{i=1}^l \sqrt{c_i a_i}.$$

Note that the a_i's are non-negative real numbers summing up to 1, and so are the b_i's and the c_i's. These follow from Fuchs and Caves's characterisation of fidelity [FC95].

Define the set S as $S \stackrel{\Delta}{=} \left\{ i \in [l] : a_i > b_i 2^{4k/B(\tau,\rho)^2} \right\}$, where $k \stackrel{\Delta}{=} D(\rho\|\sigma)$. Note that $\forall i \in S, b_i \neq 0$ as $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$, k being finite. Define the POVM element G on $\mathcal{H}$ as $G \stackrel{\Delta}{=} \sum_{i \in S} F_i$. Let $a \stackrel{\Delta}{=} \mathrm{Tr}\,(G\rho)$ and $b \stackrel{\Delta}{=} \mathrm{Tr}\,(G\sigma)$. Then $a = \sum_{i\in S} a_i$, $b = \sum_{i \in S} b_i$, $b > 0$ and $a > b\, 2^{4k/B(\tau,\rho)^2}$. We have that

$$D(\rho\|\sigma) = k \geq a \log \frac{a}{b} > \frac{4ka}{B(\tau,\rho)^2} \Rightarrow a < \frac{B(\tau,\rho)^2}{4}.$$

Now, using the Cauchy-Schwarz inequality one can now check that $B(\tau,\rho) < \frac{B(\tau,\rho)}{2} + 2^{2k/B(\tau,\rho)^2} B(\tau,\sigma)$. Since $k' = 4k+2$, we can now conclude that

$$|\langle\phi|\theta\rangle|^2 \geq \frac{|\langle\psi|\theta\rangle|^2}{2^{k'/|\langle\psi|\theta\rangle|^2}},$$

completing the proof of the lemma. ■

We now prove a lemma which, roughly speaking, removes the dependence on F in the above lemma.

Lemma 4 *Consider two finite dimensional Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, $\dim(\mathcal{K}) \geq \dim(\mathcal{H})$. Let ρ, σ be density matrices in $\mathcal{H}$ and $|\psi\rangle$ be a purification of ρ in $\mathcal{H} \otimes \mathcal{K}$. Let $0 \leq p \leq 1$. There exists a density matrix ω in $\mathcal{H} \otimes \mathcal{K}$ such that $\mathrm{Tr}_{\mathcal{K}}\,\omega = \sigma$, and for all POVM elements F on $\mathcal{H} \otimes \mathcal{K}$ such that $\mathrm{Tr}\,(F|\psi\rangle\langle\psi|) \geq p$, $\mathrm{Tr}\,(F\omega) \geq p/2^{k'/p}$, where $k' \stackrel{\Delta}{=} 4D(\rho\|\sigma) + 2$.*

Proof: **(Sketch)** We assume without loss of generality that $0 < D(\rho\|\sigma) < +\infty$ and that $p > 0$. Consider the set A_1 of all extensions ω of σ in $\mathcal{H} \otimes \mathcal{K}$ i.e. $\mathrm{Tr}_{\mathcal{K}}\,\omega = \sigma$. A_1 is a non-empty, compact, convex set. Consider the set A_2 of all POVM operators F in $\mathcal{H} \otimes \mathcal{K}$ such that $\mathrm{Tr}\,(F|\psi\rangle\langle\psi|) \geq p$. A_2 is a compact convex set. Without loss of generality, A_2 is non-empty. Let $u(\omega, F) \stackrel{\Delta}{=} \mathrm{Tr}\,(\omega F)$. The lemma now follows from Fact 3 (note that we think of our matrices, which in general have complex entries, as vectors in a larger real vector space). ■

The previous lemma depends on the parameter p. We now remove this restriction, to get an 'observational divergence lifting' theorem.

Theorem 1 (Lifting of observational divergence) *Consider two finite dimensional Hilbert spaces $\mathcal{H}, \mathcal{K}$, $\dim(\mathcal{K}) \geq \dim(\mathcal{H})$. Let ρ, σ be density matrices in $\mathcal{H}$. Let $|\psi\rangle$ be a purification of ρ in $\mathcal{H} \otimes \mathcal{K}$. Then there exists a density matrix ω in $\mathcal{H} \otimes \mathcal{K}$ such that $\mathrm{Tr}_{\mathcal{K}}\,\omega = \sigma$ and $D((|\psi\rangle\langle\psi|)\,\|\omega) < 8D(\rho\|\sigma) + 6$.*

Proof: **(Sketch)** Follows from Lemma by a "discrete integration" argument with respect to parameter p. ■

We are now finally in a position to prove the Substate Theorem.

Theorem 2 (Substate Theorem) *Consider two finite dimensional Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, $\dim(\mathcal{K}) \geq \dim(\mathcal{H})$. Let $\mathbb{C}^2$ denote the two dimensional complex Hilbert space. Let ρ, σ be density matrices in $\mathcal{H}$. Let $r > 1$ be any real number. Let $|\psi\rangle$ be a purification of ρ in $\mathcal{H} \otimes \mathcal{K}$. Then there exist pure states $|\phi\rangle, |\theta\rangle \in \mathcal{H} \otimes \mathcal{K}$ (depending on r) and $|\zeta\rangle \in \mathcal{H} \otimes \mathcal{K} \otimes \mathbb{C}^2$ such that $|\zeta\rangle$ is a purification of σ and $\||\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|\|_t \leq 4/\sqrt{r}$, where*

$$|\zeta\rangle \stackrel{\Delta}{=} \sqrt{\frac{r-1}{r2^{rk}}}\,|\phi\rangle|0\rangle + \sqrt{1 - \frac{r-1}{r2^{rk}}}\,|\theta\rangle|1\rangle$$

and $k \stackrel{\Delta}{=} 8S(\rho\|\sigma) + 14$.

Proof: **(Sketch)** Follows from Lemma 2 and Theorem 1. ■

Remarks:
1. Note that Result 4 in the introduction follows from above by tracing out $\mathcal{K} \otimes \mathbb{C}^2$ and monotonicity of trace distance.
2. From Result 4, one can easily see that $\|\rho - \sigma\|_t \leq 2 - 2^{-O(k)}$. This implies a $2^{-O(k)}$ lower bound on the fidelity of ρ and σ.

4 Pointer chasing

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

4.1 The pointer chasing problem P_k

The input: Alice's input is a function $F_A : V_A \to V_B$. Bob's input is a function $F_B : V_B \to V_A$. 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 $F \stackrel{\Delta}{=} F_A \cup F_B$; let $\mathsf{ans} \stackrel{\Delta}{=} F^{(k+1)}(s)$. We assume that vertices in V_A and V_B have binary encodings of length $\log n$.
The communication: Alice and Bob exchange messages $M_1, \ldots, M_k$, having lengths $c_1 n, \ldots, c_k n$, 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 on measuring Ans in the computational basis[1] the answer obtained should be equal to ans with probability at least $\frac{3}{4}$, for all F_A, F_B.

4.2 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 hy-

[1] From now on, all measurements are to be performed using the computational basis.

pothesis 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_ℓ^A implies $Q_{\ell-1}^B$ and Q_ℓ^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 $Q_k^A(c_1, \ldots, c_k, n_a, n_b, \epsilon)$ 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 $Q_k^A(c_1, \ldots, c_k, n_a, n_b, \epsilon)$ holds if there is a quantum protocol of the following form.

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 by applying a unitary transformation U_B on the registers M_1, $F_B[s]$, $(F_B[b] : b \in V_B - \{s\})$ and 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.

For our predicate $Q_k^A(c_1, \ldots, c_k, n_a, n_b, \epsilon)$ 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.

Communication: After U_A, U_B have been applied, Alice and Bob follow a quantum protocol exchanging further messages $M_2, \ldots, M_k$ of lengths $c_2 n, \ldots, c_k n$. 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: Let ans denote the value observed in Ans at the end of the protocol, 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 $f \triangleq f_A \cup f_B$.
Requirement 2: $\Pr[\mathsf{ans} = f^{(k+1)}(s)] \geq \epsilon$.

Lemma 5 *If there is a safe quantum protocol for P_k^A with $v_0 = s \in V_B$, messages of lengths $c_1 n, \ldots, c_k n$, and worst case error at most $\frac{1}{4}$, then $Q_k^A(c_1, \ldots, c_k, n_A = 0, n_B = 0, \frac{3}{4})$ is true.*

Proof: Ommited, easy to check. ■

Lemma 6 *If there is a safe quantum protocol for $Q_1^A(c_1, n_A, n_B, \epsilon)$ (with $n_A < n$), then $\epsilon^{-1} 2^{c_1/n} \geq n$.*

Proof: Omitted, easy to check. ■

The following lemma is the key to our inductive argument.

Lemma 7 (Round elimination) *For $k \geq 2$, if $Q_k^A(c_1, \ldots, c_k, n_A, n_B, \epsilon)$ holds (with $n_A < n$) then $Q_{k-1}^B(c_1 + c_2, c_3, \ldots, c_k, n_A, n_B + 1, \epsilon')$ holds with $\epsilon' = \frac{5n\epsilon/(n-a)}{8*2^{(256/(n\epsilon/n-a)^2)(8nc_1/(n-a)+14)}}$*

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

Theorem 3 *Suppose $k \leq n^{\frac{1}{4}}$ and $Q_k^A(c_1, \ldots, c_k, 0, 0, \frac{1}{4})$ holds. Then $c_1 + c_2 + \cdots + c_k = \Omega(\log^{(k)} n)$.*

Proof: Follows from Lemma 7 and Lemma 6. ■

Now, by using Lemma 5, 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 $\Omega(n \log^{(k)} n)$ qubits.*

5 Round elimination: proof of Lemma 7

We consider Part (a) first. Part (b) follows using similar argument, and we do not describe them explicitly. Suppose $Q_k^A(c_1, c_2, \ldots, c_k, n_A, n_B, \epsilon)$ is true and let protocol $\mathcal{P}$ satisfy the requirements.

In what follows, subscripts of pure and mixed states will denote the registers which are in those states. 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

$$|\psi_{\text{in}}\rangle = |\psi^A\rangle \otimes \frac{1}{\sqrt{n}} \sum_{a \in V_A} |a\rangle_s |\mathbf{0}\rangle_R, \tag{5}$$

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

$$|\psi^A\rangle = \frac{1}{\sqrt{\ell_a}} \sum_{b \in V_B} |b\rangle_a |\psi^A_{a \to b}\rangle, \tag{6}$$

where $\ell_a = 1$ if $a \in X_A$ and $\ell_a = n$ otherwise. (If $\Pr[f_A[a] = b] = 0$, then $|\psi^A_{a \to b}\rangle \triangleq 0$.) From (5) and (6), we have

$$|\psi_{\text{in}}\rangle = \frac{1}{\sqrt{n}} \sum_{a \in V_A} \frac{1}{\sqrt{\ell_a}} \sum_{b \in V_B} |b\rangle_a |\psi^A_{a \to b}\rangle |a\rangle_s |\mathbf{0}\rangle_R \tag{7}$$

At this point the first message M_1 is sent to Bob. Let the rest of the protocol starting from this point be $\mathcal{P}'$.

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

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

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

$$\epsilon = \mathop{\mathrm{E}}_{a,b}[\epsilon_{a \to b}] \geq \frac{n - n_a}{n} \mathop{\mathrm{E}}_{a \in_u V_A - X_A, b \in_u V_B} [\epsilon_{a \to b}]. \tag{8}$$

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.

Let $(M_1, \widetilde{M_1})$ be the canonical purification of the first message of the protocol $\mathcal{P}$. Suppose $S(M_{1,a \to b} \| M_1) \triangleq \delta_{a \to b}$. Then, by the Substate Theorem, Theorem 2, there exists a unitary transformation $U_{a \to b}$ and a measurement $G_{a \to b}$, 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 $\tilde{\psi}^A_{a \to b}$ (with probability $\widehat{\delta_{a \to b}} \triangleq \frac{r-1}{r 2^{r(8\delta_{a \to b} + 14)}}$, where $r \triangleq 256/\epsilon^2_{a \to b}$) such that

$$\left\| |\psi^A_{a \to b}\rangle\langle\psi^A_{a \to b}| - |\tilde{\psi}^A_{a \to b}\rangle\langle\tilde{\psi}^A_{a \to b}| \right\|_t \leq 4/\sqrt{r}. \tag{9}$$

In particular, if the protocol $\mathcal{P}'$ is run starting from the state $|\tilde{\psi}^A_{a \to b}\rangle |\mu\rangle_s |\mathbf{0}\rangle_R$ (instead of $|\psi^A_{a \to b}\rangle |\mu\rangle_s |\mathbf{0}\rangle_R$), the probability of success is at least $\epsilon_{a \to b} - \epsilon_{a \to b}/4$.

5.1 The protocol $\mathcal{P}_{a \to b}$

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 .

Step 1: Alice generates the canonical purification $(M_1, \widetilde{M_1})$. Alice applies $U_{a \to b}$ and a Measurement $G_{a \to b}$ to $\widetilde{M_1}$ (plus some ancilla) to produce the state $|\tilde{\psi}^A_{a \to b}\rangle$ in the registers (M_1, F_A, W_A). She succeeds with probability $\widehat{\delta_{a \to b}}$.

Step 2: Alice and Bob proceed according to the protocol $\mathcal{P}'$ starting from the state $|\tilde{\psi}_{a \to b}\rangle = |\tilde{\psi}^A_{a \to b}\rangle |a\rangle_s |\mathbf{0}\rangle_R$, 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 .

Remark on the inputs generated: Let $f'_{A,a \to b}$ be the random variable with distribution $D'_{a \to b}$, resulting on measuring F_A after $U_{a \to b}$ has been applied. Let $f_{A,a \to b}$ be the random variable with distribution D, resulting on measuring F_A in $|\psi^A_{a \to b}\rangle$. Then, it follows from (**??**) and Theorem **??** that

$$\|D_{a \to b} - D'_{a \to b}\|_1 \leq 4/\sqrt{r}. \tag{10}$$

In $\mathcal{P}_{a \to b}$, Bob's input registers continue to satisfy the following requirements:

B1. $f_{B,a \to b}$ is constant on $X_A \cup \{s\}$ (in fact, $f_B[s] = a$), and

B2. the set of random variables $(f_{B,a \to b}[v] : v \in V_B - X_B - \{s\})$ are independent and uniformly distributed over V_A.

Probability of success in $\mathcal{P}_{a \to b}$: By (**??**) and Theorem **??**, the probability of success of $\mathcal{P}_{a \to b}$, which we denote by $\tilde{\epsilon}_{a \to b}$, is at least $\widehat{\delta_{a \to b}}(\epsilon_{a \to b} - \epsilon_{a \to b}/4)$.

5.2 Revised Protocol $\mathcal{P}_{a \to b}$

Step 1: Alice does the **Step1** as in $\mathcal{P}_{a \to b}$ followed by correction of the input registers as follows:
Correcting Alice's input registers: Let $C_{a \to b}$ be the unitary transformation corresponding to $D'_{a \to b}$ and $D_{a \to b}$ according to Fact 2. 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

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

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 $(f'_{A,a\to b}, \hat{f}_{A,a\to b})$ have distribution precisely $D'_{a\to b}$ and $D_{a\to b}$. Furthermore, (see Fact 2),

$$\Pr[f'_{A,a\to b} \neq \hat{f}_{A,a\to b}] \leq \frac{1}{2} \cdot 4/\sqrt{r} = \epsilon_{a\to b}/8. \qquad (11)$$

Step 2: From this point on, Alice and Bob just follow $\mathcal{P}'$ Let $|\phi_{a\to b}\rangle$ denote the state of the entire system just after M_2 is sent to Alice. 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.

Success probability in revised $\mathcal{P}_{a\to b}$: Let $\hat{\epsilon}_{a\to b}$ be the success probability of the revised protocol. It is easy to check that :

$$\hat{\epsilon}_{a\to b} \quad \geq \quad \widehat{\delta_{a\to b}}(\epsilon_{a\to b} - \epsilon_{a\to b}/4 - \epsilon_{a\to b}/8) \qquad (12)$$

5.3 The final protocol: $\mathcal{P}_a$

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

$$|b\rangle_{\hat{F}[a]}|\theta\rangle_{\widetilde{M_1},A} \mapsto |b\rangle_{\hat{F}[a]}U_{a\to b}|\theta\rangle_{\widetilde{M_1},A}.$$

Note that this transformation is safe on $\hat{F}[a]$. Then he measures the register $\hat{F}_A[a]$ and then performs the measurement $G_{a\to b}$.

(b) Alice applies the correction used in the revised Step 1 of $\mathcal{P}_{a\to b}$. After this $\hat{F}_A$ are to be treated as A's input registers.

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.

Note 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, we will check if the answer ans' agrees with the answer $\mathsf{ans}(\hat{f}_A, f_B)$, where $\hat{f}_A$ is the random variable obtained by measuring the new input registers $\hat{F}_A$.

The probability of success of $\mathcal{P}_a$: For $a \in V_A - X_A$, let $\hat{\epsilon}_a$ be the probability of success of $\mathcal{P}_a$. Then, by (12), we have

$$\begin{aligned} \hat{\epsilon}_a \quad &= \quad \mathop{\mathrm{E}}_{b\in_u V_B}[\hat{\epsilon}_{a\to b}] & (13) \\ &\geq \quad \mathop{\mathrm{E}}_{b\in_u V_B}[\frac{r-1}{r2^{r(8\delta_{a\to b}+14)}}(5\epsilon_{a\to b}/8)] & (14) \end{aligned}$$

Since the function inside the expectation is jointly convex in $\epsilon_{a\to b}$ and $\delta_{a\to b}$ and using Fact 1, we conclude

$$\hat{\epsilon}_a \geq \frac{5\epsilon_a}{8 * 2^{(256/\epsilon_a^2)(8I(f_A[a]:M_1)+14)}} \qquad (15)$$

where $\epsilon_a = \mathrm{E}_{b\in_u V_B}[\epsilon_{a\to b}]$.

Claim 1 $\mathrm{E}_{a\in_u V_A - X_A}[I(f_A[a] : M_1)] \leq \left(\frac{n}{n-n_a}\right) c_1$.

Proof: Using Fact **??** and (10), we have $c_1 n \geq I(f_A : M_1) \geq \sum_{a\in V_A} I(f_A[a] : M_1) \geq \sum_{a\in V_A - X_A} I(f_A[a] : M_1)$. ■

Now again since the above function is jointly convex in ϵ_a and $I(f_A[a])$ we conclude from (15) and above claim that :

$$\mathop{\mathrm{E}}_{a\in_u V_A - X_A}[\hat{\epsilon}_a] \;\geq\; \frac{5n\epsilon/(n-a)}{8 * 2^{(256/(n\epsilon/n-a)^2)(8nc_1/(n-a)+14)}}, \qquad (16)$$

where on the right a is chosen uniformly from $V_A - X_A$ and b is chosen independently and uniformly from V_B.

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

$$\hat{\epsilon}_a \geq \frac{5n\epsilon/(n-a)}{8 * 2^{(256/(n\epsilon/n-a)^2)(8nc_1/(n-a)+14)}}.$$

Now, it can be verified, that the protocol $\mathcal{P}_a$ satisfies the requirements for $Q^B_{k-1}(c_1 + c_2, c_3, \ldots, c_k, n_A, n_B + 1, \hat{\epsilon}_a)$. This shows Lemma 7.

Acknowledgments

We are very grateful to Ashwin Nayak for his contribution to this work. He patiently went through several versions of our proofs; his counter examples and insights were invaluable in arriving at our definition of privacy. We are grateful to K R Parthasarathy and Rajendra Bhatia for sharing with us their insights in operator theory.

References

[ANTV99] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani. Dense quantum coding and a lower bound for 1-way quantum automata. In *Proceedings of the 31st Annual ACM Symposium on Theory of Computing*, pages 376–383, 1999.

[CT91] T. Cover and J. Thomas. *Elements of Information Theory*. Wiley Series in Telecommunications. John Wiley and Sons, 1991.

[CvDNT98] R. Cleve, W. van Dam, M. Nielsen, and A. Tapp. Quantum entanglement and the communication complexity of the inner product function. In *Proceedings of the 1st NASA International Conference on Quantum Computing and Quantum Communications*, Lecture Notes in Computer Science, vol. 1509, pages 61–74. Springer-Verlag, 1998. Also quant-ph/9708019.

[DGS87] P. Duris, Z. Galil, and G. Schnitger. Lower bounds on communication complexity. *Information and Computation*, 73:1–22, 1987.

[DJS98] C. Damm, S. Jukna, and J. Sgall. Some bounds on multiparty communication complexity of pointer jumping. *Computational Complexity*, 7:109–127, 1998.

[FC95] C. Fuchs and C. Caves. Mathematical techniques for quantum communication theory. *Open Systems and Information Dynamics*, 3(3):345–356, 1995. Also quant-ph/9604001.

[Joz94] R. Jozsa. Fidelity for mixed quantum states. *Journal of Modern Optics*, 41(12):2315–2323, 1994.

[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.

[Kla02] H. Klauck. On quantum and approximate privacy. In *Proceedings of the 19th Annual Symposium on Theoretical Aspects of Computer Science*, Lecture Notes in Computer Science, vol. 2285, pages 335–346. Springer-Verlag, 2002. Also quant-ph/0110038.

[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.

[Nay99] A. Nayak. Optimal lower bounds for quantum automata and random access codes. In *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science*, pages 369–377, 1999.

[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.

[OR94] M. Osborne and A. Rubinstein. *A course in game theory*. MIT Press, 1994.

[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.

[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.