# Optimal Direct Sum and Privacy Trade-off Results for Quantum and Classical Communication Complexity

## [Full version]*

Rahul Jain[1]**, Jaikumar Radhakrishnan[2]***, and Pranab Sen[3]†

[1] University of Waterloo, Waterloo, ON, Canada, N2L 3G1. rjain@cs.uwaterloo.ca
[2] Tata Institute of Fundamental Research, Mumbai, India. jaikumar@tifr.res.in
[3] Tata Institute of Fundamental Research, Mumbai, India. pgdsen@tcs.tifr.res.in

**Abstract.** We show optimal *Direct Sum* result for the *one-way entanglement-assisted quantum communication complexity* for any relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. We show:

$$\mathsf{Q}^{1,\mathsf{pub}}(f^{\oplus m}) = \Omega(m \cdot \mathsf{Q}^{1,\mathsf{pub}}(f)),$$

where $\mathsf{Q}^{1,\mathsf{pub}}(f)$, represents the one-way entanglement-assisted quantum communication complexity of $f$ with error at most $1/3$ and $f^{\oplus m}$ represents $m$-copies of $f$. Similarly for the *one-way public-coin classical communication complexity* we show:

$$\mathsf{R}^{1,\mathsf{pub}}(f^{\oplus m}) = \Omega(m \cdot \mathsf{R}^{1,\mathsf{pub}}(f)),$$

where $\mathsf{R}^{1,\mathsf{pub}}(f)$, represents the one-way public-coin classical communication complexity of $f$ with error at most $1/3$. We show similar optimal Direct Sum results for the *Simultaneous Message Passing* (SMP) quantum and classical models. For two-party two-way protocols we present optimal *Privacy Trade-off* results leading to a Weak Direct Sum result for such protocols.

We show our Direct Sum and Privacy Trade-off results via *message compression* arguments. These arguments also imply a new *round elimination* lemma in quantum communication, which allows us to extend classical lower bounds on the *cell probe complexity* of some *data structure problems*, e.g. *Approximate Nearest Neighbor Searching* (ANN) on the Hamming cube $\{0,1\}^n$ and *Predecessor Search* to the quantum setting.

In a separate result we show that Newman's [New91] technique of reducing the number of public-coins in a classical protocol cannot be lifted to the quantum setting. We do this by defining a general notion of *black-box* reduction of prior entanglement that subsumes Newman's technique. We prove that such a black-box reduction is impossible for quantum protocols by exhibiting a particular one-round quantum protocol for the *Equality* function where the black-box technique fails to reduce the amount of prior entanglement by more than a constant factor.

In the final result in the theme of message compression, we provide an upper bound on the problem of *Exact Remote State Preparation* (ERSP).

## 1 Introduction

Communication complexity studies the communication required to solve a computational problem in a distributed setting. Consider a relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. In a two-party protocol to solve $f$,

* A preliminary version of this paper with many results mentioned here appeared in the 20th IEEE Conference on Computational Complexity, 2005.
** Research supported in part by ARO/NSA USA. Part of this work was done while the author was at U.C. Berkeley, Berkeley, USA, where it was supported by Army Research Office (ARO), North California, under grant DAAD 19-03-1-00082. Part of the work done while the author was at Tata Institute of Fundamental Research, Mumbai, India.
*** Part of the work done while the author was at Toyota Technological Institute Chicago, USA.
† Work done while the author was at University of Waterloo, Waterloo, Canada.

one party say Alice would be given input $x \in \mathcal{X}$, and the other party say Bob would be given input $y \in \mathcal{Y}$. The goal for Alice and Bob would be to communicate and find an element $z \in \mathcal{Z}$ that satisfies the relation, i.e., to find a $z$ such that $(x, y, z) \in f$. The protocols they follow could be deterministic, randomized or quantum leading to different notions of deterministic, randomized and quantum communication complexity. Please refer to Sec. 2.2 for detailed exposition to various models, definitions and notations related to classical and quantum communication complexity.

### 1.1 Direct Sum

Let us consider a natural question in communication complexity as follows. Suppose Alice and Bob wish to solve several, say $k$, instances of relation $f$ simultaneously, with constant success on the overall output. A *Direct Sum* theorem states that the communication required for accomplishing this would be at least $k$ time the communication required for solving single instance of $f$, with constant success. It is a natural and fundamental question in communication complexity.

Although they seem highly plausible, it is well-known that Direct Sum results fail to hold for some settings of communication. For example for the *Equality* function ($\mathrm{EQ}_n$), in which Alice and Bob need to determine if their $n$-bit inputs are equal or not, its randomized private-coins communication complexity, denoted $\mathsf{R}(\mathrm{EQ}_n)$ does not satisfy the Direct Sum property. It is known that $\mathsf{R}(\mathrm{EQ}_n) = \Theta(\log n)$ whereas for testing Equality of $k = \log n$ [4] pairs of $n$-bit strings $\mathsf{R}(\mathrm{EQ}_n^{\oplus k}) = O(k \log k + \log n) = O(\log n \log \log n)$ (see, e.g., [KN97, Example 4.3, page 43]), where we might expect $\mathsf{R}(\mathrm{EQ}_n^{\oplus k}) = \Omega(k \log n) = \Omega(\log^2 n)$. Similarly, Shaltiel [Sha03] gives an example for which a related notion called the *Strong Direct Product* property fails to hold for average case (i.e., distributional) communication complexity. (A Strong Direct Product theorem would show that even with probability of success that is exponentially small in $k$, the cost of solving $k$ instances of $f$, would be $k$ times the cost of solving one instance.)

**Previous works:** Notwithstanding these examples, Direct Sum results have met with some success in several settings of communication. It is straightforward to show that $\mathsf{D}^1(f)$ : the deterministic one-way communication complexity of every relation $f$, satisfies the Direct Sum property. It is also known that for two-way protocols, for any function $f$, $\mathsf{D}(f^{\oplus k}) = \Omega(k \cdot \sqrt{\mathsf{D}(f)})$ (see, e.g., [KN97, Exercise 4.11, page 46]). For classical distributional complexity, under the uniform distribution on the inputs, Chakrabarti, Shi, Wirth, and Yao [CSWY01] showed Direct Sum in the one-way and SMP models of communication. They introduced an important notion of *information cost* and obtained their Direct Sum result via a message compression argument. The notion of information cost has also been effectively used to obtain two-way classical and quantum communication complexity bounds for example see [BYJKS04,JRS03b]. Jain, Radhakrishnan, and Sen [JRS03a] extended the result of [CSWY01], and provided a Direct Sum result for classical distributional complexity under any product distribution on inputs, for bounded-round two-way protocols. They [JRS03a] again used the information cost approach however achieved their message compression via different techniques (than [CSWY01]) involving the *Substate Theorem* [JRS02]. Recently, Harsha, Jain, McAllester, and Radhakrishnan [HJMR07] have strengthened the Direct Sum result of [JRS03a] by reducing to a large extent its dependence on the number of rounds. The message compression results in [JRS03a,HJMR07] have been used in the work of Chakrabarti and Regev [CR04] to show lower bounds on the Approximate Nearest Neighbor problem (ANN) in the cell probe model.

[4] All logarithms in this article are taken to base 2 unless otherwise specified.

Pătraşcu and Thorup [PT06b] also use Direct Sum type results to prove better lower bounds for this problem.

**Our results:** In this paper we prove that for any relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, the classical public-coin one-way communication complexity $\mathsf{R}^{1,\mathsf{pub}}(f)$ and the one-way entanglement assisted quantum communication complexity $\mathsf{Q}^{1,\mathsf{pub}}(f)$ satisfy the Direct Sum property. Similarly in the SMP model $\mathsf{R}^{\|,\mathsf{pub}}(f)$ and $\mathsf{Q}^{\|,\mathsf{pub}}(f)$ satisfy the Direct Sum property. Our precise results are as follows.

**Theorem 1 (Direct Sum).** *Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation. Let $\epsilon, \delta \in (0, 1/2)$ with $\epsilon + \delta < 1/2$. For one-round protocols we have:*

1. $\mathsf{R}^{1,\mathsf{pub}}_{\epsilon}(f^{\oplus m}) \geq \Omega\left(\delta^3 m \cdot \mathsf{R}^{1,\mathsf{pub}}_{\epsilon+\delta}(f)\right)$.
2. $\mathsf{Q}^{1,\mathsf{pub}}_{\epsilon}(f^{\oplus m}) \geq \Omega\left(\delta^3 m \cdot \mathsf{Q}^{1,\mathsf{pub}}_{\epsilon+\delta}(f)\right)$.

*Similarly for SMP protocols (with shared resource as specified in Section 2.2), we have:*

1. $\mathsf{R}^{\|,\mathsf{pub}}_{\epsilon}(f^{\oplus m}) \geq \Omega\left(\delta^3 m \cdot \mathsf{R}^{\|,\mathsf{pub}}_{\epsilon+\delta}(f)\right)$.
2. $\mathsf{Q}^{\|,\mathsf{pub}}_{\epsilon}(f^{\oplus m}) \geq \Omega\left(\delta^3 m \cdot \mathsf{Q}^{\|,\mathsf{pub}}_{\epsilon+\delta}(f)\right)$.

We obtain our Direct Sum results via message compression results. Our message compression result for classical one-way protocols is as follows:

**Result 1 (Classical one-way message compression, informal statement)** *Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation and let $\mu$ be a probability distribution (possibly non-product) on $\mathcal{X} \times \mathcal{Y}$. Let $\mathcal{P}$ be a one-way private-coins classical protocol for $f$ (with single message from Alice to Bob) having bounded average error under $\mu$. Suppose Alice's message in $\mathcal{P}$ has* mutual information *(please refer to Sec. 2 for definition) at most $k$ about her input. Then there is a one-way deterministic protocol $\mathcal{P}'$ for $f$ having similar average error probability under $\mu$, in which Alice's message is $O(k)$ bits long.*

We show similar message compression result for one-way quantum protocols.

**Result 2 (Quantum one-way message compression, informal statement)** *Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation and let $\mu$ be a probability distribution (possibly non-product) on $\mathcal{X} \times \mathcal{Y}$. Let $\mathcal{P}$ be a one-way quantum protocol without prior entanglement for $f$ having bounded average error probability under $\mu$. Suppose Alice's message in $\mathcal{P}$ has mutual information at most $k$ about her input. Then there is a one-way protocol $\mathcal{P}'$ for $f$ with prior entanglement having similar average error probability under $\mu$, where Alice's message is classical and $O(k)$ bits long.*

The proof of the above result uses a technical quantum information-theoretic fact called the Substate Theorem [JRS02]. Essentially, it says that if a quantum encoding of a classical random variable $x \mapsto \sigma_x$ has information at most $k$ about $x$, then for most $x$, $\frac{\sigma_x}{2^{O(k)}} \leq \sigma$ (for Hermitian matrices $A$, $B$, $A \leq B$ is a shorthand for the statement "$B - A$ is positive semidefinite"), where $\sigma \stackrel{\mathsf{def}}{=} \mathbb{E}_x[\sigma_x]$. Similarly the classical message compression result uses the classical version of the Substate Theorem. The classical Substate Theorem was also used by [JRS03a] to prove their classical message compression results.

Res. 2 also allows us to prove a new *round elimination* result for quantum communication. To state the round elimination lemma, we first need the following definition.

**Definition 1.** *Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a function. The communication game $f^{(k),A}$ is defined as follows:* Alice *gets $k$ strings $x_1, \ldots, x_k \in \mathcal{X}$.* Bob *gets an integer $j \in [k]$, a copy of strings $x_1, \ldots, x_{j-1}$, and a string $y \in \mathcal{Y}$. They are supposed to communicate and determine $f(x_j, y)$. The communication game $f^{(k),B}$ is defined analogously with roles of* Alice *and* Bob *reversed.*

**Result 3 (Round elim., informal stmt.)** *Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a function. Suppose $\mathcal{P}$ is a $t$-round quantum protocol for $f^{(k),A}$ with prior entanglement having bounded worst case error. Suppose* Alice *starts the communication and the first and second messages of $\mathcal{P}$ are $l_1$ and $l_2$ qubits long respectively. Then there is a $(t-1)$-round protocol for $f$ with prior entanglement having similar worst case error where* Bob *starts the communication and the first message is $l_2 \cdot 2^{O(l_1/k)}$ qubits long. The subsequent communication in $\mathcal{P}'$ is similar to that in $\mathcal{P}$.*

The classical analogue of the above result was shown by Chakrabarti and Regev [CR04], where they used the message compression arguments of [JRS03a,HJMR07] to arrive at their result. The above round elimination lemma is useful in situations where Alice's message length $l_1$ is much smaller than Bob's message length $l_2$. Such a situation arises in proving cell probe lower bounds for data structure problems like ANN in $\{0,1\}^n$ and Set Predecessor. [CR04] used it crucially in proving optimal randomized cell probe lower bounds for ANN. Recently, Patrascu and Thorup [PT06a,PT06b] used the same classical technique to prove sharper lower bounds for the Set Predecessor problem. We remark that both these results carry over to the *address-only quantum cell probe model* (defined in [SV01]) as a consequence of Res. 3.

### 1.2 Privacy trade-offs

Let us consider another natural question in communication complexity as follows. Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation. We are interested in the *privacy loss* of Alice and Bob that is inherent in computing $f$. Privacy in communication complexity was studied in the classical setting by Bar-Yehuda et al. [BCKO93], and in the quantum setting by Klauck [Kla02] and Jain, Radhakrishnan, and Sen [JRS02]. For studying privacy issues in quantum communication, we only consider protocols without prior entanglement. To define the privacy loss of Alice, imagine that Alice follows the protocol $\mathcal{P}$ honestly but Bob is malicious and deviates arbitrarily from $\mathcal{P}$ in order to extract the maximum amount of information about Alice's input. The only constraint on Bob is that Alice should not be able to figure out at any point of time whether he is cheating or not; we call such a cheating strategy of Bob *undetectable*. Suppose $\mu = \mu_{\mathcal{X}} \times \mu_{\mathcal{Y}}$ is a product probability distribution on $\mathcal{X} \times \mathcal{Y}$. Let register $X$ denote the input qubits of Alice, and $B$ denote all the qubits in the possession of Bob at the end of $\mathcal{P}$. We assume the input registers of Alice and Bob are never modified and are never sent as messages in $\mathcal{P}$. Then the privacy loss of Alice under distribution $\mu$ at the end of $\mathcal{P}$ is the maximum mutual information $I(X : B)$ over all undetectable cheating strategies of Bob. The privacy loss of Bob can be defined analogously. In the quantum setting Bob has a big bag of undetectable cheating tricks that he can use in order to extract information about $X$. For instance, he can start the protocol $\mathcal{P}$ by placing a superposition of states $|\mu_{\mathcal{Y}}\rangle$ (for a probability distribution $\pi$ on $\mathcal{Z}$, $|\pi\rangle \stackrel{\text{def}}{=} \sum_z \sqrt{\pi(z)}|z\rangle$) in his input register $Y$ and running the rest of the protocol honestly. This trick works especially well for so-called *clean* protocols that leave the work qubits of Alice and Bob at the end of the protocol in the state $|0\rangle$. For example, consider the following exact clean protocol $\mathcal{P}$ computing the inner product modulo 2, $x \cdot y$, of two bit strings $x, y \in \{0,1\}^n$: Alice sends her input $x$ to Bob, Bob computes $x \cdot y$ and sends back $x$ to

Alice keeping the bit $x \cdot y$ with himself, and finally Alice zeroes out Bob's message by XORing with her input $x$. If Bob does the above 'superposition cheating' trick for $\mathcal{P}$, his final state at the end of $\mathcal{P}$ becomes $\left(\sum_{y \in \{0,1\}^n} |y, x \cdot y\rangle\right)$. It is easy to see that Bob has $\frac{n}{2}$ bits of information about $x$, if $x$ is distributed uniformly in $\{0,1\}^n$. Thus, the privacy loss from Alice to Bob for this protocol is at least $\frac{n}{2}$, under the uniform distribution on $\{0,1\}^n \times \{0,1\}^n$. See [CvDNT98] for more details. Thus, it is conceivable that Alice and Bob use an 'unclean' protocol to compute $f$ in order to minimize their privacy losses. We shall be concerned with proving tradeoffs between the privacy losses of Alice and Bob for any quantum protocol computing $f$, including 'unclean' ones. Please refer to Sec. 4, Def. 6 for precise definition of privacy loss. Note that defining the privacy loss only for quantum protocols without prior entanglement is without loss of generality, since we can convert a protocol with prior entanglement into one without prior entanglement by sending the entanglement as part of the first message of the protocol; this process does not affect the privacy loss since after the first message is sent, the qubits in the possession of Alice and Bob are exactly the same as before.

For private-coin randomized classical protocols, a related notion called *information cost*, was defined in [CSWY01,BYJKS04] to be the mutual information $I(XY : M)$ between the players' inputs and the complete message transcript $M$ of the protocol. For quantum protocols there is no clear notion of a message transcript, hence we use our definition of privacy instead. Also, other than cryptographic reasons there is also another reason why we allow the players to use undetectable cheating strategies. In the above clean protocol $\mathcal{P}$ for the inner product function, if both Alice and Bob were honest the final state of $\mathcal{P}$ would be $|x\rangle \otimes |y, x \cdot y\rangle$, where the first state belongs to Alice and the second to Bob. Under the uniform distribution on $x, y$ the privacy loss from Alice to Bob is 1, whereas the classical information cost is at least $n$. This shows that in the quantum setting, because of the ability of players to 'forget' information by uncomputing, it is better to allow undetectable cheating strategies for players in the definition of privacy loss in order to bypass examples such as the above.

**Our results:** In this paper we relate the privacy loss incurred in computation of any relation $f$ to the one-way communication complexity $f$. We show that in multi-round protocols with low privacy loss, all the messages could be replaced by a single short message. For quantum protocols, again using the Substate Theorem [JRS02], we prove the following result.

**Result 4 (Quantum multiple rounds compression, informal stmt.)** *Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation and let $\mu$ be a product probability distribution on $\mathcal{X} \times \mathcal{Y}$. Let $\mathcal{P}$ be a multi-round two-way quantum protocol without prior entanglement for $f$ having bounded average error probability under $\mu$. Let $k_a$, $k_b$ denote the privacy losses of* Alice *and* Bob *respectively under distribution $\mu$ in $\mathcal{P}$. Then there is a one-way protocol $\mathcal{P}'$ for $f$ with prior entanglement having similar average error probability under $\mu$, such that the single message of $\mathcal{P}'$ is from* Alice *to* Bob*, it is classical and $k_a 2^{O(k_b)}$ bits long. Similarly statement also holds with the roles of* Alice *and* Bob *reversed.*

We would like to remark that Res. 2 does not follow from Res. 4. Res. 2 holds for any probability distribution on $\mathcal{X} \times \mathcal{Y}$ whereas our proof of Res. 4 requires product distributions. It is open whether a similar multi-round compression result can be proved for non-product distributions for quantum protocols.

Similarly for classical protocols we show the following result. Please refer to Sec. 4, Def. 7 for precise definition of privacy loss for classical protocols.

**Result 5 (Classical multiple rounds compression, informal stmt.)** *Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation and let $\mu$ be a product probability distribution on $\mathcal{X} \times \mathcal{Y}$. Let $\mathcal{P}$ be a multi-round two-way private-coins classical protocol for $f$ having bounded average error probability under $\mu$. Let $k_a$, $k_b$ denote the privacy losses of* Alice *and* Bob *respectively under distribution $\mu$ in $\mathcal{P}$. Then there is a one-way deterministic protocol $\mathcal{P}'$ for $f$ having similar average error probability under $\mu$, such that the single message of $\mathcal{P}'$ is from* Alice *to* Bob *and is $k_a 2^{O(k_b)}$ bits long. Similarly statement also holds with the roles of* Alice *and* Bob *reversed.*

We would like to point out that the proof of this result does not follow entirely on the lines of Res. 4, essentially due to the difference in the definition between the notions of privacy loss for classical and quantum protocols. Therefore its proof is presented separately.

These message compression results immediately imply the following privacy trade-off results (similar results hold with the roles of Alice and Bob reversed.)

**Result 6 (Quantum privacy trade-off)** *Let the privacy loss of* Alice *be $k_a$ and the privacy loss of* Bob *be $k_b$ at the end of a quantum protocol without entanglement $\mathcal{P}$ for computing a relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. Then,*

$$k_a 2^{O(k_b)} \quad \geq \quad \mathsf{Q}^{1,A\to B,\mathsf{pub},[\,]}(f),$$

*where $\mathsf{Q}^{1,A\to B,\mathsf{pub},[\,]}(f)$ is the maximum over all product distributions $\mu$ on $\mathcal{X} \times \mathcal{Y}$, of the one-round quantum communication complexity (with* Alice *communicating) of $f$ with prior entanglement having bounded average error under $\mu$.*

**Result 7 (Classical privacy trade-off)** *Let the privacy loss of* Alice *be $k_a$ and the privacy loss of* Bob *be $k_b$ at the end of a classical private coins protocol $\mathcal{P}$ for computing a relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. Then,*

$$k_a 2^{O(k_b)} \quad \geq \quad \mathsf{R}^{1,A\to B,[\,]}(f),$$

*where $\mathsf{R}^{1,A\to B,[\,]}(f)$ is the maximum over all product distributions $\mu$ on $\mathcal{X} \times \mathcal{Y}$, of the one-round classical distributional communication complexity (with* Alice *communicating) of $f$ having bounded average error under $\mu$.*

*Remarks:*

1. Note that Res. 6 also shows that the privacy loss for computing $f$ is lower bounded by $\Omega(\log \mathsf{Q}^{1,\mathsf{pub},[\,]}(f))$. This latter result can be viewed as the privacy analogue of Kremer's result [Kre95] that the bounded error quantum communication complexity of $f$ is lower bounded by the logarithm of its deterministic one-round communication complexity.
2. Res. 4 and Res. 5 also allow us to show weak general Direct Sum result for quantum protocols and classical protocols as mentioned in Corr. 3 and Corr. 5 respectively in Sec. 4.
3. All these results are optimal in general as evidenced by the Index function problem [ANTV02]. In the Index function problem, Alice is given a database $x \in \{0,1\}^n$ and Bob is given an index $i \in [n]$. They have to communicate and determine $x_i$, the $i$-th bit of $x$. The one-round quantum communication complexity from Alice to Bob for this problem is $\Omega(n)$, even for bounded average error under the uniform distribution and in the presence of prior entanglement. Thus, we get the privacy tradeoff $k_a 2^{O(k_b)} = \Omega(n)$ for the Index function problem. This is optimal; consider a deterministic protocol where Bob sends the first $b$ bits of his index and Alice replies by sending all the $\frac{n}{2^b}$ bits of her database consistent with Bob's message.

4. Earlier, Jain, Radhakrishnan, and Sen [JRS02] had proved the same privacy tradeoff for the Index function problem specifically. Our general tradeoff above can be viewed as an extension of their result to all functions and relations.

### 1.3 Impossibility of black-box entanglement reduction

Let us return to the third main question we investigate in this work which appears different but is intimately related to the theme of message compression and we mention this connection later.

We know that for some quantum communication problems, presence of prior entanglement helps in reducing the communication. For example, the technique of superdense coding [BW92] allows us to often reduce the communication complexity by a multiplicative factor of 2. So a natural question that arises is how much prior entanglement is really required by a quantum protocol? For classical communication, Newman [New91] has shown that $O(\log n)$ shared random bits are sufficient for any protocol. This is tight, as evidenced by the Equality function on $\{0,1\}^n$ which requires $\Theta(\log n)$ bits with private randomness and $O(1)$ bits with shared randomness. One might hope to extend Newman's [New91] proof that a classical protocol needs only $O(\log n)$ shared random bits to the quantum setting. Newman's proof uses a Chernoff-based sampling argument on the shared random bit strings to reduce their number to $O(n)$. Moreover, the reduction is done in a black-box fashion i.e. it does not change the computation of Alice and Bob in the protocol. In the quantum setting, one might similarly hope to reduce the amount of entanglement of the prior entangled state $|\phi\rangle$ to $O(\log n)$ and leave the unitary transforms of Alice and Bob unaffected i.e. the hope is to find a black-box Newman-style prior entanglement reduction technique. We show that such a black-box reduction is impossible.

To state our result precisely, we need the following definitions.

**Definition 2 (Similar protocols).** *Two protocols $\mathcal{P}$ and $\mathcal{P}'$ with prior entanglement and outputting values in $\mathcal{Z}$ are called* similar protocols *if both use the same number of qubits and the same unitary transformations and measurements, have the same amount of communication and for all $(x,y) \in \{0,1\}^n \times \{0,1\}^n$, $\|\mathcal{P}(x,y) - \mathcal{P}'(x,y)\|_1 < 1/20$. Here, $\mathcal{P}(x,y)$, $\mathcal{P}'(x,y)$ are the probability distributions on $\mathcal{Z}$ of the output of protocol $\mathcal{P}$, $\mathcal{P}'$ on input $(x,y)$. $\mathcal{P}$, $\mathcal{P}'$ may use different quantum states as their input independent prior entanglement.*

**Definition 3 (Amt. of entanglement).** *For a bipartite pure state $|\phi\rangle_{AB}$, consider its* Schmidt decomposition, *$|\phi\rangle = \sum_{i=1}^{k} \sqrt{\lambda_i}|a_i\rangle \otimes |b_i\rangle$, where $\{a_i\}$ is an orthonormal set and so is $\{b_i\}$, $\lambda_i \geq 0$ and $\sum_i \lambda_i = 1$. The* amount of entanglement *of $|\phi\rangle_{AB}$ is defined to be $E(|\phi\rangle_{AB}) \stackrel{\mathsf{def}}{=} -\sum_i \lambda_i \log \lambda_i$. The* Schmidt rank *of $|\phi\rangle_{AB}$ is defined to be $k$.*

One might hope that the following conjecture is true.

*Conjecture 1.* For any protocol $\mathcal{P}$ for $f : \{0,1\}^n \times \{0,1\}^n \to \mathcal{Z}$ with prior entanglement, there exists a similar protocol $\mathcal{P}'$ that starts with prior entanglement $|\phi\rangle_{AB}$, $E(|\phi\rangle_{AB}) = O(\log n)$.

We prove that the above conjecture is **not** correct for quantum communication protocols.

**Result 8 (No black-box red. of prior entang.)** *Let us denote the Equality function on $n$-bit strings by $\mathrm{EQ}_n$ . There exists a one-round quantum protocol $\mathcal{P}$ for $\mathrm{EQ}_n$ with $\frac{2n}{3} + \log n + \Theta(1)$ EPR pairs of prior entanglement and communicating 4 bits, such that there is no similar protocol $\mathcal{P}'$ that starts with a prior entangled state $|\phi\rangle_{AB}$, $E(|\phi\rangle_{AB}) \leq n/600$.*

Our proof of this result follows essentially by sharpening the geometric arguments behind the proof of the 'recipient-non-invasive incompressibility' result of Jain, Radhakrishnan, and Sen[JRS03a]. Jain, Radhakrishnan, and Sen [JRS03a] showed that for classical constant round private-coin protocols with a product probability distribution on their inputs, one can compress the messages to the information cost of the protocol. Their compression technique for classical protocols was 'recipient-non-invasive' in the sense that, for one round protocols, it did not change the computation of the recipient except up to a trivial relabeling of the messages. They however also showed that such a recipient-non-invasive compression result does not hold for quantum protocols; they exhibited a one-round quantum protocol without prior entanglement for the Equality function on $n$-bit strings with constant privacy loss, where any recipient-non-invasive compression strategy cannot compress Alice's message by more than a multiplicative factor of 6! We essentially convert their "incompressibility of message" result to "incompressibility of prior-entanglement" result.

*Remarks:*

1. The above Res. 8 shows that in order to reduce prior entanglement in quantum communication, one has to look beyond black-box arguments and change the unitary transforms of Alice and Bob.
2. Recently Gavinsky [Gav08] showed that even if Alice and Bob are allowed to change their operations, there is an exponential increase that can occur in the required message length for computation of a relation in case the prior-entanglement is reduced only by a logarithmic factor. However Gavinsky measures shared entanglement with the number of qubits in the shared state between Alice and Bob, and not with the measure of entanglement as considered by us in Def. 3. Hence Res. 8, which first appeared in [JRS05], is incomparable to Gavinsky's result.

### 1.4 Exact Remote State Preparation (ERSP)

The final result we present in the theme of message compression concerns the communication complexity of the Exact Remote State Preparation (ERSP) problem. The ERSP problem is as follows. Let $E : x \to \rho_x$ be an encoding from a set $\mathcal{X}$ to the set of quantum states.
**Problem** ERSP($E$)**:**

1. Alice and Bob start with prior entanglement.
2. Alice gets $x \in \mathcal{X}$.
3. They interact at the end of which Bob should end up with $\rho_x$ in some register.

*Remark:* The adjective 'exact' signifies that we do not allow for any fidelity loss in the state that Bob should end up.

We provide the following upper bound on the communication complexity of this problem.

**Theorem 2.** *Let $E : x \to \rho_x$ be an encoding where $\rho_x$ is a pure state for all $x$ and let $\sigma$ be any state with full rank. There is a protocol $\mathcal{P}$ for* ERSP($E$) *with expected communication bounded by* $\max_x\{\log(\mathsf{Tr}\sigma^{-1}\rho_x) + 2\log\log(\mathsf{Tr}\sigma^{-1}\rho_x)\}$.

### 1.5 Organization of the paper

In the next section, we collect some preliminaries that will be required in the proofs of the message compression results. In Sec. 3, we prove our results on first round compression and round elimination

in quantum protocols. We prove our multi-round compression result in Sec. 4. In Sec. 5, we show that black-box reduction of prior entanglement in quantum communication is impossible. Finally in Sec. 6 we provide the proof of the upper bound on the ERSP problem.

## 2 Preliminaries

### 2.1 Information Theory

A quantum state is a positive semi definite trace one operator. For a quantum state $\rho$, its *von-Neumann entropy* is defined as $S(\rho) \stackrel{\mathsf{def}}{=} \sum_i -\lambda_i \log \lambda_i$, where $\lambda_i$s represent the various eigenvalues of $\rho$. For an $l$ qubit quantum system $A$, $S(A) \leq l$. For correlated quantum systems $A, B$ their mutual information is defined as $I(A:B) = S(A) + S(B) - S(AB)$. Given a tri-partite system $A, B, C$, mutual information satisfies the *monotonicity property* that is $I(A:BC) \geq I(A:B)$. Let us define $I(A:B|C) \stackrel{\mathsf{def}}{=} I(A:BC) - I(A:C)$. We have the following very useful *Chain Rule* for mutual information.

$$I(A:B_1 \dots B_k) = \sum_{i=1}^{k} I(A:B_i|B_1 \dots B_{i-1}).$$

Therefore if $B_1$ through $B_k$ are independent systems then,

$$I(A:B_1 \dots B_k) \geq \sum_{i=1}^{k} I(A:B_i).$$

For classical random variables the analogous definitions and facts hold *mutates mutandis* and we skip making explicit statements here for brevity.

### 2.2 Communication complexity

**Quantum communication complexity:** Consider a two-party quantum communication protocol $\mathcal{P}$ for computing a relation $f : \{0,1\}^n \times \{0,1\}^n \to \mathcal{Z}$. The relations we consider are always total in the sense that for every $(x,y) \in \mathcal{X} \times \mathcal{Y}$, there is at least one $z \in \mathcal{Z}$, such that $(x,y,z) \in f$. In a two-way protocol $\mathcal{P}$ for computing $f$, Alice and Bob get inputs $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ respectively. They send messages (qubits) to each other in turns, and their intention is to determine an answer $z \in \mathcal{Z}$ such that $(x,y,z) \in f$. We assume that $\mathcal{P}$ starts with the internal work qubits of Alice and Bob in the state $|0\rangle$. Both the parties use only unitary transformations for their internal computation, except at the very end when the final recipient of a message makes a von-Neumann measurement of some of her qubits to declare the output. Thus, the joint state of Alice and Bob is always pure during the execution of $\mathcal{P}$. We also assume that the players make *safe* copies of their respective inputs using CNOT gates before commencing the protocol. These safe copies of the inputs are not affected by the subsequent operations of $\mathcal{P}$, and are never sent as messages. In this paper, we consider protocols with and without prior entanglement. By prior entanglement, we mean a pure quantum state $|\phi\rangle$ that is shared between Alice and Bob and that is independent of their input $(x,y)$. The state $|\phi\rangle$ can be supported on an extremely large number of qubits. The unitary transforms of Alice in $\mathcal{P}$ are allowed to address her share of the qubits of $|\phi\rangle$; similarly for Bob. The classical analogue of prior entanglement is shared random bits. Often, the prior entanglement in a quantum protocol is in the

form of some number of EPR pairs, one-half of which belongs to Alice and the other half belongs to Bob.

Given $\epsilon \in (0, 1/2)$, the two-way quantum communication complexity $\mathsf{Q}_\epsilon(f)$ is defined to be the communication of the best two-way quantum protocol without prior entanglement, with error at most $\epsilon$ on all inputs. Whenever error parameter $\epsilon$ is not specified it is assumed to be $1/3$. Given a distribution $\mu$ on $\mathcal{X} \times \mathcal{Y}$, we can similarly define the quantum distributional two-way communication complexity of $f$, denoted $\mathsf{Q}^\mu_\epsilon(f)$, to be the communication of the best one-way quantum protocol without entanglement for $f$, such that the average error of the protocol over the inputs drawn from the distribution $\mu$ is at most $\epsilon$. We define $\mathsf{Q}^{[\,]}_\epsilon(f) \stackrel{\mathsf{def}}{=} \max_{\mu \text{ product}} \mathsf{Q}^\mu_\epsilon(f)$. The corresponding quantities for protocols with entanglement are denoted with the superscript pub.

The following result due to Yao [Yao77] is a very useful fact connecting worst-case and distributional communication complexities. It is a consequence of the *MiniMax* theorem in game theory [KN97, Thm. 3.20, page 36].

**Lemma 1 (Yao's principle [Yao77]).** $\mathsf{Q}^{\mathsf{pub}}_\epsilon(f) = \max_\mu \mathsf{Q}^{\mathsf{pub},\mu}_\epsilon(f)$.

Similar relationships as above also hold in the various other models that we mention below *mutates mutandis*.

In the one-way protocols we consider, the single message is always assumed to be from Alice to Bob unless otherwise specified. Sometimes we specify the direction of the message for example with superscript $A \to B$. The complexities $\mathsf{Q}^1_\epsilon(f), \mathsf{Q}^{1,\mathsf{pub}}_\epsilon(f), \mathsf{Q}^{1,\mu}_\epsilon(f), \mathsf{Q}^{1,[\,]}_\epsilon(f)$ could be analogously defined in the one-way case.

In the Simultaneous Message Passing (SMP) model, Alice and Bob each send a message each to a third party called Referee. In the SMP protocols we consider, we let prior entanglement to be shared between Alice and Referee and Bob and Referee and public coins to be shared between Alice and Bob. The communication complexity in this described model is denoted by $\mathsf{Q}^{\|,\mathsf{pub}}_\epsilon(f)$.

**Classical communication complexity:** Let us now consider classical communication protocols. We let $\mathsf{D}(f)$ represent the deterministic two-way communication complexity, that is the communication of the best deterministic two-way protocol computing $f$ correctly on all inputs. Let $\mu$ be a probability distribution on $\mathcal{X} \times \mathcal{Y}$ and $\epsilon \in (0, 1/2)$. We let $\mathsf{D}^\mu_\epsilon(f)$ represent the distributional two-way communication complexity of $f$ under $\mu$ with expected error $\epsilon$, i.e., the communication of the best private-coin two-way protocol for $f$, with distributional error (average error over the coins and the inputs) at most $\epsilon$ under $\mu$. It is easily noted that $\mathsf{D}^\mu_\epsilon(f)$ is always achieved by a deterministic two-way protocol, and henceforth we will restrict ourselves to deterministic protocols in the context of distributional communication complexity. We let $\mathsf{R}^{\mathsf{pub}}_\epsilon(f)$ represent the public-coin randomized two-way communication complexity of $f$ with worst case error $\epsilon$, i.e., the communication of the best public-coin randomized two-way protocol for $f$ with error for each input $(x, y)$ being at most $\epsilon$. The analogous quantity for private coin randomized protocols is denoted by $\mathsf{R}_\epsilon(f)$. The public- and private-coin randomized communication complexities are not much different, as shown in Newman's result [New91] that

$$\mathsf{R}(f) = O(\mathsf{R}^{\mathsf{pub}}(f) + \log\log|\mathcal{X}| + \log\log|\mathcal{Y}|). \tag{1}$$

We define $\mathsf{R}^{[\,]}_\epsilon(f) \stackrel{\mathsf{def}}{=} \max_{\mu \text{ product}} \mathsf{D}^\mu_\epsilon(f)$. The analogous communication complexities for one-way protocols could be similarly defined. As before, we put superscript 1 to signify that they stand

for one-way protocols and superscript $\|$ to signify SMP protocols. In classical public coin SMP protocols that we consider, we let the public coins to be shared between Alice and Bob.

### 2.3 Substate Theorem and $(\delta, \alpha)$-corrector

All our message compression arguments are based on the following common idea: If Alice does not reveal much information about her input, then it must be the case that Bob's state after receiving Alice's messages does not vary much (as Alice's input varies). In this situation, Alice and Bob can start in a suitable input independent state and Alice can account for the variation by applying appropriate local transformations on her registers. We formalize this idea using the notion of a $(\delta, \alpha)$-corrector, and establish the existence of such correctors by appealing to the following information-theoretic result, the Substate Theorem due to Jain, Radhakrishnan, and Sen [JRS02].

**Fact 1 (Substate Theorem, [JRS02])** *Let $\mathcal{H}, \mathcal{K}$ be two finite dimensional Hilbert spaces and $\dim(\mathcal{K}) \geq \dim(\mathcal{H})$. Let $\mathbb{C}^2$ denote the two dimensional complex Hilbert space. Let $\rho, \sigma$ be density matrices in $\mathcal{H}$ such that $S(\rho\|\sigma) < \infty$. Let $|\overline{\rho}\rangle$ be a purification of $\rho$ in $\mathcal{H} \otimes \mathcal{K}$. Then, for $r > 1$, there exist pure states $|\phi\rangle, |\theta\rangle \in \mathcal{H} \otimes \mathcal{K}$ and $|\overline{\sigma}\rangle \in \mathcal{H} \otimes \mathcal{K} \otimes \mathbb{C}^2$, depending on $r$, such that $|\overline{\sigma}\rangle$ is a purification of $\sigma$ and $\| |\overline{\rho}\rangle\langle\overline{\rho}| - |\phi\rangle\langle\phi| \|_{\mathrm{tr}} \leq \frac{2}{\sqrt{r}}$, where*

$$|\overline{\sigma}\rangle \stackrel{\mathsf{def}}{=} \sqrt{\frac{r-1}{r2^{rc}}}\,|\phi\rangle|1\rangle + \sqrt{1 - \frac{r-1}{r2^{rc}}}\,|\theta\rangle|0\rangle$$

*and $c \stackrel{\mathsf{def}}{=} S(\rho\|\sigma) + O(\sqrt{S(\rho\|\sigma)}) + O(1)$. Note that one can, by means of a local unitary operator on $\mathcal{K} \otimes \mathbb{C}^2$, transform any known purification $|\overline{\sigma'}\rangle$ of $\sigma$ to $|\overline{\sigma}\rangle$. Also, measuring the last qubit of $|\overline{\sigma}\rangle$ and observing a $|1\rangle$ puts the remaining qubits into the state $|\phi\rangle$. It follows that for every purification $|\overline{\sigma'}\rangle$ of $\sigma$, there is an unnormalized superoperator $\mathcal{M}$, depending on $|\overline{\sigma'}\rangle$, acting on the qubits of $|\overline{\sigma'}\rangle$ other than those of $\sigma$, such that $\mathcal{M}(|\sigma'\rangle\langle\sigma'|)$ normalized is equal to $|\phi\rangle$. Furthermore, this superoperator succeeds with probability at least $\frac{r-1}{r2^{rc}}$.*

**Definition 4 ($(\delta, \alpha)$-corrector).** *Let* Alice *and* Bob *form a bipartite quantum system. Let $X$ denote* Alice*'s input register, whose values range over the set $\mathcal{X}$. For $x \in \mathcal{X}$, let $\sigma_x$ be a state wherein the state of the register $X$ is $|x\rangle$; that is, $\sigma_x$ has the form $|x\rangle\langle x| \otimes \rho_x$. Let $\mu$ be a probability distribution on $\mathcal{X}$. Let $\sigma$ be some other joint state of* Alice *and* Bob*. A $(\delta, \alpha)$-corrector for the ensemble $\{\{\sigma_x\}_{x\in\mathcal{X}}; \sigma\}$ with respect to the distribution $\mu$ is a family of unnormalized superoperators $\{\mathcal{M}_x\}_{x\in\mathcal{X}}$ acting only on* Alice*'s qubits such that:*

1. *$r_x \stackrel{\mathsf{def}}{=} \mathsf{Tr}\mathcal{M}_x(\sigma) = \alpha$ for all $x \in \mathcal{X}$, that is, $\mathcal{M}_x$ when applied to $\sigma$ succeeds with probability exactly $\alpha$.*
2. *$\mathcal{M}_x(\sigma)$ has the form $|x\rangle\langle x| \otimes \rho'_x$, that is, the state of the register $X$ of* Alice *is $|x\rangle$ when $\mathcal{M}_x$ succeeds.*
3. *$\mathbb{E}_\mu \left[ \left\| \sigma_x - \frac{1}{\alpha}\mathcal{M}_x(\sigma) \right\|_{\mathrm{tr}} \right] \leq \delta$, that is, $\mathcal{M}_x$ on success* corrects *the state $\sigma$ by bringing it to within trace distance $\delta$ from $\sigma_x$.*

We shall also need the following observation.

**Proposition 1.** *Suppose a boolean-valued measurement $\mathcal{M}$ succeeds with probabilities $p$, $q$ on quantum states $\rho$, $\sigma$ respectively. Let $\rho'$, $\sigma'$ be the respective quantum states if $\mathcal{M}$ succeeds. Then, $\|\rho' - \sigma'\|_{\mathrm{tr}} \leq \frac{1}{\max\{p,q\}} \|\rho - \sigma\|_{\mathrm{tr}}$.*

*Proof.* We formalize the intuition that if some measurement distinguishes $\rho'$ and $\sigma'$, then there is a measurement that distinguishes $\rho$ and $\sigma$. Assume $p \geq q$ (otherwise interchange the roles of $\rho$ and $\sigma$). Now there exists (see e.g. [AKN98]) an orthogonal projection $M'$, such that $\mathsf{Tr}M'(\rho'-\sigma') = \frac{\|\rho'-\sigma'\|_{\mathrm{tr}}}{2}$. Let $M''$ be the POVM element obtained by first applying POVM $\mathcal{M}$ and on success applying $M'$. Then the probability of success of $M''$ on $\rho$ is $p \cdot \mathsf{Tr}M'\rho'$, and the probability of success of $M''$ on $\sigma$ is $q \cdot \mathsf{Tr}M'\sigma' \leq p \cdot \mathsf{Tr}M'\sigma'$. Thus,

$$\begin{aligned}\frac{1}{2}\left\|\rho-\sigma\right\|_{\mathrm{tr}} &\geq \mathsf{Tr}M''\rho - \mathsf{Tr}M''\sigma \\ &\geq p(\mathsf{Tr}M'\rho' - \mathsf{Tr}M'\sigma') \\ &= \frac{p}{2}\cdot\left\|\rho'-\sigma'\right\|_{\mathrm{tr}},\end{aligned}$$

implying that $\left\|\rho'-\sigma'\right\|_{\mathrm{tr}} \leq \frac{\|\rho-\sigma\|_{\mathrm{tr}}}{p}$. □

We are now ready to use the Substate Theorem to show the existence of good correctors when Bob's state does not contain much information about Alice's input. While applying the Substate Theorem below, it will be helpful to think of Alice's Hilbert space as $\mathcal{K}\otimes\mathbb{C}^2$ and Bob's Hilbert space as $\mathcal{H}$ in Fact 1.

**Lemma 2.** *For $x \in \mathcal{X}$, let $|\phi_x\rangle \stackrel{\mathsf{def}}{=} |x\rangle|\psi_x\rangle$ be a joint pure state of* Alice *and* Bob*, where $|x\rangle$ and possibly some other qubits of $|\psi_x\rangle$ belong to* Alice*'s subsystem $A$, and the remaining qubits of $|\psi_x\rangle$ belong to* Bob*'s subsystem $B$. Let $\mu$ be a probability distribution on $\mathcal{X}$; let $\sigma \stackrel{\mathsf{def}}{=} \mathbb{E}_\mu|\phi_x\rangle\langle\phi_x|$ and $|\phi\rangle \stackrel{\mathsf{def}}{=} \sum_x \sqrt{\mu(x)}|\phi_x\rangle$. Let $X$ denote the register of* Alice *containing $|x\rangle$. Suppose $I(X:B) = k$, when the joint state of $AB$ is $\sigma$. Then for $\delta > 0$, there is a $(\delta,\alpha)$-corrector $\{\mathcal{M}_x\}_{x\in\mathcal{X}}$ for the ensemble $\{\{|\phi_x\rangle\};|\phi\rangle\}$ where $\alpha = 2^{-O(k/\delta^3)}$.*

*Proof.* Let $\rho_x \stackrel{\mathsf{def}}{=} \mathsf{Tr}_A|\phi_x\rangle\langle\phi_x|$ and $\rho \stackrel{\mathsf{def}}{=} \mathsf{Tr}_A|\phi\rangle\langle\phi|$. Note that $\rho = \mathbb{E}_\mu\rho_x$. Now, $k = I(X:B) = \mathbb{E}_\mu S(\rho_x\|\rho)$. By Markov's inequality, there is a subset $\mathsf{Good} \subseteq \mathcal{X}$, $\Pr_\mu[\mathsf{Good}] \geq 1-\delta/4$, such that for all $x \in \mathsf{Good}$, $S(\rho_x\|\rho) \leq 4k/\delta$. We will define superoperators $\mathcal{M}_x$ for $x \in \mathsf{Good}$ and $x \notin \mathsf{Good}$ separately, and then show that they form a $(\delta,\alpha)$-corrector.

Fix $x \in \mathsf{Good}$. Using Fact 1 with $r$ to be chosen later, we conclude that for all $x \in \mathsf{Good}$, there is an unnormalized superoperator $\tilde{\mathcal{M}}_x$ acting on $A$ only such that if $q_x \stackrel{\mathsf{def}}{=} \mathsf{Tr}\tilde{\mathcal{M}}_x(|\phi\rangle\langle\phi|)$, $\tilde{\sigma}_x \stackrel{\mathsf{def}}{=} \frac{\tilde{\mathcal{M}}_x(|\phi\rangle\langle\phi|)}{q_x}$ then, $q_x \geq \frac{r-1}{r2^{4rk/\delta}}$ and $\|\tilde{\sigma}_x - |\phi_x\rangle\langle\phi_x|\|_{\mathrm{tr}} \leq \frac{2}{\sqrt{r}}$. Now, measure register $X$ in $\tilde{\sigma}_x$ and declare success if the result is $x$. Let $\sigma'_x$ be the resulting normalized state when $x$ is observed. Measuring $X$ in $|\phi_x\rangle$ results gives the value $x$ with probability 1. Hence, by Proposition 1,

$$\left\|\sigma'_x - |\phi_x\rangle\langle\phi_x|\right\|_{\mathrm{tr}} \leq \frac{2}{\sqrt{r}}.$$

Furthermore, since $\|\tilde{\sigma}_x - |\phi_x\rangle\langle\phi_x|\|_{\mathrm{tr}} \leq \frac{2}{\sqrt{r}}$, the probability $q'_x$ of observing $x$ when $X$ is measured in the state $\tilde{\sigma}_x$ is at least $1-\frac{1}{\sqrt{r}}$, and the overall probability of success is at least $q_xq'_x \geq (1-\frac{1}{\sqrt{r}})(\frac{r-1}{r2^{4rk/\delta}}) \stackrel{\mathsf{def}}{=} \alpha$. In order to ensure that the overall probability of success is exactly $\alpha$, we do a further *rejection* step: Even on success we artificially declare failure with probability $1-\frac{\alpha}{q_xq'_x}$. Let $\mathcal{M}_x$ be the unnormalized superoperator which first applies $\tilde{\mathcal{M}}_x$, then measures the register $X$,

and on finding $x$ accepts with probability $\frac{\alpha}{q_x q'_x}$. Thus, for all $x \in \mathsf{Good}$, the probability of success $r_x \stackrel{\mathsf{def}}{=} \mathsf{Tr}\mathcal{M}_x(|\phi\rangle\langle\phi|)$ is exactly equal to $\alpha$. This completes the definition of $\mathcal{M}_x$ for $x \in \mathsf{Good}$.

For $x \notin \mathsf{Good}$, $\mathcal{M}_x$ swaps $|x\rangle$ into register $X$ from some outside ancilla initialized to $|0\rangle$ and declares success artificially with probability $r_x = \alpha$. For all $x \in \mathcal{X}$, let $\sigma'_x \stackrel{\mathsf{def}}{=} \frac{\mathcal{M}_x(|\phi\rangle\langle\phi|)}{r_x}$.

Thus for all $x \in \mathcal{X}$, $\sigma'_x$ contains $|x\rangle$ in register $X$ and $r_x = \alpha$. Finally, we have

$$\begin{aligned}
&\mathbb{E}_\mu \left\| \sigma'_x - |\phi_x\rangle\langle\phi_x| \right\|_{\mathrm{tr}} \\
&\leq \sum_{x \in \mathsf{Good}} \mu(x) \left\| \sigma'_x - |\phi_x\rangle\langle\phi_x| \right\|_{\mathrm{tr}} + \sum_{x \notin \mathsf{Good}} \mu(x) \cdot 2 \\
&\leq \frac{2}{\sqrt{r}} + \frac{\delta}{4} \cdot 2.
\end{aligned}$$

For $r = \frac{16}{\delta^2}$, this quantity is at most $\delta$, and we conclude that the family $\{\mathcal{M}_x\}_{x \in \mathcal{X}}$ forms the required $(\delta, \alpha)$-corrector for the ensemble $\{\{|\phi_x\rangle\}_{x \in \mathcal{X}}; |\phi\rangle\}$ with $\alpha = 2^{-O(k/\delta^3)}$. □

### 2.4 Miscellaneous

We have the following Lemma.

**Lemma 3.** *Let $\delta > 0$. Let $P, Q$ be probability distributions with support on set $\mathcal{X}$ such that $S(P||Q) \leq c$. Then, we get a set $\mathsf{Good} \subseteq \mathcal{X}$ such that*

$$\Pr_P[x \in \mathsf{Good}] \geq 1 - \delta \quad \text{and} \quad \forall x \in \mathsf{Good}, \frac{P(x)}{Q(x)} \leq 2^{\frac{c+1}{\delta}}. \tag{2}$$

*Proof.* We first have the following claim:

*Claim 1.* Let $P$ and $Q$ be two distributions on the set $\mathcal{X}$. For any set $\mathcal{X}' \subseteq \mathcal{X}$, we have

$$\sum_{x \in \mathcal{X}'} P(x) \log \frac{P(x)}{Q(x)} \quad \geq \quad -\frac{\log e}{e} \quad > \quad -1.$$

*Proof.* We require the following facts.

1. *log-sum inequality:* For non-negative integers $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$,
$$\sum a_i \log \frac{a_i}{b_i} \geq \left(\sum a_i\right) \log \frac{\sum a_i}{\sum b_i}.$$
2. The function $x \log x \geq -(\log e)/e$ for all $x > 0$.

From the above, we have the following sequence of inequalities.

$$\begin{aligned}
\sum_{x \in \mathcal{X}'} P(x) \log \frac{P(x)}{Q(x)} &= \sum_{x \in \mathcal{X}'} P(x) \log \frac{P(x)}{Q(x)} + \sum_{x \notin \mathcal{X}'} Q(x) \log \frac{Q(x)}{Q(x)} \\
&\geq \left( \sum_{x \in \mathcal{X}'} P(x) + \sum_{x \notin \mathcal{X}'} Q(x) \right) \log \left( \frac{\sum_{x \in \mathcal{X}'} P(x) + \sum_{x \notin \mathcal{X}'} Q(x)}{\sum_{x \in \mathcal{X}} Q(x)} \right) \\
&= \left( \sum_{x \in \mathcal{X}'} P(x) + \sum_{x \notin \mathcal{X}'} Q(x) \right) \log \left( \sum_{x \in \mathcal{X}'} P(x) + \sum_{x \notin \mathcal{X}'} Q(x) \right) \\
&\geq -\frac{\log e}{e}
\end{aligned}$$

□

Now:

$$\begin{aligned} c \geq S(P||Q) &= \sum_{x:P(x)\geq Q(x)} P(x)\log\frac{P(x)}{Q(x)} + \sum_{x:P(x)<Q(x)} P(x)\log\frac{P(x)}{Q(x)} \\ &> \sum_{x:P(x)\geq Q(x)} P(x)\log\frac{P(x)}{Q(x)} - 1 \\ \Rightarrow c+1 &> \sum_{x:P(x)\geq Q(x)} P(x)\log\frac{P(x)}{Q(x)} \end{aligned}$$

Now we get our desired set Good immediately by using Markov's inequality. □

We also need the following lemma.

**Lemma 4.** *Let $XAB$ be a tri-partite system with $X$ classical and $A, B$ quantum systems. If $I(X : A) = 0$ then $I(X : AB) \leq 2S(B)$.*

*Proof.* We have the following Araki-Lieb [AL70] inequality for any two systems $M_1, M_2$: $|S(M_1) - S(M_2)| \leq S(M_1M_2)$. This implies:

$$I(M_1 : M_2) = S(M_1) + S(M_2) - S(M_1M_2) \leq \min\{2S(M_1), 2S(M_2)\}.$$

Now,

$$\begin{aligned} I(X : AB) &= I(X : A) + I(XA : B) - I(A : B) \\ &\leq I(XA : B) \leq 2S(B). \end{aligned}$$

## 3 One-way Message Compression and Optimal Direct Sum

Although in this section are concerned with message compression in one-way protocols, we state our results in a general setting of compressing the first message of multi-round two-way protocols. This way of stating our message compression results is helpful in expressing our round-elimination results. We state our results and proofs here only for quantum protocols and the corresponding results for classical protocols can be obtained in analogous fashion. We skip making explicit statements and proofs for classical protocols for brevity.

### 3.1 Message Compression and Round Elimination

We begin with the following definition.

**Definition 5 ($[t; l_1, \ldots, l_t]^A$ protocol).** *In a $[t; l_1, \ldots, l_t]^A$ protocol, there are $t$ rounds of communication with* Alice *starting, the $i$th message being $l_i$ qubits long. A $[t; l_1, \ldots, l_t]^B$ protocol is the same but* Bob *starts the communication.*

**Theorem 3 (Compressing the first message).** *Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a function and $\mu$ be a probability distribution on $\mathcal{X} \times \mathcal{Y}$. Suppose $\mathcal{P}$ is a $[t; l_1, l_2, \ldots, l_t]^A$ quantum protocol without prior entanglement for $f$ having average error less than $\epsilon$ under $\mu$. Let $X$ denote the random variable corresponding to* Alice*'s input and $N_1$ denote the qubits of* Alice*'s first message in $\mathcal{P}$. Suppose $I(X : N_1) \leq k$. Let $\delta > 0$ be a sufficiently small constant. Then, there is a $[t; \beta, l_2 \ldots, l_t]^A$ quantum protocol $\mathcal{P}'$ with prior entanglement for $f$ with average error less than $\epsilon + \delta$ under $\mu$, where $\beta = O\left(\frac{k}{\delta^3}\right)$. Also, the first message of* Alice *in $\mathcal{P}'$ is classical.*

*Proof.* Let $|\phi_x\rangle$ denote the state vector in $\mathcal{P}$ of Alice's qubits (including her input register) and her first message $N_1$ just after she sends $N_1$ to Bob, when she is given input $x \in \mathcal{X}$. Let $|\phi\rangle$ denote the corresponding state vector in $\mathcal{P}$ when the protocol starts with Alice's input registers in the state $\sum_x \sqrt{p_x}|\phi_x\rangle$, where $p_x \stackrel{\mathsf{def}}{=} \Pr_\mu[X = x]$. Since $I(X : N_1) \leq k$, Lem. 2 implies that there is a $(\delta/2, \alpha)$-corrector $\{\mathcal{M}_x\}_{x\in\mathcal{X}}$ for the ensemble $\{\{|\phi_x\rangle\}_{x\in\mathcal{X}}; |\phi\rangle\}$ where $\alpha = 2^{-O(k/\delta^3)}$. That is, with $r_x \stackrel{\mathsf{def}}{=} \mathsf{Tr}(\mathcal{M}_x|\phi\rangle\langle\phi|)$ and $\sigma'_x \stackrel{\mathsf{def}}{=} \frac{\mathcal{M}_x(|\phi\rangle\langle\phi|)}{r_x}$, we have $\mathbb{E}_\mu\left[\|\sigma'_x - |\phi_x\rangle\langle\phi_x|\|_{\mathrm{tr}}\right] \leq \frac{\delta}{2}$.

We now describe the protocol $\mathcal{P}'$. The protocol $\mathcal{P}'$ starts with $2^\beta \stackrel{\mathsf{def}}{=} \alpha^{-1}\log(2/\delta)$ copies of $|\phi\rangle$ as prior entanglement. Alice applies $\mathcal{M}_x$ to each copy of $|\phi\rangle$ and sends the index of the first copy on which she achieves success. Thus, her first message in $\mathcal{P}'$ is classical and $\beta = \log(1/\alpha) + \log\log(2/\delta) = O(k/\delta^3)$ bits long. Alice and Bob use that copy henceforth; the rest of $\mathcal{P}'$ is exactly as in $\mathcal{P}$. The probability that Alice achieves success with $\mathcal{M}_x$ on at least one copy of $|\phi\rangle$ is more than $1 - \frac{\delta}{2}$. Furthermore, the state of Alice's registers and the first message $N_1$ on this copy is exactly $\sigma'_x$. Thus, the probability of error for the protocol $\mathcal{P}'$ is at most

$$\epsilon + \frac{\delta}{2} + \mathbb{E}_\mu\left[\left\|\sigma'_x - |\phi_x\rangle\langle\phi_x|\right\|_{\mathrm{tr}}\right] \leq \epsilon + \frac{\delta}{2} + \frac{\delta}{2} \leq \epsilon + \delta.$$

This completes the proof of the theorem. □

*Remark:* We can eliminate prior entanglement in quantum protocols by assuming that Alice generates the prior entangled state herself, and then sends Bob's share of the state along with her first message. This can make Alice's first message long, but if the information about $X$ in Alice's first message together with Bob's share of prior entanglement qubits in the original protocol is small, then the conclusions of the theorem still hold.

**Corollary 1 (Eliminating the first round).** *Under the conditions of Thm. 3, if $t \geq 3$ there is a $[t-1; 2^\beta l_2, l_3 + \beta, l_4, \ldots, l_t]^B$ quantum protocol $\tilde{\mathcal{P}}$ with prior entanglement for $f$ with average error at most $\epsilon + \delta$ under $\mu$. If $t = 2$, we get a $[1; 2^\beta l_2]^B$ quantum protocol $\tilde{\mathcal{P}}$ with prior entanglement for $f$ with average error at most $\epsilon + \delta$ under $\mu$.*

*Proof.* Suppose $t \geq 3$. Let $N_2$, $N_3$ denote the second and third messages of $\mathcal{P}'$. Consider a $(t-1)$-round protocol $\tilde{\mathcal{P}}$ where Bob begins the communication by sending his messages $N_2$ for all the $2^\beta$ copies of $|\phi\rangle$. This makes Bob's first message in $\tilde{\mathcal{P}}$ to be $2^\beta l_2$ qubits long. Alice replies by applying $\mathcal{M}_x$ to each copy of $|\phi\rangle$ and sending the index of the first copy on which she achieves success. She also sends her response $N_3$ corresponding to that copy of $|\phi\rangle$. Thus, her first message in $\tilde{\mathcal{P}}$ is $l_3 + \beta$ qubits long. Note that the operations of Bob and the applications of $\mathcal{M}_x$ by Alice during the first two messages of $\tilde{\mathcal{P}}$ are on disjoint sets of qubits, hence they commute. Thus, the global state vector of $\tilde{\mathcal{P}}$ after the second message is exactly the same as the global state vector of $\mathcal{P}'$ after the third message. Hence the error probability remains the same. This proves the first statement of the corollary. The second statement of the corollary (case $t = 2$) can be proved similarly. □

*Remark:* The above corollary can be thought of as the quantum analogue of the 'message switching' lemma of [CR04].

Using Corr. 1, we can now prove our new round elimination result for quantum protocols.

**Theorem 4 (Round elimination lemma).** *Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be a function and $k$, $t$ be positive integers. Suppose $t \geq 3$. Suppose $\mathcal{P}$ is a $[t; l_1, l_2, l_3, \ldots, l_t]^A$ quantum protocol with prior*

*entanglement for $f^{(k),A}$ (recall definition Def. 1) with worst case error less than $\epsilon$. Let $\delta > 0$ be a sufficiently small constant. Let $\beta \stackrel{\mathsf{def}}{=} O(\frac{l_1}{\delta^3 k})$. Then there is a $[t-1; 2^\beta l_2, l_3+\beta, \ldots, l_t]^B$ quantum protocol with prior entanglement for $f$ with worst case error at most $\epsilon+\delta$.*

*Proof.* (**Sketch**) The proof follows in a standard fashion by combining the proof technique of Lem. 4 of [Sen03] with Corr. 1. We skip making a complete proof for brevity. □

*Remark:* The above round elimination lemma is quantum analogue of a classical round elimination result of Chakrabarti and Regev [CR04]. It allows us to extend their optimal randomized cell probe lower bound for Approximate Nearest Neighbor Searching in the Hamming cube $\{0,1\}^n$ to the quantum address-only cell probe model defined by Sen and Venkatesh [SV01]. It also allows us to extend the sharper lower bounds for Predecessor Searching of Patrascu and Thorup [PT06a] to the quantum case. We skip making explicit statements and their proofs for brevity.

### 3.2 One-Way Optimal Direct Sum

We get the following implication of Thm. 3 to the Direct Sum problem for one-round quantum communication protocols. Recall that $f^{\oplus m}$ is the $m$-fold Direct Sum problem corresponding to the relation $f$.

**Theorem 5 (Direct Sum).** *Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation. Let $\epsilon, \delta \in (0, 1/2)$ with $\epsilon + \delta < 1/2$. For one-round quantum protocols with prior entanglement, we get*

$$\mathsf{Q}_\epsilon^{1,A\to B,\mathsf{pub}}(f^{\oplus m}) \quad \geq \quad \Omega\left(\delta^3 m \cdot \mathsf{Q}_{\epsilon+\delta}^{1,A\to B,\mathsf{pub}}(f)\right).$$

*Similar result also holds by switching the roles of* Alice *and* Bob. *For simultaneous message protocols, we get*

$$\mathsf{Q}_\epsilon^{\|,\mathsf{pub}}(f^{\oplus m}) \quad \geq \quad \Omega\left(\delta^3 m \cdot \mathsf{Q}_\epsilon^{\|,\mathsf{pub}}(f)\right).$$

*Proof.* We present the proof for one-round protocols and the proof for SMP protocols follows very similarly. Below we assume that in the one-way protocols we consider the single message is from Alice to Bob, and hence we do not explicitly mention it in the superscripts. Let $\epsilon, \delta$ be as in the statement of the theorem and let $c \stackrel{\mathsf{def}}{=} \mathsf{Q}_\epsilon^{1,\mathsf{pub}}(f^{\oplus m})$. For showing our result we will show that for all distributions $\lambda$ on $\mathcal{X} \times \mathcal{Y}$,

$$\mathsf{Q}_{\epsilon+\delta}^{1,\mathsf{pub},\lambda}(f) \quad = \quad O(\frac{c}{\delta^3 m}). \tag{3}$$

Using Yao's principle and Eq. (3), we immediately get the desired result as follows:

$$\mathsf{Q}_{\epsilon+\delta}^{1,\mathsf{pub}}(f) \quad = \quad \max_{\lambda \text{ on } \mathcal{X}\times\mathcal{Y}} \mathsf{Q}_{\epsilon+\delta}^{1,\mathsf{pub},\lambda}(f) \quad = \quad O(\frac{c}{\delta^3 m}) \quad = \quad O(\frac{1}{\delta^3 m} \cdot \mathsf{Q}_\epsilon^{1,\mathsf{pub}}(f^{\oplus m})).$$

Let us now turn to showing Eq. (3). Since $\mathsf{Q}_\epsilon^{1,\mathsf{pub}}(f^{\oplus m}) = c$, let $\mathcal{P}$ be a protocol (possibly using entanglement) for $f^{\oplus m}$ with communication $c$ and error on every input being at most $\epsilon$. Let us consider a distribution $\mu$ (possibly non-product) on $\mathcal{X} \times \mathcal{Y}$. Our intention is to exhibit a protocol $\tilde{\mathcal{P}}$ for $f$ with communication $O(\frac{c}{\delta^3 m})$ and distributional error at most $\epsilon+\delta$ under $\mu$ and this would imply from definition that $\mathsf{Q}_{\epsilon+\delta}^{1,\mathsf{pub},\mu}(f) = O(\frac{c}{\delta^3 m})$, and we would be done.

From $\mathcal{P}$ let us get a protocol $\mathcal{P}'$ without prior entanglement in which Alice generates both parts of the shared state herself and then sends Bob's part as part of her first message. Alice and Bob then behave identically as in $\mathcal{P}$. Now let us provide inputs to $\mathcal{P}'$ as follows. Let $\mu_X$ be the marginal of $\mu$ on $\mathcal{X}$. Recall that Alice has $m$ parts of the inputs in $\mathcal{P}'$. Let the input of Alice be distributed according to $\mu_X$ in each part independently, and let Bob get input 0 in every part. Let $X$ be the random variable representing the combined input of Alice. Let $X_i, i \in [m]$ be the random variable representing the input of Alice on the $i$-th co-ordinate. Note that $X_i, i \in [m]$ are all independent. Let $M$ represent the message of Alice. Now using Lem. 4 (irrespective of the number of qubits of prior entanglement in $\mathcal{P}$) we have $2c \geq I(X : M)$. Now from Chain Rule of mutual information we get:

$$2c \quad \geq \quad I(X : M) \quad = \quad \sum_{i=1}^{m} I(X_i : M).$$

Therefore there exists a co-ordinate $i_0 \in [m]$ such that $I(X_{i_0} : M) \leq \frac{2c}{m}$. Now let us define a protocol $\mathcal{P}''$ for $f$, in which on getting input $x \in \mathcal{X}, y \in \mathcal{Y}$ respectively (sampled jointly according to $\mu$), Alice and Bob simulate $\mathcal{P}'$ by assuming $x$ and $y$ to be inputs for the $i_0$-th co-ordinate. For the rest of the co-ordinates Alice generates the inputs independently according to the distribution $\mu_X$. Bob simply inserts 0 as inputs in the rest of the co-ordinates. Alice then acts identically as in $\mathcal{P}$ and sends her message $M''$ to be Bob, who then outputs his decision as in $\mathcal{P}$. Note that in this case too $I(X''_{i_0} : M'') \leq \frac{2c}{m}$, where $X''_{i_0}$ represents the input of Alice in the $i_0$-th co-ordinate. Also note that since the error of $\mathcal{P}$ on every input was at most $\epsilon$, we have that the distributional error under $\mu$ in $\mathcal{P}''$ is also at most $\epsilon$.

We are now ready to define our intended protocol $\tilde{\mathcal{P}}$. Protocol $\tilde{\mathcal{P}}$ is obtained by compressing the message of Alice in protocol $\mathcal{P}''$ as according to Thm. 3 (by assuming $t = 1$). Hence the message of Alice in $\tilde{\mathcal{P}}$ has length $O(\frac{c}{m\delta^3})$ and the distributional error of $\tilde{\mathcal{P}}$ under $\mu$ is at most $\epsilon + \delta$. □

## 4 Multi-Round Message Compression and Weak Direct Sum

### 4.1 Quantum Protocols

In this section, we state and formally prove our results for compressing messages in multi-round quantum communication protocols for computing a relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. In our discussion below, $A, X, B, Y$ denote Alice's work qubits, Alice's input qubits, Bob's work qubits and Bob's input qubits respectively, at a particular point in time.

**Definition 6 (Privacy loss).** *Let $\mu \stackrel{\text{def}}{=} \mu_{\mathcal{X}} \times \mu_{\mathcal{Y}}$ be a product probability distribution on $\mathcal{X} \times \mathcal{Y}$. Suppose $\mathcal{P}$ is a quantum protocol for a relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. Consider runs of $\mathcal{P}$ when* Alice*'s input register $X$ starts in the mixed state $\sum_{x \in \mathcal{X}} \mu_{\mathcal{X}}(x)|x\rangle\langle x|$ and* Bob*'s input register $Y$ starts in the pure state $\sum_{y \in \mathcal{Y}} \sqrt{\mu_{\mathcal{Y}}(y)}|y\rangle$. Let $B$ denote the qubits in the possession of* Bob *including $Y$, at some point during the execution of $\mathcal{P}$. Let $I(X : B)$ denote the mutual information of* Alice*'s input register $X$ with* Bob*'s qubits $B$. The* privacy loss *of $\mathcal{P}$ for relation $f$ on the distribution $\mu$ from* Alice *to* Bob *at that point in time is $L^{\mathcal{P}}(f, \mu, A, B) \stackrel{\text{def}}{=} I(X : B)$. The privacy loss from* Bob *to* Alice*, $L^{\mathcal{P}}(f, \mu, B, A)$, is defined similarly. The privacy loss of $\mathcal{P}$ from* Alice *to* Bob *for $f$, $L^{\mathcal{P}}(f, A, B)$, is the maximum over all product distributions $\mu$ of $L^{\mathcal{P}}(f, \mu, A, B)$. The privacy loss of $\mathcal{P}$ from* Bob *to* Alice *for $f$, $L^{\mathcal{P}}(f, B, A)$, is defined similarly. The privacy loss from* Alice *to* Bob *for $f$, $L(f, A, B)$, is the infimum over all protocols $\mathcal{P}$ of $L^{\mathcal{P}}(f, A, B)$ at the end of $\mathcal{P}$. The quantity $L(f, B, A)$ is defined similarly.*

**Theorem 6 (Compressing many rounds).** *Suppose $\mathcal{P}$ is a $[t; l_1, l_2, \ldots, l_t]^A$ quantum protocol without prior entanglement for a relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. Let $\mu \stackrel{\text{def}}{=} \mu_{\mathcal{X}} \times \mu_{\mathcal{Y}}$ be a product probability distribution on $\mathcal{X} \times \mathcal{Y}$. Suppose the average error of $\mathcal{P}$ when the inputs are chosen according to $\mu$ is at most $\epsilon$. Let $k_a$, $k_b$ denote the privacy losses of* Alice *and* Bob *respectively after $t'$ rounds of communication. Suppose $t'$ is odd (similar statements hold for even $t$, as well as for interchanging the roles of* Alice *and* Bob*). Then, for all sufficiently small constants $\delta > 0$, there exists a $[t - t' + 1; \lambda_1, \lambda_2, l_{t'+2}, \ldots, l_t]^A$ protocol $\mathcal{P}'$ in the presence of prior entanglement such that:*

1. *the average error of $\mathcal{P}'$ with respect to $\mu$ is at most $\epsilon + \delta$;*
2. *$\lambda_1 \leq k_a \cdot 2^{O(k_b/\delta^6)}$ and $\lambda_2 \leq l_{t'+1} + O(k_b/\delta^6)$.*

*Proof.* Consider the situation after $t'$ rounds of $\mathcal{P}$. Let the joint state of Alice and Bob be denoted by

$\sigma_{xy}$: when Alice starts $\mathcal{P}$ with $x$ in her input register and Bob starts with $y$ in his input register;
$\sigma_x$: when Alice starts with $x$ in her input register and Bob starts with the superposition $\sum_{y \in \mathcal{Y}} \sqrt{\mu_{\mathcal{Y}}(y)}|y\rangle$ in his input register;
$\sigma_y$: when Bob starts with $y$ in his input register and Alice starts with the superposition $\sum_{x \in \mathcal{X}} \sqrt{\mu_{\mathcal{X}}(x)}|x\rangle$ in her input register;
$\sigma$: when Alice and Bob start with the superposition $\sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} \sqrt{\mu(xy)}|x\rangle|y\rangle$ in their input registers.

Note that $\sigma_{xy}$, $\sigma_x$, $\sigma_y$ and $\sigma$ are pure states.

We overload the symbols $\mathcal{X}, \mathcal{Y}$ to also denote the superoperators corresponding to measuring in the computational basis the input registers $X, Y$ of Alice and Bob respectively. Whether $\mathcal{X}, \mathcal{Y}$ denote sets or superoperators will be clear from the context. When several superoperators are applied to a state in succession we omit the parenthesis; for example, we write $\mathcal{X}\mathcal{Y}(\rho)$ instead of $\mathcal{X}(\mathcal{Y}(\rho))$ which corresponds to measuring the input registers of Alice and Bob (in this case, their order does not matter).

We will choose $\delta_a, \delta_b > 0$ later. Since the privacy loss of Alice is at most $k_a$, Lem. 2 implies that there is a $(\delta_a, \alpha)$-corrector $\{\mathcal{M}_x\}_{x \in \mathcal{X}}$ for $\{\{\sigma_x\}_{x \in \mathcal{X}}; \sigma\}$ with $\alpha = 2^{-O(k_a/\delta_a^3)}$. Similarly, since the privacy loss of Bob is at most $k_b$, there is a $(\delta_b, \beta)$-corrector $\{\mathcal{M}_y\}_{y \in \mathcal{Y}}$ for $\{\{\sigma_y\}_{y \in \mathcal{Y}}; \sigma\}$ with $\beta = 2^{-O(k_b/\delta_b^3)}$. In particular, with $\mathcal{M}_X \stackrel{\text{def}}{=} \mathbb{E}_{\mu_{\mathcal{X}}}[\mathcal{M}_x]$ and $\mathcal{M}_Y \stackrel{\text{def}}{=} \mathbb{E}_{\mu_{\mathcal{Y}}}[\mathcal{M}_y]$, we have

$$\begin{aligned} \left\| \frac{\mathcal{M}_X(\sigma)}{\alpha} - \mathcal{X}(\sigma) \right\|_{\text{tr}} &\leq \delta_a, \\ \left\| \frac{\mathcal{M}_Y(\sigma)}{\beta} - \mathcal{Y}(\sigma) \right\|_{\text{tr}} &\leq \delta_b. \end{aligned} \tag{4}$$

In our proof, we will take

$$\delta_b \stackrel{\text{def}}{=} \left(\frac{\delta}{10}\right)^2, \quad \delta_a \stackrel{\text{def}}{=} \frac{\delta_b \beta}{2}. \tag{5}$$

The proof has two steps. In the first step, we analyze the protocol $\mathcal{P}'$ given in Figure 1. In $\mathcal{P}'$, Alice and Bob try to recreate the effect of the first $t'$ rounds of the original protocol, but without sending any messages. For this, they start from the state $\sigma$ (their prior entanglement) and on receiving $x$ and $y$, apply suitable correcting transformations. In the second step, we shall consider a protocol $\mathcal{P}''$ that starts with several parallel executions of $\mathcal{P}'$.

Alice *and* Bob *start with the joint state* $\sigma$ *as prior entanglement.*

**Input:** Alice is given $x \in X$; Bob is given $y \in Y$.
Alice**:** Applies superoperator $\mathcal{M}_x$ to her registers.
Bob**:** Applies superoperator $\mathcal{M}_y$ to his registers.

**Fig. 1.** The intermediate protocol $\mathcal{P}'$

Let $r_{xy} \stackrel{\text{def}}{=} \mathsf{Tr}\mathcal{M}_y\mathcal{M}_x(\sigma)$ and let $r \stackrel{\text{def}}{=} \mathbb{E}_\mu[r_{xy}]$. Then, $r_{xy}$ is the probability that both Alice and Bob succeed on input $(x, y)$, and $r$ is the probability that they succeed when their input is chosen according to the distribution $\mu$. Let $\rho$ denote the state after $t'$ rounds of $\mathcal{P}$ when the inputs are chosen according to $\mu$ i.e. $\rho \stackrel{\text{def}}{=} \mathbb{E}_\mu[\sigma_{xy}]$. Observe that $\rho = \mathcal{Y}\mathcal{X}(\sigma)$. Let $\rho'$ be the state at the end of $\mathcal{P}'$, when the inputs are chosen according to $\mu$ and we condition on both parties succeeding i.e. $\rho' = \frac{\mathcal{M}_Y\mathcal{M}_X(\sigma)}{r}$.

*Claim 2.* (a) $1 - \frac{\delta_b}{2} \leq \frac{r}{\alpha\beta} \leq 1 + \frac{\delta_b}{2}$.

(b) $\|\rho - \rho'\|_{\text{tr}} \leq 2\delta_b$.

(c) $\Pr_\mu\left[\left|\frac{r_{xy}}{r} - 1\right| \geq 2\delta_b^{1/2}\right] \leq \delta_b^{1/2}$.

*Proof.* (a)

$$\begin{aligned}\frac{r}{\alpha\beta} &= \frac{\mathsf{Tr}\mathcal{M}_Y\mathcal{M}_X(\sigma)}{\alpha\beta} \\ &= \frac{1}{\beta}\mathsf{Tr}\left(\mathcal{M}_Y\left(\frac{\mathcal{M}_X(\sigma)}{\alpha}\right)\right) \\ &= \frac{1}{\beta}\mathsf{Tr}\mathcal{M}_Y\mathcal{X}(\sigma) + \frac{1}{\beta}\mathsf{Tr}\mathcal{M}_Y\left(\frac{\mathcal{M}_x(\sigma)}{\alpha} - \mathcal{X}(\sigma)\right).\end{aligned}$$

The first term on the right is 1 since $\mathcal{M}_\mathcal{Y}$ and $\mathcal{X}$ commute as they act on disjoint sets of qubits. For the second term, we have using (4), (5) and the fact that an unnormalized superoperator cannot increase the trace norm, that

$$\left|\frac{1}{\beta}\mathsf{Tr}\mathcal{M}_Y\left(\frac{\mathcal{M}_X(\sigma)}{\alpha} - \mathcal{X}(\sigma)\right)\right| \leq \frac{\delta_a}{\beta} = \frac{\delta_b}{2}.$$

(b) Using (4), (5), the fact that a measurement or an unnormalized superoperator cannot increase the trace norm, and that $\mathcal{M}_{\mathcal{Y}}$ and $\mathcal{X}$ commute as they act on disjoint sets of qubits, we get

$$\begin{aligned}
\left\|\rho-\rho'\right\|_{\mathrm{tr}}\left\|\mathcal{X}\mathcal{Y}(\sigma)-\rho'\right\|_{\mathrm{tr}} &\leq \left\|\mathcal{X}\frac{\mathcal{M}_{\mathcal{Y}}(\sigma)}{\beta}-\rho'\right\|_{\mathrm{tr}}+\left\|\mathcal{X}\left(\mathcal{Y}(\sigma)-\frac{\mathcal{M}_{\mathcal{Y}}(\sigma)}{\beta}\right)\right\|_{\mathrm{tr}} \\
&\leq \left\|\mathcal{M}_{\mathcal{Y}}\frac{\mathcal{X}(\sigma)}{\beta}-\rho'\right\|_{\mathrm{tr}}+\delta_b \\
&\leq \left\|\frac{1}{\beta}\mathcal{M}_Y\frac{\mathcal{M}_X(\sigma)}{\alpha}-\rho'\right\|_{\mathrm{tr}}+\delta_b+\frac{1}{\beta}\left\|\mathcal{M}_Y\left(\mathcal{X}(\sigma)-\frac{\mathcal{M}_X(\sigma)}{\alpha}\right)\right\|_{\mathrm{tr}} \\
&\leq \left\|\frac{1}{\beta}\mathcal{M}_Y\frac{\mathcal{M}_X(\sigma)}{\alpha}-\rho'\right\|_{\mathrm{tr}}+\delta_b+\frac{\delta_a}{\beta} \\
&\leq \left\|\frac{r}{\alpha\beta}\frac{\mathcal{M}_Y\mathcal{M}_X(\sigma)}{r}-\rho'\right\|_{\mathrm{tr}}+\frac{3\delta_b}{2} \\
&= \left\|\left(\frac{r}{\alpha\beta}-1\right)\rho'\right\|_{\mathrm{tr}}+\frac{3\delta_b}{2} \\
&\leq 2\delta_b.
\end{aligned}$$

(c) Let $\tau$ describe the joint state of the input registers when the combined state of Alice and Bob is $\rho$; similarly, let $\tau'$ be the state of their input registers when the combined state is $\rho'$; thus,

$$\tau=\sum_{xy}p_{xy}|x\rangle\langle x|\otimes|y\rangle\langle y|$$

and

$$\tau'=\sum_{xy}p_{xy}\frac{r_{xy}}{r}|x\rangle\langle x|\otimes|y\rangle\langle y|.$$

Using part (b), we have

$$\sum_{xy}p_{xy}\left|1-\frac{r_{xy}}{r}\right|=\left\|\tau-\tau'\right\|_{\mathrm{tr}}\leq\left\|\rho-\rho'\right\|_{\mathrm{tr}}\leq 2\delta_b.$$

Thus, $\mathbb{E}_\mu\left[\left|\frac{r_{xy}}{r}-1\right|\right]\leq 2\delta_b$, and by Markov's inequality, $\Pr_\mu\left[\left|\frac{r_{xy}}{r}-1\right|\geq 2\delta_b^{1/2}\right]\leq\delta_b^{1/2}$. □

We can now move to the second step of our proof of Thm. 6. Figure 2 presents a protocol $\mathcal{P}''$ with $t-t'+1$ rounds of communication where the initial actions of Alice and Bob are derived from the protocol $\mathcal{P}'$ analyzed above.

*Claim 3.* (a) The number of bits sent by Alice in the first round is at most $k_a 2^{O(k_b/\delta^6)}$; the number of bits sent by Bob is at most $O(k_b/\delta^6)$.

(b) If the inputs are chosen according to the distribution $\mu$, the protocol $\mathcal{P}''$ computes $f$ correctly with probability of error at most $\epsilon+\delta$.

*Proof.* Recall that $\delta_b=(\delta/10)^2$, $\beta=2^{-O(k_b/\delta_b^3)}$ and $\delta_a=\delta_b\beta/2$ and $\alpha=2^{-O(k_a/\delta_a^3)}$. By part (a) of Claim 2 it follows that $r\geq\alpha\beta/2$. The number of bits needed by Alice to encode her set $S$ is at most

$$\log\binom{K}{2\alpha K}\leq 2\alpha K\log\left(\frac{e}{2\alpha}\right)=k_a 2^{O(k_b/\delta^6)}.$$

Alice *and* Bob *start with* $K \stackrel{\text{def}}{=} \frac{10}{r}(\log \frac{1}{\delta})$ *copies of* $\sigma$ *as prior entanglement. We refer to these copies as* $\sigma^1, \ldots, \sigma^K$.

**Input:** Alice gets $x \in X$ and Bob gets $y \in Y$.

Alice**:** Applies $\mathcal{M}_x$ to each $\sigma^i$. Let $\hat{S} \stackrel{\text{def}}{=} \{i : \mathcal{M}_x \text{ succeeded on } \sigma^i\}$. If $\hat{S}$ has less than $2\alpha K$ elements, Alice aborts the protocol; otherwise, she sends $S \subseteq \hat{S}$ to Bob, $|S| = 2\alpha K$.

Bob**:** Applies $\mathcal{M}_y$ to each $\sigma_i$ for $i \in S$ and sends Alice the index $i^*$ where he (and hence both) succeeded. If there is is no such $i^*$ he aborts the protocol.

Alice and Bob now revert to protocol $\mathcal{P}$ after round $t'$, and operate on the registers corresponding to $\sigma^{i^*}$.

**Fig. 2.** The final protocol $\mathcal{P}''$

The number of bits sent by Bob is at most $\log 2\alpha K = O\left(\frac{k_b}{\delta^6}\right)$. This justifies part (a) of our claim.

For part (b), we will use Claim 2 to bound the probability of error $\mathcal{P}''$. Call a pair $(x, y) \in \mathcal{X} \times \mathcal{Y}$ *good* if $|\frac{r_{xy}}{r} - 1| \leq 2\delta_b^{1/2}$; let $\chi$ denote the indicator random variable for the event "$(x, y)$ is good." Let $\chi'$ be the indicator random variable for the event "Alice and Bob do not abort protocol $\mathcal{P}''$." Note that if Alice and Bob do not abort protocol $\mathcal{P}''$, they enter round $t' + 1$ of protocol $\mathcal{P}$ with their registers in the state $\sigma'_{xy} \stackrel{\text{def}}{=} \frac{\mathcal{M}_x \mathcal{M}_y(\sigma)}{r_{xy}}$. Thus under distribution $\mu$, the average probability of error of $\mathcal{P}''$ differs from the average probability of error $\epsilon$ of the original protocol $\mathcal{P}$ by at most

$$\mathbb{E}_\mu \left[\chi\chi' \left\|\sigma'_{xy} - \sigma_{xy}\right\|_{\text{tr}}\right] + \Pr[\chi = 0] + \Pr[\chi = 1 \text{ and } \chi' = 0]. \tag{6}$$

The first term in the above sum can be bounded as follows:

$$\begin{aligned}
\mathbb{E}_\mu \left[\chi\chi' \left\|\sigma'_{xy} - \sigma_{xy}\right\|_{\text{tr}}\right] &= \mathbb{E}_\mu \left[\chi\chi' \left\|\frac{1}{r_{xy}}\mathcal{M}_x\mathcal{M}_y(\sigma) - \sigma_{xy}\right\|_{\text{tr}}\right] \\
&\leq \mathbb{E}_\mu \left[\chi\chi' \left\|\frac{1}{r}\mathcal{M}_x\mathcal{M}_y(\sigma) - \sigma_{xy}\right\|_{\text{tr}}\right] + \\
&\quad \mathbb{E}_\mu \left[\chi\chi' \left|1 - \frac{r_{xy}}{r}\right| \frac{1}{r_{xy}} \left\|\mathcal{M}_x\mathcal{M}_y(\sigma)\right\|_{\text{tr}}\right] \\
&\leq \left\|\frac{1}{r}\mathcal{M}_Y\mathcal{M}_X(\sigma) - \mathcal{X}\mathcal{Y}(\sigma)\right\|_{\text{tr}} + \\
&\quad \mathbb{E}_\mu \left[\chi\chi' \left|1 - \frac{r_{xy}}{r}\right| \frac{1}{r_{xy}} \left\|\mathcal{M}_x\mathcal{M}_y(\sigma)\right\|_{\text{tr}}\right] \\
&\leq 2\delta_b + 2\delta_b^{1/2}.
\end{aligned}$$

For the second last inequality, we used the fact that in the states $\sigma'_{xy}$ and $\sigma_{xy}$, the input registers of Alice and Bob contain $x$ and $y$. For the last inequality, we used part (b) of Claim 2 and the definition of good $(x, y)$. The second term of (6) is at most $\delta_b^{1/2}$ by part (c) of Claim 2. It remains to bound the last term of (6), which corresponds to the probability that Alice or Bob abort the protocol for some good $(x, y)$.

Alice **aborts:** The probability of success of $\mathcal{M}_x$ for any one copy of $\sigma$ is exactly $\alpha$. Thus, the expected number of successes is $\alpha K$, and by Chernoff's bound (see e.g. [AS00, Appendix A]), the probability that there are less than $2\alpha K$ successes is at most $\left(\frac{e}{4}\right)^{\alpha K} \leq \delta^{10}$.

Bob **aborts:** Bob aborts when the two parties do not simultaneously succeed in any of the $K$ attempts, even though their probability of success was at least $r_{xy} \geq (1 - 2\delta_b^{1/2})r \geq r/2$ (recall that we are now considering a good pair $(x, y)$). The probability of this is at most $\left(1 - \frac{r}{2}\right)^K \leq \exp\left(-\frac{rK}{2}\right) \leq \delta^5$.

Thus overall, the average probability of error of $\mathcal{P}''$ is at most

$$\epsilon + 2\delta_b + 2\delta_b^{1/2} + \delta_b^{1/2} + \delta^{10} + \delta^5 \leq \epsilon + \delta.$$

□

This completes the proof of Thm. 6. □

The following corollaries result from the above theorem.

**Corollary 2 (Privacy tradeoff).** *For any relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, $L(f, A, B)2^{O(L(f,B,A))} \geq \mathsf{Q}^{1,A\to B,\mathsf{pub},[\,]}(f)$. Similarly, $L(f, B, A)2^{O(L(f,A,B))} \geq \mathsf{Q}^{1,B\to A,\mathsf{pub},[\,]}(f)$.*

*Remark:* It was shown by Kremer [Kre95] that $Q(f) \geq \Omega(\log D^1(f))$, where $D^1(f)$ is the one-round deterministic communication complexity of $f$. The above corollary can be viewed as the privacy analogue of that result. It is optimal as evidenced by the Index function problem and the Pointer Chasing problem, both of which have communication complexity $O(\log n)$ [JRS02].

**Corollary 3 (Weak Direct Sum).** *For any relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$,*

$$\mathsf{Q}^{\mathsf{pub},[\,]}(f^{\oplus m}) \quad \geq \quad m \cdot \Omega(\log \mathsf{Q}^{1,\mathsf{pub},[\,]}(f)).$$

*Remark:* Jain, Radhakrishnan, and Sen [JRS03a,HJMR07] proved Direct Sum results for classical multi-round protocols. Their results were stronger because it avoided the logarithm. However, if we want a Direct Sum result independent of the number of rounds, the above is the best possible as evidenced by the Index function problem and the Pointer Chasing problem [JRS02].

### 4.2 Classical Protocols

Let $\mathcal{P}$ be a classical private-coins two-way protocol for a relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. Let $\mu_X, \mu_Y$ be probability distributions on $\mathcal{X}, \mathcal{Y}$, and let $\mu \stackrel{\mathsf{def}}{=} \mu_X \times \mu_Y$ denote a product distribution on $\mathcal{X} \times \mathcal{Y}$. Consider a run of $\mathcal{P}$, in which the inputs of Alice and Bob, are drawn according to distribution $\mu$. Let $X$ and $Y$ denote the random variables corresponding to the input of Alice and Bob respectively. Let $M$ denote the complete transcript of the messages sent by Alice and Bob during the protocol. Let $I(X : M)$ denote the mutual information between random variables $X$ and $M$ at the end of this run of $\mathcal{P}$.

**Definition 7 (Privacy loss).** *The* privacy loss *of $\mathcal{P}$ for relation $f$ on the product distribution $\mu$ from* Alice *to* Bob *is defined as $L^{\mathcal{P}}(f, \mu, A, B) \stackrel{\mathsf{def}}{=} I(X : M)$. The privacy loss from* Bob *to* Alice*, is defined similarly as $L^{\mathcal{P}}(f, \mu, B, A) \stackrel{\mathsf{def}}{=} I(Y : M)$.*

**Theorem 7.** *Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation and let $\epsilon \in (0, 1/2)$. Let $\mu$ be a product distribution on $\mathcal{X} \times \mathcal{Y}$. Let $\mathcal{P}$ be a private-coins protocol for $f$ with distributional error at most $\epsilon$ under $\mu$. Let us assume without loss of generality that* Alice *sends the first message and* Bob *computes the final answer. Let $L^{\mathcal{P}}(f, \mu, A, B) \leq k_a$ and $L^{\mathcal{P}}(f, \mu, B, A) \leq k_b$. Let $\tilde{\delta} > 0$ be such that $\epsilon + \tilde{\delta} \in (0, 1/2)$. Then there exists a one-round public-coin protocol (and hence also a deterministic protocol) $\tilde{\mathcal{P}}$ with single communication from* Alice*, such that,*

1. *Communication from* Alice *in $\tilde{\mathcal{P}}$ is $O\left(\frac{\log \frac{1}{\tilde{\delta}}}{\tilde{\delta}^3} \cdot (k_a + 1) \cdot 2^{O((k_b+1)/\tilde{\delta}^2)}\right)$.*
2. *The distributional error of $\tilde{\mathcal{P}}$ under $\mu$ is at most $\epsilon + \tilde{\delta}$.*

*Proof.* Let the marginals of $\mu$ on $\mathcal{X}, \mathcal{Y}$ be $\mu_X, \mu_Y$ respectively. Therefore $\mu = \mu_X \otimes \mu_Y$. Let the distribution of $M$ (the combined message transcript in $\mathcal{P}$), when $X = x$ and $Y = y$, be $P_{x,y}$. Let $P_x \stackrel{\mathsf{def}}{=} \mathbb{E}_{y \leftarrow \mu_Y}[P_{x,y}]$, $P_y \stackrel{\mathsf{def}}{=} \mathbb{E}_{x \leftarrow \mu_X}[P_{x,y}]$ and $P \stackrel{\mathsf{def}}{=} \mathbb{E}_{(x,y) \leftarrow \mu}[P_{x,y}]$. Let there be $k$ messages in protocol $\mathcal{P}$. Let $M_1, M_2, \ldots M_k$ denote the random variables corresponding to the first, second and so on till the $k$-th message of the protocol $\mathcal{P}$. Let $S$ be the set of all message strings $s$. For $s \in S$, let $s_1, s_2, \ldots, s_k$ denote the parts corresponding to $M_1, M_2, \ldots M_k$ respectively. For $i \in [k]$, let $p^{x,y}(s, i)$ denote the probability with which $s_i$ appears in $P_{x,y}$ conditioned on the first $i-1$ messages as being $s_1, s_2, \ldots s_{i-1}$. Similarly we define $p^x(s, i)$, $p^y(s, i)$ and $p(s, i)$ corresponding to distributions $P_x$, $P_y$ and $P$. Let $p^{x,y}(s)$ denote the probability with which message $s$ appears in $P_{x,y}$. Similarly let us define $p^x(s)$, $p^y(s)$ and $p(s)$ corresponding to distributions $P_x$, $P_y$ and $P$. Now we have the following claim.

*Claim 4.* For all $x \in \mathcal{X}, y \in \mathcal{Y}, s \in S$,

$$p^x(s) \cdot p^y(s) \quad = \quad p(s) \cdot p^{x,y}(s).$$

*Proof.* Note that since $\mathcal{P}$ is a private coins protocol and Bob sends even numbered messages, we have for all even $i, \forall x \in \mathcal{X}, \forall s \in S, p^x(s, i) = p(s, i)$. Therefore $\forall x \in \mathcal{X}, \forall s \in S$,

$$\frac{p^x(s)}{p(s)} = \frac{\prod_{i=1}^{k} p^x(s, i)}{\prod_{i=1}^{k} p(s, i)} = \frac{\prod_{i:\mathsf{odd}} p^x(s, i)}{\prod_{i:\mathsf{odd}} p(s, i)}. \tag{7}$$

Similarly we have for all odd $i, \forall y \in \mathcal{Y}, \forall s \in S, p^y(s, i) = p(s, i)$ and hence,

$$\frac{p^y(s)}{p(s)} = \frac{\prod_{i:\mathsf{even}} p^y(s, i)}{\prod_{i:\mathsf{even}} p(s, i)}. \tag{8}$$

We can note further that for $\forall x \in \mathcal{X}, \forall y \in \mathcal{Y}, \forall s \in S$; for all odd $i, p^{x,y}(s, i) = p^x(s, i)$ and for all even $i, p^{x,y}(s, i) = p^y(s, i)$. Therefore,

$$p^{x,y}(s) = \prod_{i=1}^{k} p^{x,y}(s, i) = \prod_{i:\mathsf{odd}} p^x(s, i) \cdot \prod_{i:\mathsf{even}} p^y(s, i). \tag{9}$$

Our claim now follows by combining Eq. (7), Eq. (8) and Eq. (9). □

Let $\delta = \frac{\tilde{\delta}}{5}$. Since $k_a \geq I(M : X) = \mathbb{E}_{x \leftarrow \mu_X}[S(P_x || P)]$, using Markov's inequality we get a set $\mathsf{Good}_X \subseteq \mathcal{X}$ such that

$$\Pr_{\mu_X}[x \in \mathsf{Good}_X] \geq 1 - \delta \quad \text{and} \quad \forall x \in \mathsf{Good}_X, S(P_x || P) \leq \frac{k_a}{\delta}. \tag{10}$$

Let $x \in \mathsf{Good}_X$. Since $\frac{k_a}{\delta} \geq S(P_x||P) = \mathbb{E}_{s \leftarrow P_x}\left[\log \frac{p^x(s)}{p(s)}\right]$, using Lem. 3, we get a set $\mathsf{Good}^x \subseteq S$ such that

$$\Pr_{P_x}[s \in \mathsf{Good}^x] \geq 1 - \delta \quad \text{and} \quad \forall s \in \mathsf{Good}^x, \frac{p^x(s)}{p(s)} \leq 2^{\frac{k_a+1}{\delta^2}}. \tag{11}$$

Similarly there exists a set $\mathsf{Good}_Y \subseteq \mathcal{Y}$ such that

$$\Pr_{\mu_Y}[y \in \mathsf{Good}_Y] \geq 1 - \delta \quad \text{and} \quad \forall y \in \mathsf{Good}_Y, S(P_y||P) \leq \frac{k_b}{\delta}. \tag{12}$$

Similarly for $y \in \mathsf{Good}_Y$, there exists a set $\mathsf{Good}^y \subseteq S$ such that

$$\Pr_{P_y}[s \in \mathsf{Good}^y] \geq 1 - \delta \quad \text{and} \quad \forall s \in \mathsf{Good}^y, \frac{p^y(s)}{p(s)} \leq 2^{\frac{k_b+1}{\delta^2}}. \tag{13}$$

Let us now present an intermediate protocol $\mathcal{P}'$ in Fig. 3 from which we will finally obtain our desired protocol $\tilde{\mathcal{P}}$.

Alice *and* Bob*, using shared prior randomness, generate an array of strings (each string belonging to the set* $S$*) with infinite columns and* $K \stackrel{\text{def}}{=} \left(\frac{1}{1-\delta} \cdot \ln \frac{1}{\delta}\right) \cdot 2^{(k_b+1)/\delta^2}$ *rows. Each string in the array is sampled independently according to the distribution* $P$*. Let the random variables representing various strings be* $S^{i,j}, i \in [K], j \in \mathbb{N}$ *(*$\mathbb{N}$ *is the set of natural numbers).*

**Input:** Alice gets $x \in X$ and Bob gets $y \in Y$.

Alice**:** She sets $i = 1, j = 1$.

1. In case $x \notin \mathsf{Good}_X$, she aborts the protocol and sends a special abort message to Bob (using constant number of bits). Otherwise she moves to step 2.
2. She considers string $S^{i,j}$. In case $S^{i,j} \in \mathsf{Good}_x$, she accepts $S^{i,j}$ with probability $\frac{1}{2^{(k_a+1)/\delta^2}} \cdot \frac{p^x(S^{i,j})}{p(S^{i,j})}$. In case $S^{i,j} \notin \mathsf{Good}_x$, she accepts $S^{i,j}$ with probability 0.
3. In case she accepts $S^{i,j}$, she communicates $j$ to Bob using a prefix free binary encoding. If $i = K$, she stops, otherwise she sets $i = i + 1, j = 1$ and goes to step 2. In case she rejects $S^{i,j}$, she sets $j = j + 1$ and moves to step 2.

Let the various index communicated to Bob be denoted $J_i, i \in [K]$.

Bob**:** He sets $l = 1$. If he gets abort message from Alice, he aborts the protocol, otherwise he goes to step 1.

1. If $y \notin \mathsf{Good}_y$, he aborts the protocol. Otherwise he goes to step 2.
2. He considers the string $S^{l,J_l}$, where $J_l$ is as obtained from Alice. If $S^{l,J_l} \in \mathsf{Good}_y$, he accepts $S^{l,J_l}$ with probability $\frac{1}{2^{(k_b+1)/\delta^2}} \cdot \frac{p^y(S^{l,J_l})}{p(S^{l,J_l})}$. If $S^{l,J_l} \notin \mathsf{Good}_y$, he accepts $S^{l,J_l}$ with probability 0.
3. In case he accepts $S^{l,J_l}$, he considers it to be the the final message transcript $M$ of protocol $\mathcal{P}$ and simulates $\mathcal{P}$ from now on to output $z \in \mathcal{Z}$. In case he rejects $S^{l,J_l}$, if $l = K$ he aborts the protocol, otherwise he sets $l = l + 1$ and goes to step 2.

**Fig. 3.** The intermediate protocol $\mathcal{P}'$

Protocol $\mathcal{P}'$ is clearly one-way protocol. Now let us now analyze the expected communication from Alice to Bob in $\mathcal{P}'$ and expected error of $\mathcal{P}'$.

**Expected communication of $\mathcal{P}'$:** When $x \notin \mathsf{Good}_X$, there is constant communication. Let $x \in \mathsf{Good}_X$, and fix $i \in [K]$. Then the probability that $J_i = j$ given that the previous samples were rejected in the row $i$, is:

$$\begin{aligned}
&\sum_{s \in S} \Pr(S^{i,j} = s) \cdot \Pr(s \text{ is accepted}) \\
&= \sum_{s \in \mathsf{Good}_x} \Pr(S^{i,j} = s) \cdot \Pr(s \text{ is accepted}) + \sum_{s \notin \mathsf{Good}_x} \Pr(S^{i,j} = s) \cdot \Pr(s \text{ is accepted}) \\
&= \sum_{s \in \mathsf{Good}_x} \Pr(S^{i,j} = s) \cdot \Pr(s \text{ is accepted}) + 0 \\
&= \sum_{s \in \mathsf{Good}_x} p(s) \cdot \frac{1}{2^{(k_a+1)/\delta^2}} \cdot \frac{p^x(s)}{p(s)} = \frac{1}{2^{(k_a+1)/\delta^2}} \cdot \Pr_{P_x}(s \in \mathsf{Good}_x) \geq \frac{1-\delta}{2^{(k_a+1)/\delta^2}}.
\end{aligned}$$

The last inequality follows from Eq. (11). Therefore expected value of $J_i$ is $\frac{2^{(k_a+1)/\delta^2}}{1-\delta}$. Therefore, from concavity of the log function it follows that the expected communication from $\mathsf{Alice}$ to communicate $J_i$ to $\mathsf{Bob}$ (using a prefix free binary encoding) is $O(\log \frac{2^{(k_a+1)/\delta^2}}{1-\delta})$. This is true for every $i \in [K]$. Therefore for $x \in \mathsf{Good}_x$, expected communication from $\mathsf{Alice}$ is $O\left(\frac{\log \frac{1}{\delta}}{\delta^2} \cdot (k_a+1) \cdot 2^{O((k_b+1)/\delta^2)}\right)$. Therefore overall expected communication from $\mathsf{Alice}$ is $O\left(\frac{\log \frac{1}{\delta}}{\delta^2} \cdot (k_a+1) \cdot 2^{O((k_b+1)/\delta^2)}\right)$.

**Expected error of $\mathcal{P}'$:** $\mathsf{Alice}$ aborts the protocol when $x \notin \mathsf{Good}_X$, which happens with probability at most $\delta$. Assume that $\mathsf{Alice}$ does not abort. $\mathsf{Bob}$ aborts the protocol when $y \notin \mathsf{Good}_Y$, which happens with probability at most $\delta$. When $y \in \mathsf{Good}_Y$, using a similar calculation as above we can conclude that $\mathsf{Bob}$ accepts the $l$-th sample (for any $l \in [K]$), given that he has rejected the samples before is at least $\frac{1-\delta}{2^{(k_b+1)/\delta^2}}$. Therefore,

$$\Pr(\mathsf{Bob} \text{ rejects all } K \text{ samples}) \leq \left(1 - \frac{1-\delta}{2^{(k_b+1)/\delta^2}}\right)^K \leq \exp(-K \cdot \frac{1-\delta}{2^{(k_b+1)/\delta^2}}) = \delta.$$

Therefore, when $(x, y) \in \mathsf{Good}_X \times \mathsf{Good}_Y$,

$$\Pr(\mathsf{Bob} \text{ aborts given input of } \mathcal{P}' \text{ is } (x, y)) \leq \delta.$$

We have the following claim.

*Claim 5.* Let $(x, y) \in \mathsf{Good}_X \times \mathsf{Good}_Y$ and $\mathsf{Bob}$ does not abort. Then,

1. If $s \in \mathsf{Good}_x \cap \mathsf{Good}_y$ then $\Pr(\mathsf{Bob} \text{ sets } M = s) = \frac{p^{x,y}(s)}{\Pr(s \in \mathsf{Good}_x \cap \mathsf{Good}_y)}$.
2. If $s \notin \mathsf{Good}_x \cap \mathsf{Good}_y$ then $\Pr(\mathsf{Bob} \text{ sets } M = s) = 0$.

We defer the proof of this claim to later. Let us now analyze the expected error of the protocol $\mathcal{P}'$ assuming Claim 5 to be true. Let $\epsilon'_{x,y}$ stand for error of $\mathcal{P}'$ when input is $(x, y)$. From above claim, if $(x, y) \in \mathsf{Good}_X \times \mathsf{Good}_Y$ and $\mathsf{Bob}$ does not abort, then the $\ell_1$ distance between the distribution of $M$ in $\mathcal{P}'$ and $P_{x,y}$ is $2(1 - \Pr(s \in \mathsf{Good}_x \cap \mathsf{Good}_y)) \leq 2\delta$. Therefore if $(x, y) \in \mathsf{Good}_X \times \mathsf{Good}_Y$ and $\mathsf{Bob}$ does not abort, then $\epsilon'_{x,y} \leq \epsilon_{x,y} + \delta$, where $\epsilon_{x,y}$ is the error of $\mathcal{P}$ on input $(x, y)$. Therefore, for $(x, y) \in \mathsf{Good}_X \times \mathsf{Good}_Y$,

$$\Pr(\mathcal{P}' \text{ errs on input } (x, y) \text{ given } \mathsf{Bob} \text{ does not abort}) \leq \epsilon_{x,y} + \delta.$$

This implies:

$$\begin{aligned}
&\mathbb{E}_{(x,y)\in \mathsf{Good}_X \times \mathsf{Good}_Y}[\Pr(\mathcal{P}' \text{ errs on input } (x,y) \text{ given } \mathsf{Bob} \text{ does not abort})] \\
&\leq \mathbb{E}_{(x,y)\in \mathsf{Good}_X \times \mathsf{Good}_Y}[\epsilon_{x,y}] + \delta \\
&\leq \frac{1}{\Pr((x,y) \in \mathsf{Good}_X \times \mathsf{Good}_Y)} \mathbb{E}_{(x,y)\in \mathcal{X}\times\mathcal{Y}}[\epsilon_{x,y}] + \delta \\
&\leq \frac{\epsilon}{1-2\delta} + \delta
\end{aligned}$$

$$\begin{aligned}
&\text{Expected error of } \mathcal{P}' \\
&\leq \Pr(x \notin \mathsf{Good}_X) + \Pr(y \notin \mathsf{Good}_Y) + \Pr((x,y) \in \mathsf{Good}_X \times \mathsf{Good}_Y) \cdot \mathbb{E}_{x\in \mathsf{Good}_X, y\in \mathsf{Good}_Y}[\epsilon'_{x,y}] \\
&\leq \delta + \delta \\
&+ \Pr((x,y) \in \mathsf{Good}_X \times \mathsf{Good}_Y) \cdot (\mathbb{E}_{x\in \mathsf{Good}_X, y\in \mathsf{Good}_Y}[\Pr(\mathsf{Bob} \text{ aborts given input of } \mathcal{P}' \text{ is } (x,y))] \\
&+ \mathbb{E}_{x\in \mathsf{Good}_X, y\in \mathsf{Good}_Y}[\Pr(\mathcal{P}' \text{ errs given } x \in \mathsf{Good}_X, y \in \mathsf{Good}_Y \text{ and } \mathsf{Bob} \text{ does not abort})]) \\
&\leq 2\delta + \delta + (\Pr((x,y) \in \mathsf{Good}_X \times \mathsf{Good}_Y) \cdot \left( \frac{\epsilon}{\Pr((x,y) \in \mathsf{Good}_X \times \mathsf{Good}_Y)} + \delta \right) \\
&\leq 4\delta + \epsilon.
\end{aligned}$$

□

We are now finally ready to describe the protocol $\tilde{\mathcal{P}}$.

*Let $c$ be the expected communication from* Alice *to* Bob *in protocol $\mathcal{P}'$.*

**Input:** Alice gets $x \in X$ and Bob gets $y \in Y$.
**Alice:** She simulates protocol $\mathcal{P}'$. If for some choice of the public coins the bits needed to communicate all $J_i, i \in [K]$ exceeds $c/\delta$, she aborts the protocol and sends a special abort message to Bob in constant bits.
**Bob:** In case he does not get abort message from Alice, he proceeds as in protocol $\mathcal{P}'$.

**Fig. 4.** The final protocol $\tilde{\mathcal{P}}$

Now it is clear that the communication of $\tilde{\mathcal{P}}$ is as claimed. Also it is easily noted that the expected error of $\tilde{\mathcal{P}}$ is at most expected error of $\mathcal{P}'$ plus $\delta$ which is $\epsilon + 5\delta = \epsilon + \tilde{\delta}$ as claimed (since $\delta = \frac{\tilde{\delta}}{5}$).

**Proof of Claim 5:** Let $l \in [K]$. Then conditioned on $\mathsf{Bob}$ rejecting first $l-1$ samples, for $s \in \mathsf{Good}_x \cap \mathsf{Good}_y$,

$$\begin{aligned}
&\Pr(\mathsf{Bob}\text{'s outputs } S^{l,J_l} \text{ and } S^{l,j_l} = s) \\
&= \Pr(S^{l,j_l} = s) \cdot \Pr(\mathsf{Alice} \text{ accepts } S^{l,j_l}) \cdot \Pr(\mathsf{Bob} \text{ accepts } S^{l,j_l}) \\
&= p(s) \cdot \frac{p^x(s)}{2^{(k_a+1)/\delta^2} p(s)} \cdot \frac{p^y(s)}{2^{(k_b+1)/\delta^2} p(s)} \\
&= \frac{p^{x,y}(s)}{2^{(k_a+k_b+2)/\delta^2}}.
\end{aligned}$$

Therefore conditioned on $\mathsf{Bob}$ rejecting first $l-1$ samples,

$$\begin{aligned}
\Pr(\mathsf{Bob}\text{'s outputs } S^{l,j_l}) &= \sum_{s \in S} \Pr(\mathsf{Bob}\text{'s outputs } S^{l,j_l} \text{ and } S^{l,j_l} = s) \\
&= \sum_{s \in \mathsf{Good}_x \cap \mathsf{Good}_y} \Pr(\mathsf{Bob}\text{'s outputs } S^{l,j_l} \text{ and } S^{l,j_l} = s) \\
&= \frac{1}{2^{(k_a+k_b+2)/\delta^2}} \cdot \Pr(s \in \mathsf{Good}_x \cap \mathsf{Good}_y).
\end{aligned}$$

Therefore, conditioned on $\mathsf{Bob}$ rejecting first $l-1$ samples, for $s \in \mathsf{Good}_x \cap \mathsf{Good}_y$,

$$\begin{aligned}
\Pr(\mathsf{Bob}\text{'s outputs } s \text{ given } \mathsf{Bob} \text{ outputs } S^{l,j_l}) &= \frac{\Pr(\mathsf{Bob}\text{'s outputs } S^{l,j_l} \text{ and } S^{l,j_l} = s)}{\Pr(\mathsf{Bob}\text{'s outputs } S^{l,j_l})} \\
&= \frac{p^{x,y}(s)}{\Pr(s \in \mathsf{Good}_x \cap \mathsf{Good}_y)}
\end{aligned}$$

Clearly for $s \notin \mathsf{Good}_x \cap \mathsf{Good}_y, \Pr(\mathsf{Bob}\text{'s outputs } s \text{ given } \mathsf{Bob} \text{ outputs } S^{l,j_l}) = 0$. Our claim now immediately follows. □

As before we get the following corollaries from the above theorem.

**Corollary 4 (Privacy tradeoff).** *For any relation $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$, $L(f,A,B)2^{O(L(f,B,A))} \geq \mathsf{R}^{1,A\to B,[\,]}(f)$. Similarly, $L(f,B,A)2^{O(L(f,A,B))} \geq \mathsf{R}^{1,B\to A,[\,]}(f)$.*

**Corollary 5 (Weak Direct Sum).** *For any relation $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$,*

$$\mathsf{R}^{[\,]}(f^{\oplus m}) \quad \geq \quad m \cdot \Omega(\log \mathsf{R}^{1,[\,]}(f)).$$

## 5 Entanglement Reduction

We will need the following geometric result. It is similar to a result proved earlier in [JRS03a].

**Lemma 5.** *Suppose $M$, $N$ are positive integers with $M = \Theta(N^{2/3} \log N)$. Let the underlying Hilbert space be $\mathbb{C}^M$. There exist $16N$ subspaces $V_{ij} \leq \mathbb{C}^M$, $1 \leq i \leq N$, $1 \leq j \leq 16$, each of dimension $\frac{M}{16}$, such that if we define $\Pi_{ij}$ to be the orthogonal projection onto $V_{ij}$ and $\rho_{ij} \stackrel{\mathsf{def}}{=} \frac{16}{M} \cdot \Pi_{ij}$, then*

1. $\forall i,j \; \mathsf{Tr}(\Pi_{ij}\rho_{ij}) = 1$.
2. $\forall i,j,i',j', \; i \neq i', \; \mathsf{Tr}(\Pi_{ij}\rho_{i'j'}) < 1/4$.

3. $\forall i, j, j',\ j \neq j',\ \mathsf{Tr}(\Pi_{ij}\rho_{ij'}) = 0$.
4. $\forall i,\ I_M = \sum_{j=1}^{16} \Pi_{ij}$, *where* $I_M$ *is the identity operator on* $\mathbb{C}^M$.
5. *For all subspaces* $W$ *of dimension at most* $N^{1/6}$, *for all families of density matrices* $\{\sigma_{ij}\}_{i\in[N],1\leq j\leq 16}$, $\sigma_{ij}$ *supported in* $W$,

$$|\{i : \exists j,\ 1 \leq j \leq 16,\ \mathsf{Tr}(\Pi_{ij}\sigma_{ij}) > 9/16\}| \leq N/4.$$

*Proof.* (**Sketch**) The proof follows by combining the proofs of Thm. 5 and Lem. 7 of [JRS03a]. We skip a full proof for brevity. □

We shall also need the following easy proposition.

**Proposition 2.** *Let* $|\phi\rangle_{AB}$ *be a bipartite pure quantum state. Define* $e \stackrel{\mathsf{def}}{=} E(|\phi\rangle)$. *Then there is a bipartite pure quantum state* $|\phi'\rangle_{AB}$ *having Schmidt rank at most* $2^{100e}$ *such that* $\||\phi\rangle\langle\phi| - |\phi'\rangle\langle\phi'|\|_{\mathrm{tr}} \leq 1/20$.

*Proof.* Let $|\phi\rangle_{AB} = \sum_i \sqrt{\lambda_i}|a_i\rangle_A|b_i\rangle_B$ be the Schmidt decomposition of $|\phi\rangle$, $\lambda_i \geq 0$, $\sum_i \lambda_i = 1$. Define a set $\mathsf{Good} \stackrel{\mathsf{def}}{=} \{i : \lambda_i \geq 2^{-100e}\}$. Since $e = -\sum_i \lambda_i \log \lambda_i$, by Markov's inequality $\sum_{i\in\mathsf{Good}} \lambda_i \geq 99/100$. Define the bipartite pure state $|\phi'\rangle_{AB} \stackrel{\mathsf{def}}{=} \sum_{i\in\mathsf{Good}} \sqrt{\lambda_i}|a_i\rangle_A|b_i\rangle_B$ normalized. The Schmidt rank of $|\phi'\rangle_{AB}$ is at most $2^{100e}$ and $\||\phi\rangle\langle\phi| - |\phi'\rangle\langle\phi'|\|_{\mathrm{tr}} \leq 1/20$. □

We are now ready to prove our impossibility result about black-box reduction of prior entanglement.

**Theorem 8 (No black-box red. of prior entan.).** *Let* $\mathrm{EQ}_n$ *denote the Equality function on* $n$*-bit strings. There exists a one-round quantum protocol* $\mathcal{P}$ *for* $\mathrm{EQ}_n$ *with* $\frac{2n}{3} + \log n + \Theta(1)$ *EPR pairs of prior entanglement and communicating* 4 *bits such that, there is no similar protocol* $\mathcal{P}'$ *that starts with a prior entangled state* $|\phi\rangle$, $E(|\phi\rangle) \leq \frac{n}{600}$.

*Proof.* We use the notation of Lem. 5 with $M \stackrel{\mathsf{def}}{=} 2^m$ and $N \stackrel{\mathsf{def}}{=} 2^n$. Let $0 \leq i \leq 2^n - 1$ i.e. $i \in \{0,1\}^n$. Choose $m = \frac{2n}{3} + \log n + \Theta(1)$. Let $\mathcal{P}$ be a one-round protocol with $m$ EPR pairs of prior entanglement. In $\mathcal{P}$, on input $i$ Alice measures her EPR halves according to the von-Neumann measurement $\{\Pi_j\}_{1\leq j\leq 16}$ and sends the result $j$ as a 4-bit classical message to Bob. The state of Bob's EPR halves now becomes $\rho_{ij}$. On input $i'$ and message $j'$, Bob performs the two-outcome measurement $\{\Pi_{i'j'}, I_M - \Pi_{i'j'}\}$ on his EPR halves. Therefore in $\mathcal{P}$, Bob outputs 1 with probability 1 if $i' = i$ and with probability at most 1/4 if $i' \neq i$. Thus, $\mathcal{P}$ is a protocol for $\mathrm{EQ}_n$.

Suppose there exists a protocol $\mathcal{P}'$ similar to $\mathcal{P}$ that starts with an input independent shared state $|\phi'\rangle_{AB}$ on $m+m$ qubits. Suppose $E(|\phi\rangle) \leq n/10$. By Proposition 2, there is a bipartite pure state $|\phi''\rangle_{AB}$ on $m+m$ qubits having Schmidt rank at most $2^{n/6}$ such that $\||\phi'\rangle\langle\phi'| - |\phi''\rangle\langle\phi''|\|_{\mathrm{tr}} \leq 1/20$. Consider the protocol $\mathcal{P}''$ similar to $\mathcal{P}'$ starting with $|\phi''\rangle_{AB}$ as prior entanglement. Since $\mathcal{P}''$ is similar to $\mathcal{P}'$, it is also a one-round protocol with 4 classical bits of communication. Let $\sigma_{ij}$ be the state of Bob's share of prior entanglement qubits after the first round of communication from Alice when Alice's input is $i$ and her message is $j$. Since the Schmidt rank of $|\phi''\rangle$ is at most $2^{n/6}$, the $\sigma_{ij}$, $0 \leq i \leq 2^n - 1$, $1 \leq j \leq 16$ have support in a $2^{n/6}$-dimensional space. Let $p_{ij}$ be the probability with which Alice sends message $j$ when her input is $i$. It follows that for all $i$, $\sum_{j=1}^{16} p_j \mathsf{Tr} M_{ij}\sigma_{ij} \geq \frac{3}{4} - \frac{1}{20} - \frac{1}{20} = \frac{13}{20}$. This implies that for all $i$ there exists a $j$, $1 \leq j \leq 16$, such that $\mathsf{Tr} M_{ij}\sigma_{ij} \geq 13/20 > 9/16$. From Lem. 5 this is not possible, and hence no such protocol $\mathcal{P}'$ exists. □

## 6 Exact Remote State Preparation

**Proof of Thm. 2:** We start with the following lemma which may be of independent interest.

**Lemma 6.** *Let $\rho \stackrel{\mathsf{def}}{=} |\phi\rangle\langle\phi| \in \mathcal{H}$ be a pure state and $\sigma \in \mathcal{H}$ be any positive definite matrix. Then the maximum value of $k$ such that, $\sigma - k\rho \geq 0$, is $(\langle\phi|\sigma^{-1}|\phi\rangle)^{-1}$.*

*Proof.* First we show that, $\langle\phi|\sigma^{-1}|\phi\rangle\sigma - \rho \geq 0$. Let $|v\rangle \in \mathcal{H}$. Let $|w_1\rangle \stackrel{\mathsf{def}}{=} \sigma^{-1/2}|\phi\rangle$ and $|w_2\rangle \stackrel{\mathsf{def}}{=} \sigma^{1/2}|v\rangle$. Now Cauchy-Schwartz inequality implies,

$$\begin{aligned} \langle w_1|w_1\rangle\langle w_2|w_2\rangle &\geq |\langle w_1|w_2\rangle|^2 \\ \Rightarrow \langle\phi|\sigma^{-1}|\phi\rangle\langle v|\sigma|v\rangle &\geq |\langle\phi|v\rangle|^2 \\ \Rightarrow \langle v|\langle\phi|\sigma^{-1}|\phi\rangle\sigma|v\rangle &\geq \langle v|\phi\rangle\langle\phi|v\rangle \\ \Rightarrow \langle v|(\langle\phi|\sigma^{-1}|\phi\rangle\sigma - |\phi\rangle\langle\phi|)|v\rangle &\geq 0 \end{aligned}$$

Now since above is true for every $|v\rangle \in \mathcal{H}$ we have that $\langle\phi|\sigma^{-1}|\phi\rangle\sigma - |\phi\rangle\langle\phi| \geq 0$.
Next we show that if $k > (\langle\phi|\sigma^{-1}|\phi\rangle)^{-1}$ then $\sigma - |\phi\rangle\langle\phi|$ is not positive semi-definite. For this let $|v\rangle \stackrel{\mathsf{def}}{=} \sigma^{-1}|\phi\rangle$, and in this case $|w_1\rangle = |w_2\rangle$. Now since $\sigma \geq 0$ and $k > (\langle\phi|\sigma^{-1}|\phi\rangle)^{-1}$ we have,

$$\begin{aligned} \langle v|(k^{-1}\sigma - |\phi\rangle\langle\phi|)|v\rangle &< \langle v|(\langle\phi|\sigma^{-1}|\phi\rangle\sigma - |\phi\rangle\langle\phi|)|v\rangle \\ &= \langle\phi|\sigma^{-1}|\phi\rangle\langle v|\sigma|v\rangle - |\langle\phi|v\rangle|^2 \\ &= \langle w_1|w_1\rangle\langle w_2|w_2\rangle - |\langle w_1|w_2\rangle|^2 \\ &= 0 \end{aligned}$$

Hence $k^{-1}\sigma - |\phi\rangle\langle\phi|$ is not positive semi-definite. □

Let $\rho \stackrel{\mathsf{def}}{=} |\phi\rangle\langle\phi|$, $\sigma$ be some full rank state and let $k = (\langle\phi|\sigma^{-1}|\phi\rangle)^{-1}$. Let $\rho' \stackrel{\mathsf{def}}{=} \sigma - \langle\phi|\sigma^{-1}|\phi\rangle)^{-1}\rho$. Lem. 6 implies $\rho' \geq 0$. Let $\mathcal{K}$ be a Hilbert space with $dim(\mathcal{K}) = dim(\mathcal{H})$. Let $|\theta\rangle \in \mathcal{K} \otimes \mathcal{H}$ be some purification of $\rho'$ and $|\bar{0}\rangle$ be a fixed vector in $\mathcal{K}$. We now define,

$$|\psi\rangle_\rho \stackrel{\mathsf{def}}{=} \sqrt{k}|1\rangle|\bar{0}\rangle|\phi\rangle + \sqrt{1-k}|0\rangle|\theta\rangle$$

We note that the marginal of $|\psi\rangle_\rho$ in $\mathcal{H}$ is $\sigma$.

We have the following lemma due to Jozsa and Uhlmann [Joz94,Uhl76].

**Lemma 7 (Local transition).** *Let $\rho$ be a quantum state in $\mathcal{H}$. Let $|\phi_1\rangle$ and $|\phi_2\rangle$ be two purification of $\rho$ in $\mathcal{K} \otimes \mathcal{H}$. There is a local unitary transformation $U$ acting on $\mathcal{K}$ such that $(U \otimes I)|\phi_1\rangle = |\phi_2\rangle$.*

Now consider the following protocol $\mathcal{P}$:

1. Alice and Bob start with several copies of a fixed pure state $|\psi\rangle$ such that marginal on Bob's side in $|\psi\rangle$ is $\sigma$.
2. On getting $x$, Alice transforms using a local unitary the first copy of $|\psi\rangle$ to $|\psi\rangle_{\rho_x}$. This can be done using Lemma 7, since the marginal on Bob's side in both $|\psi\rangle$ and $|\psi\rangle_{\rho_x}$ is $\sigma$. She then measures the first qubit.
3. She keeps doing this to successive copies of $|\psi\rangle$ until she gets the first 1 on measurement. She communicates to Bob the first occurrence of 1.

From the definition of $|\psi\rangle_{\rho_x}$, we note that in the copy in which Alice gets 1, Bob ends up with $\rho_x$. Also, (from concavity of the log function) it can be verified that, using a prefix-free encoding of integers that requires $\log n + 2\log\log n$ bits to encode the integer $n$, the expected communication of Alice is bounded by $\log(\mathsf{Tr}\sigma^{-1}\rho_x) + 2\log\log(\mathsf{Tr}\sigma^{-1}\rho_x)$. Hence our theorem. □

*Remarks:*

1. For any fixed state $\sigma$ of full rank, from the above proof, we get a protocol $\mathcal{P}_\sigma$ such that given the description of any pure state $\rho$ to Alice, she ends up creating $\rho$ with Bob with communication $\log(\mathsf{Tr}\sigma^{-1}\rho)$.
2. We note that when $\rho_x \stackrel{\mathsf{def}}{=} |\phi_x\rangle\langle\phi_x|$ then from concavity of log function we have, $S(\rho_x||\sigma) = \langle\phi_x|\log\sigma|\phi\rangle \leq \log\langle\phi|\sigma^{-1}|\phi_x\rangle$. Therefore the approach that we take here, which is analogous to the *rejection sampling* approach of [HJMR07], does not help us in getting the communication down to $S(\rho_x||\sigma)$ which happens in [HJMR07] for a similar problem in the classical setting.
3. It is open as to whether the communication could be brought down to $S(\rho_x||\sigma)$. Also the case when $\rho_x$ is not necessarily a pure state is open.
4. The inexact version of this problem was considered in [Jai06] where some fidelity loss in generating $\rho_x$ was allowed. There using the substate theorem, the task was accomplished with communication $S(\rho_x||\sigma)/\epsilon$ at the end of which Bob got a state $\rho'_x$ which was $\epsilon$ close in trace distance to $\rho_x$ (not necessarily pure).


### Acknowledgment

We thank the referees for their comments and suggestions. We are grateful to Harold Ollivier for his comments on the proof of Thm. 8.


## References


[AKN98] D. Aharonov, A. Kitaev, and N. Nisan. Quantum circuits with mixed states. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, pages 20–30, 1998. Also quant-ph/9806029.

[AL70] H. Araki and E.H. Lieb. Entropy inequalities. *Comm. Math. Phys.*, 18:160–170, 1970.

[ANTV02] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani. Dense quantum coding and quantum finite automata. *Journal of the ACM*, 49(4):496–511, 2002.

[AS00] N. Alon and J. Spencer. *The probabilistic method*. John Wiley and Sons, 2000.

[BCKO93] R. Bar-Yehuda, B. Chor, E. Kushilevitz, and A. Orlitsky. Privacy, additional information, and communication. *IEEE Transactions on Information Theory*, 39(6):1930–1943, 1993.

[BW92] C. Bennett and S. Wiesner. Communication via one and two particle operators on Einstein-Podolsky-Rosen states. In *Phys. Rev. Lett.*, volume 69, pages 2881–2884, 1992.

[BYJKS04] Z. Bar-Yossef, T. S. Jayram, R. Kumar, and D. Sivakumar. An information statistics approach to data stream and communication complexity. *Journal of Computer and System Sciences*, 68(4):702–732, 2004. Special issue on FOCS 2002.

[CR04] A. Chakrabarti and O. Regev. An optimal randomized cell probe lower bound for approximate nearest neighbor searching. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, pages 473–482, 2004.

[CSWY01] A. Chakrabarti, Y. Shi, A. Wirth, and A. Yao. Informational complexity and the direct sum problem for simultaneous message complexity. In *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science*, pages 270–278, 2001.

[CvDNT98] R. Cleve, Wim 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.

[Gav08] D. Gavinsky. On the role of shared entanglement. *Quantum Information and Computation*, Vol.8 No.1&2:0082–0095, 2008.

[HJMR07] P. Harsh, R. Jain, D. McAllester, and J. Radhakrishnan. The communication complexity of correlation. In *Proceedings of the 22nd Annual IEEE Conference on Computational Complexity*, pages 10–23, 2007.

[Jai06] R. Jain. Communication complexity of remote state preparation with entanglement. *Quantum Information and Computation*, 6 No.4&5:461–464, 2006.

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

[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*, pages 429–438, 2002.

[JRS03a] R. Jain, J. Radhakrishnan, and P. Sen. A direct sum theorem in communication complexity via message compression. In *Proceedings of the 30th International Colloquium on Automata, Languages and Programming*, Lecture Notes in Computer Science, vol. 2719, pages 300–315. Springer-Verlag, 2003. Also cs.CC/0304020.

[JRS03b] R. Jain, J. Radhakrishnan, and P. Sen. A lower bound for the bounded round quantum communication complexity of Set Disjointness. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, pages 220–229, 2003.

[JRS05] R. Jain, J. Radhakrishnan, and P. Sen. Prior entanglement, message compression and privacy in quantum communication. In *Proceedings of the 20th Annual IEEE Conference on Computational Complexity*, pages 285–296, 2005.

[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.

[KN97] E. Kushilevitz and N. Nisan. *Communication Complexity*. Cambridge University Press, Cambridge, UK, 1997.

[Kre95] I. Kremer. Quantum communication. Master's thesis, Hebrew University, Jerusalem, 1995.

[New91] I. Newman. Private vs. common random bits in communication complexity. *Information Processing Letters*, 39(2):67–71, 1991.

[PT06a] M. Patrascu and M. Thorup. Space-time tradeoffs for the predecessor problem. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pages 232–240, 2006.

[PT06b] M. Pătraşcu and M. Thorup. Higher lower bounds for near-neighbor and further rich problems. In *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science*, pages 646–654. IEEE Computer Society Press, Los Alamitos, CA, USA, 2006.

[Sen03] P. Sen. Lower bounds for predecessor searching in the cell probe model. In *Proceedings of the 18th Annual IEEE Conference on Computational Complexity*, pages 73–83, 2003.

[Sha03] R. Shaltiel. Towards proving strong direct product theorems. *Computational Complexity*, 12(1–2):1–22, 2003.

[SV01] P. Sen and S. Venkatesh. Lower bounds in the quantum cell probe model. In *Proceedings of the 28th International Colloquium on Automata, Languages and Programming*, Lecture Notes in Computer Science, vol. 2076, pages 358–369. Springer-Verlag, 2001.

[Uhl76] A. Uhlmann. The 'transition probability' in the state space of a *-algebra. *Rep. Math. Phys.*, 9:273–279, 1976.

[Yao77] A. C-C. Yao. Probabilistic computations: Toward a unified measure of complexity. In *Proceedings of the 18th IEEE Conference on Foundations of Computer Science*, pages 222–227, 1977.